Ergodicity of Probabilistic Cellular Automata: A Constructive Criterion

Christian Maes* and Senya B. Shlosman**

Dipartimento di Matematica di Roma–Tor Vergata, Rome Italy and Department of Mathematics, Rutgers University, New Brunswick, NJ 08903, USA

Abstract. We give a sequence of criteria (of increasing complexity) for the exponential ergodicity of discrete time interacting particle systems. Each criterion involves estimating the dependence on initial conditions of the process on finite space-time volumes. It generalizes and improves the existing single site condition and is the analog of the Dobrushin–Shlosman C_V condition in equilibrium statistical mechanics. Our "dynamic" criterion may also be used to prove the uniqueness of Gibbs state in situations where the C_V condition fails. As a converse we prove that if there is a certain form of convergence to the stationary measure faster than n^{-d}, where n is the time and d is the dimension of the lattice, then our condition holds for some space–time volumes and hence the convergence must be exponentially fast.

1. Introduction

It is well known that finite state space Markov chains have a unique stationary measure whenever there is a strictly positive probability to go from one state to any other in a finite time not depending on the pair of states (irreducible and aperiodic Markov chains), [F, Ge]. The situation is very different for the so-called Probabilistic Cellular Automata (PCA) – Markov chains, for which an infinite collection of cells, each assigned to the vertex of some lattice and taking on a finite number of positions, evolve under a discrete time stochastic dynamics. There exist in fact many examples of PCA with strictly positive transition probabilities which are not ergodic. Even so, it is in general rather easy to give sufficient conditions which ensure the ergodicity of a PCA, see for example [V, LMS, KMPSTV, Fe, S].

For simplicity consider a system of Ising spins $\sigma_i = \pm 1$, associated to the sites of the d-dimensional lattice $\mathbb{Z}^d$. We define a Markov process $\{\sigma_{n,i}, i \in \mathbb{Z}^d, n \geq 0\}$ by giving the *transition probabilities* $\mathrm{Prob}(\sigma_{n,i} = 1$ for all $i \in A$ given that $\sigma_{n-1,\cdot} = \eta) =$

$$P(\sigma_{n,i} = 1, \forall i \in A \,|\, \sigma_{n-1,\cdot} = \eta) = \prod_{i \in A} p_i(\sigma_{n,i} = 1 | \eta), \qquad (1\text{-}1)$$

for $A \subset \mathbb{Z}^d$, finite. The spins are simultaneously and independently updated. It is natural to expect that if $p_i(\cdot|\eta)$ depends sufficiently weakly on η, then the system will forget exponentially fast about the initial state, and will converge to the unique stationary state. One form of this weak dependence is given (in the translation invariant case) by

$$\sum_{j \in \mathbb{Z}^d} \sup_\eta |p_0(1|\eta) - p_0(1|\eta^j)| < 1, \qquad (1\text{-}2)$$

where $(\eta^j)_i = \eta_i$ if $i \neq j$, and $= -\eta_i$ if $i = j$. This condition is sufficient to produce such (ergodic) behavior, [V], [LMS]. Equation (1-2) is the well known $M < \varepsilon$ condition, [L], or "Dobrushin's condition," [D], for interacting particle systems applied to the PCA (1-1).

At this point it is instructive to compare the situation with the one in equilibrium statistical mechanics (ESM), where the question concerns the uniqueness of Gibbs states for some interaction H ([Ge]). The role of the transition probabilities (1-1) is now played by the specifications Q_V, for local functions f,

$$Q_V(f)(\sigma_{V^c}) = \sum_{\sigma_V \in \{-1,1\}^V} f(\sigma_V, \sigma_{V^c}) Q_V(\sigma_V; \sigma_{V^c})$$

with probability kernel

$$Q_V(\sigma_V; \sigma_{V^c}) = \frac{\exp(-H_V(\sigma_V; \sigma_{V^c}))}{Z_V(\sigma_{V^c})}, \qquad (1\text{-}3)$$

for finite $V \subset \mathbb{Z}^d$ and $V^c \equiv \mathbb{Z}^d \setminus V$. This specification is called *self-consistent* because $Q_V \circ Q_W = Q_V$, for $W \subset V$, and a Gibbs state ρ is a probability measure associated to it in the sense that its conditional probabilities are exactly given by the Boltzmann factor (1-3), i.e.

$$\rho \circ Q_V = \rho. \qquad (1\text{-}4)$$

"Dobrushin's single site condition," [D2], in ESM is the equilibrium analog of (1-2) and has often been used to define the high temperature regime of some equilibrium system. This condition is very similar to (1-2): it suffices to replace $p_0(1|\eta)$ by the one-site specification for the corresponding Gibbs state, where now η stands for the configuration *surrounding* the origin. In fact, it is known how to construct PCA for which the Gibbs measures with respect to some interaction are invariant (see also (4-5)) and one can then verify that Dobrushin's ESM condition is identical to Dobrushin's PCA condition. Alternatively, it is easy to see that for any stationary measure of a PCA with strictly positive transition probabilities (1-1) the associated space-time measure (the Markov shift) is a Gibbs measure on $\{-1, +1\}^{\mathbb{Z}^{d+1}}$ in $d+1$ dimensions, [GKLM], with specifications given by

$$Q_V(\sigma_V; \sigma_{V^c}) = \frac{\displaystyle\prod_{x \in V} p_x(\sigma_x | \sigma_{U(x)})}{Z_V(\sigma_{V^c})}, \qquad (1\text{-}5)$$

or Hamiltonian

$$H_V(\sigma_V; \sigma_{V^c}) = -\sum_{x \in \bar{V}} \log p_x(\sigma_x | \sigma_{U(x)}). \qquad (1\text{-}6)$$

Here, $p_x(\sigma_x | \sigma_{U(x)}) = p_i(\sigma_{n,i} | \sigma_{n-1,\cdot})$ for $x = (n, i) \in \mathbb{Z} \times \mathbb{Z}^d$, $\sigma_{U(x)}$ is the configuration of spins in the set $U(x)$ in the *past* of x on which σ_x depends, and $\bar{V} = \{x \in \mathbb{Z}^{d+1} : x \in V$ or $U(x) \cap V \neq \varnothing\}$. Therefore, in the language of ESM, (1-2) is "Dobrushin's single site condition" for the process $\{\sigma_{n,i}, i \in \mathbb{Z}^d, n \in \mathbb{Z}\}$ with the modification that in the supremum only the configuration η in the *past* (and not *all surrounding*) is varied.

The single site condition for ESM was extended and generalized into a constructive criterion by R. L. Dobrushin and S. B. Shlosman, [DS]. The single site is replaced by a finite volume V and the corresponding condition C_V requires that there is a sufficiently good bound on the Vasserstein distance between the conditional probabilities $Q_V(\cdot; \sigma^1)$ and $Q_V(\cdot; \sigma^2)$ for any two boundary configurations σ^1 and σ^2. The DS-criterion then takes the form: if for some finite volume V condition C_V is satisfied, then all the nice properties of high temperature fields hold, see [DS2]. The analog of exponentially weak memory of PCA has here the following form: if condition C_V holds, then there is an $\varepsilon > 0$ and a constant c such that for finite regions $M \subset \mathbb{Z}^d$ large enough and all local f,

$$\sup_{\sigma, \sigma'} |Q_M(f)(\sigma) - Q_M(f)(\sigma')| \leq c |\partial M| \sum_{x \in M} \delta_x f \exp(-\varepsilon \operatorname{dist}(x, \partial M)), \qquad (1\text{-}7)$$

with $\delta_x f = \sup_{\eta} |f(\eta) - f(\eta^x)|$. Formulation (1-7) was given by E. H. Lieb and M. Aizenman.

It is easy to show that this bound implies that there exists at most one Gibbs state, that the correlations decay exponentially fast in this state and that Gibbs states corresponding to small perturbations of the interaction have the same properties. In particular, when applied to the interaction (1-6) for the space-time measure of a PCA, the condition C_V, if it holds, implies the uniqueness of the invariant measure. However, it is interesting to investigate whether here also, as was the case for the single site condition, there exists a corresponding condition similar to C_V but which involves only the dependence of the spins in a certain region of the past. Having to consider only the influence of boundary spins in the past (and not all around) constitutes an important advantage in computer assisted proofs of the uniqueness of the invariant state.

The main purpose of this paper is to present such an extension of the condition (1-2). Our criterion has roughly speaking, the following form:
a) Take some finite set $R \subset \mathbb{Z}^d$ at time zero and choose a time interval $N > 0$.
b) Use the Markov process to construct, for any initial configuration η_R on R, the corresponding space-time measure $\mu_\Lambda^{\eta_R}$ in the appropriate trapezoid-like volume $\Lambda \equiv \Lambda_{N,R}$ with (largest) base equal to R, and with height N in the time direction.
c) If the measures $\mu_\Lambda^{\eta_R}$ for different η_R do not differ too much (in a sense to be defined), then the criterion (which we will call $C_{N,R}$) is satisfied.

This criterion has the following properties:

1. It is *constructive* in the sense that it uses only finite volumes and requires a finite calculation to check it. If it is satisfied, then the corresponding PCA is exponentially ergodic.

2. It is *natural* from the dynamical point of view in the sense that it requires to consider only the conditions in the past (and not at present or in the future).

3. The condition *improves* if the space-time volume (for which the criterion is investigated) is sufficiently enlarged. Moreover the condition is *sensitive* to a change in lattice or dimension (which is not the case for the single site condition (1-2); we will come back to this point in Sect. 5).

4. If the process is exponentially ergodic then, under some additional assumption, the criterion $C_{N,R}$ must be satisfied for some choice of N and R. Therefore our new condition is also close to being *necessary*.

In the next section we introduce some definitions and notation. We explain the construction in which terms we formulate, in Sect. 3, the main theorem. It is shown how the criterion characterizes the high noise regime of a PCA. In particular in Theorem 2 we show that the criterion is very effective in doing so since it turns out to be also a necessary condition for a certain form of rapid convergence; this is then used to derive a proof that if the convergence to the stationary state is faster than n^{-d}, then it must be exponentially fast. Section 4 is devoted to applications. We give examples of how to apply the method to various specific models. We point out that our new (dynamical) condition can be used to prove uniqueness of Gibbs measure for some equilibrium models in situations where the C_V condition fails. Finally Sect. 5 discusses some further aspects of our condition and relates it to other existing criteria.

2. Definitions and Notation

The State Space. We restrict ourselves to the case of spins $\eta_i \in X_i$ taking values in the finite set X_i, a copy of the individual spin space X_o, and associated to the sites i of the d-dimensional lattice $\mathbb{Z}^d$. More generally, (X_o, δ) is a compact metric space equipped with a metric δ. The *state space* ($=$ the set of all possible spin configurations at any given time) is $X \equiv X_o^{\mathbb{Z}^d}$. If f is a function on X, then its *sup-norm* is defined as

$$\| f \| \equiv \sup_{\eta} |f(\eta)| \tag{2-1}$$

and the set of all continuous (with respect to (2-1)) real-valued functions on X is denoted by $C(X)$. For $f \in C(X)$, put

$$\delta_i f = \sup \left\{ \frac{|f(\eta) - f(\eta')|}{\delta(\eta_i, \eta'_i)}, \eta = \eta' \text{ except at } i \right\} \tag{2-2}$$

the *variation* of f and let

$$\||f\|| \equiv \sum_i \delta_i f \tag{2-2'}$$

be its *variational (semi-) norm*. We write $D(X) \equiv \{ f \in C(X) : \||f\|| < \infty \}$ for the set of all "sufficiently local" functions. $\eta, \eta', \ldots \in X$ will be used to denote spin configurations at one given time, and $\eta_A, \eta'_A \in X_A \equiv X_o^A$ their restrictions to the set $A \subset \mathbb{Z}^d$.

The Process. The *process* is a measure μ on the space-time configurations, $\sigma, \sigma', \ldots \in \Omega \equiv X_o^{\mathbb{N} \times \mathbb{Z}^d}$. A general configuration is denoted by $\sigma = \{\sigma_x \equiv \sigma_{n,i}; x \equiv (n,i) \in \mathbb{N} \times \mathbb{Z}^d\}$ and its restriction to the equal time-layer $\mathbb{Z}_n^d \equiv \{n\} \times \mathbb{Z}^d$ is written as $\sigma_{n,\cdot}$, while $\sigma_{n,A}$ indicates a further restriction to the set $A \subset \mathbb{Z}^d$ at time n. More generally, if $B \subset \mathbb{N} \times \mathbb{Z}^d$ is a general space-time volume, then σ_B is the restriction of σ to this set B. The measure μ_B on Ω_B are defined analogously. In particular, μ_n will denote the measure at time n on Ω_n, a copy of the state space X.

We consider discrete time Markov processes (see also [La]) $\mu \equiv \mu^\nu$ indexed by a probability measure ν on Ω_0, which plays the role of initial data,

$$\mu_{\mathbb{Z}_0^d} \equiv \mu_0 = \nu, \qquad (2\text{-}3)$$

and determined by *transition probabilities*

$$p_i(\cdot | \cdot): X_i \times X_{U+i} \to [0,1], \quad i \in \mathbb{Z}^d,$$

i.e. normalized functions of η_i and the spin configuration η'_{U+i} in $U+i$, the translation over lattice vector i of a fixed finite *neighborhood* $U \subset \mathbb{Z}^d$ of the origin, taking on a finite number of different values $\lambda_1, \ldots, \lambda_M, M < \infty$:

$$p_i(\eta_i | \eta'_{U+i}) \in \{\lambda_1, \ldots, \lambda_M\} \subset [0,1], \qquad (2\text{-}4)$$

and

$$\sum_{\eta_i \in X_i} p_i(\eta_i | \eta'_{i+U}) = 1. \qquad (2\text{-}5)$$

We assume translation invariance by requiring that if $\bar{\eta}_0 = \eta_i$ and $\bar{\eta}'_u = \eta'_{u+i}$, for all $u \in U$, then $p_i(\eta_i | \eta'_{U+i}) = p_0(\bar{\eta}_0 | \bar{\eta}'_U)$. For example, if $X_o = \{-1, +1\}$, then

$$p_i(\eta_i | \eta'_{i+U}) \equiv \frac{1}{2}\left(1 + \eta_i \sum_{A \subset U} \lambda_A \prod_{j \in A+i} \eta'_j\right), \qquad (2\text{-}4')$$

is parametrized by the *coupling coefficients* $\{\lambda_A\}$ with $A \subset U$. Analogous expressions can be obtained for other X_o.

For a given configuration $\sigma_{n-1,\cdot}$ of spins at time $n-1$ the conditional probability distribution at time n is given by the measure $p(d\sigma_{n,\cdot} | \sigma_{n-1,\cdot})$ which is the product of the probability measures $p_i(\cdot | \sigma_{n-1,i+U}), i \in \mathbb{Z}^d$. We define the *transition operator* P as

$$Pf(\eta) \equiv \int f(\eta') p(d\eta' | \eta) \qquad (2\text{-}6)$$

on functions $f \in C(X)$ and if ρ is a probability measure on X, then ρP is defined by

$$\int f(\eta) \rho P(d\eta) \equiv \int Pf(\eta) \rho(d\eta), \qquad (2\text{-}7)$$

and the time evolved measure is $\mu_n \equiv \mu_{n-1} P \equiv \nu P^n, n \geq 0$. Obviously, with some abuse of notation, if $f \in C(X)$, then

$$\int Pf(\sigma_{n-1,\cdot}) \mu(d\sigma) = \int f(\sigma_{n,\cdot}) \mu(d\sigma). \qquad (2\text{-}8)$$

(this is not true for more general $f \in C(\Omega)$!)

We say that the measure ρ on X is *invariant* or *stationary* for the dynamics (2-3) iff $\rho P = \rho$. There is always at least one such measure ρ ([L]). The dynamics is said to be *ergodic* iff $\nu P^n \to \rho$ *weakly* as $n \uparrow \infty$ for any initial measure ν. Ergodicity implies having a unique invariant measure but *not* vice versa.

As a final remark: we have implicitly assumed that the Markov chain is time homogeneous. Our results will apply to the more general case where the transition probabilities may also depend on time (and space) in a periodic way.

The Coverings. To state the results we need some geometric objects.

a) *Trapezoids:* Let $\Lambda \subset \mathbb{N} \times \mathbb{Z}^d$ be any finite non-empty subset. A point $(n, i) \in \Lambda$ is called a *base point* if $n = \text{Min}\{m: (m, j) \in \Lambda\}$. $B(\Lambda)$ is the set of all base points of Λ. We call its complement the set of *inner points* denoted by $I(\Lambda) \equiv \Lambda \backslash B(\Lambda)$. For each point $x = (n, i) \in \mathbb{N} \times \mathbb{Z}^d$, $n > 0$ we write $U(x) \equiv \{n - 1\} \times (U + i)$ (see (2-4)) for the set of all relevant points for x.

Definition 1. Λ *is called a trapezoid, if*

i) *for each* $x \in I(\Lambda)$ $U(x) \subset \Lambda$,
ii) *for each* $x \in B(\Lambda)$ $x \in U(y)$ *for some* $y \in I(\Lambda)$.

The sets $\Lambda_x \equiv \{x\} \cup U(x), x = (n, i), n > 0$, are the simplest trapezoids from which all others are built, i.e. if Λ is a trapezoid, then $\Lambda = \bigcup_{x \in I(\Lambda)} \Lambda_x$. We call the Λ_x *unit trapezoids.*

From (2-8), if Λ is a trapezoid, then given the configuration $\bar{\sigma}$ on $B(\Lambda)$, the conditional probabilities $q_\Lambda(\sigma_{I(\Lambda)}; \bar{\sigma})$ on $I(\Lambda)$ are

$$q_\Lambda(\sigma_{I(\Lambda)}; \bar{\sigma}) = \prod_{x \in I(\Lambda)} p_x(\sigma_x | \sigma_{U(x)})]_{\sigma = \bar{\sigma} \text{ on } B(\Lambda)} \qquad (2\text{-}9)$$

with $p_x(\sigma_x | \sigma_{U(x)}) \equiv p_i(\sigma_{n,i} | \sigma_{n-1, U+i})$ for $x = (n, i)$, the conditional probability of σ_x given the configuration in $U(x)$.

b) *Admissible Coverings:* Let $N \geq 1$ be a fixed integer. Let $R \subset \mathbb{Z}^d$ be a finite region containing the origin for which there exists a tiling of the lattice, i.e. associated to the *fundamental cell* R there is a sublattice $\mathbb{L}$ of $\mathbb{Z}^d$ of finite index such that each site $j \in \mathbb{Z}^d$ belongs to $R(i) \equiv R + i$ for exactly one $i \in \mathbb{L}$. Choose $\Gamma_{R,N} \equiv \{\Lambda_\alpha\}_{\alpha \in J}$ a finite collection of trapezoids contained in the strip $S_N \equiv [0, N] \times \mathbb{Z}^d$ such that for any $n, 0 \leq n \leq N$, there exists $i(n) \in \mathbb{Z}^d$ for which $\bigcup_{n=0}^{N} \{(n, j + i(n)): j \in R\} \subset \bigcup_\alpha \Lambda_\alpha$.

Definition 2. *We say that* $\Gamma_{R,N}$ *generates an admissible covering* Π *of* S_N *with*
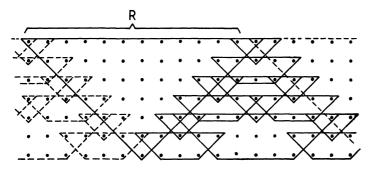

Fig. 1. $N = 6$; $U = \{-1, 0, 1\}$; $d = 1$; $i(n) = n$; time is downward

fundamental cell R if $\Pi = \{\Lambda_\alpha(i)\}_{i\in\mathbb{L}, \alpha\in J}$, $\Lambda_\alpha(i) \equiv \Lambda_a + (0,i) \subset S_N$ *is a covering of* S_N *such that*

iii) *each point* $x\in S_N\backslash\mathbb{Z}_0^d$ *is an inner point of exactly one trapezoid* $\Lambda_\alpha(i)\in\Pi$.

We will write q_α, I_α and B_α instead of $q_{\Lambda_\alpha}, I(\Lambda_\alpha)$ and $B(\Lambda_\alpha)$. If $f\in C(I_\alpha)$ is a function on I_α, then

$$q_\alpha(f)(\eta) \equiv \sum_\sigma f(\sigma)q_\alpha(\sigma;\eta). \qquad (2\text{-}10)$$

The simplest example of admissible covering in all dimensions and for all neighborhoods U is $\Pi = \{\Lambda_{(1,i)}, i\in\mathbb{Z}^d\}$ with $N = 1$, $R = \{0\}$ and $\Gamma_{R,N} = \{\Lambda_{(1,0)}\}$. Our construction is dimension- and neighborhood-dependent and can be thought of as a generalization of this example. See Fig. 1.

3. Main Result

The Dependence Matrix. We suppose that for any $\alpha\in J$ there is given a matrix $k_{x,y}^\alpha, x\in I_\alpha, y\in B_\alpha$. Its purpose will be to measure the dependence of the measure $q_\alpha(\sigma_{I_\alpha}; \bar\sigma_{B_\alpha})$ on $\bar\sigma_{B_\alpha}$, see further (3-3).

Let $\Lambda_\alpha(i)\in\Pi$ be any trapezoid in the admissible covering Π generated by $\Gamma_{R,N}$. For any $x\in I(\Lambda_\alpha(i))$ we draw *arrows* l from x to all $y\in B(\Lambda_\alpha(i))$ and associate with the arrow $l:x\to y$ from x to $l(x) \equiv y$ the value

$$k_l \equiv k_{x-(0,i), y-(0,i)}^\alpha. \qquad (3\text{-}1)$$

By property iii) (definition of admissible covering) the trapezoid $\Lambda_\alpha(i)$ with $x\in I(\Lambda_\alpha(i))$ is uniquely defined given $x\in S_N\backslash\mathbb{Z}_0^d$. Let $x,y\in S_N$. An *oriented path* $\omega:x\to y:x_0\to x_1\to\cdots\to x_m$, is a sequence of arrows $l:x_k\to x_{k+1}$ with $x_{k+1}\in B(\Lambda_\alpha(i))$ if $x_k\in I(\Lambda_\alpha(i))$, $0\leq k\leq m-1$ starting at $x_0 = x$ and ending at $x_m = y$. We assign to it the value

$$k_\omega \equiv \prod_{l\in\omega} k_l. \qquad (3\text{-}1')$$

Finally, we define the *dependence matrix*

$$K_{i,j} \equiv \sum_{\omega:(N,i)\to(0,j)} k_\omega, \quad i,j\in\mathbb{Z}^d, \qquad (3\text{-}2)$$

where the (finite) sum is over all oriented paths ω from (N,i) at time N to $(0,j)$ at time 0. The dependence matrix $K_{i,j} = K_{i+l,j+l}, l\in\mathbb{L}$, is periodic with the period of the lattice $\mathbb{L}$ and, for fixed $i\in\mathbb{Z}^d$, is non-zero for only a finite number of $j\in\mathbb{Z}^d$.

Constructive Criterion. We say that the process defined by (2-4)–(2-8) satisfies condition $C_{N,R}$ if the following three properties hold:

1. There exists an admissible covering $\{\Lambda_\alpha(i)\}$ of the strip S_N with fundamental cell R.
2. There exist numbers $k_{x,y}^\alpha, x\in I_\alpha, y\in B_\alpha, \alpha\in J$ such that for any two configurations $\bar\sigma^1, \bar\sigma^2$ in B_α there exists a coupling $P_\alpha^{\bar\sigma^1,\bar\sigma^2}$ for the conditional measures $q_\alpha(\cdot,\bar\sigma^1), q_\alpha(\cdot;\bar\sigma^2)$ on I_α such that, for all $x\in I_\alpha$,

$$\sum_{\sigma^1,\sigma^2} \delta(\sigma_x^1,\sigma_x^2)P_\alpha^{\bar\sigma^1,\bar\sigma^2}(\sigma^1,\sigma^2) \leq \sum_{y\in B(\Lambda_\alpha)} k_{x,y}^\alpha\delta(\bar\sigma_y^1,\bar\sigma_y^2). \qquad (3\text{-}3)$$

3. The dependence matrix K defined through the $\{k_{x,y}^\alpha\}$ via (3-1)–(3-2) satisfies the inequality

$$\gamma \equiv \frac{1}{|R|} \sum_{i \in R} \sum_{j \in \mathbb{Z}^d} K_{i,j} < 1. \qquad (3\text{-}4)$$

As a reminder: the measure $P(\cdot,\cdot)$ on a space $Y \times Y$ is called a *coupling* for the measures q^1, q^2 on Y if for any measurable $A \subset Y, P(A, Y) = q^1(A)$ and $P(Y, A) = q^2(A)$.

For the simplest admissible covering $\Lambda_\alpha(i) = \Lambda_x, x = (1, i)$, for example in the case $X_o = \{-1, +1\}$ of Ising spins, we can choose $k_{x,y}^{(1,0)} = \sup_\eta |p_i(1|\eta) - p_i(1|\eta^j)|$ if $x = (1, i), y = (0, j)$. Note that here (3-4) reduces to the single site condition (1-2).

Main Theorem. *Condition $C_{N,R}$ implies that the process defined in (2-4)–(2-8) is ergodic and that for all $f \in D(X)$, $0 < a, 0 \leq b < N$,*

$$\|P^{aN+b}f - \int f(\eta)\rho(d\eta)\| \leq \||P^b f\|| \frac{\gamma^a}{1 - \gamma}, \qquad (3\text{-}5)$$

where ρ is the unique invariant measure.

Proof. The crucial observation is that the process can be reconstructed from the conditional measures q_α alone. Call *first generation* trapezoids $\Lambda_\alpha^1(i)$ those trapezoids of the covering having their base points

$$B(\Lambda_\alpha^1(i)) \subset \mathbb{Z}_0^d$$

on the initial time layer. Given the trapezoids up to the *n-th generation*, $n \geq 1$, we define the $(n+1)th$ *generation* trapezoids $\{\Lambda_\alpha^{n+1}(i)\}$ of the covering by the condition

$$B(\Lambda_\alpha^{n+1}(i)) \subset \bigcup_{j \in \mathbb{L}} \bigcup_{k=1} I(\Lambda_\alpha^k(j)). \qquad (3\text{-}6)$$

In this way, with a finite number of generations, one has classified all of the trapezoids covering the strip S_N. Moreover, given the spin configuration σ in the trapezoids up to the nth generation, the measure induced by the process on the $(n + 1)^{\text{th}}$ generation is a product over the trapezoids it contains and, inside each of such trapezoid Λ, is of the form $q_\Lambda(d\sigma_{I(\Lambda)}|\sigma_{B(\Lambda)})$ with probabilities as in (2-9). Therefore, we obtain a coupling $P^{\eta,\eta'}$ for the processes $\mu^\eta, \mu^{\eta'}$ with initial configurations η, η' by taking products of translates of the couplings $P_\alpha^{\bar\sigma^1, \bar\sigma^2}$ of (3-3). This implies that, writing

$$\delta_x P^{\eta,\eta'} \equiv \int \delta(\sigma_x^1, \sigma_x^2) P^{\eta,\eta'}(d\sigma^1, d\sigma^2),$$

for $x \in S_N \setminus \mathbb{Z}_0^d$, we have

$$\delta_x P^{\eta,\eta'} \leq \sum_{l:x \to \cdot} k_l \delta_{l(x)} P^{\eta,\eta'},$$

where the sum is over all arrows l going out from x. If $l(x) \in S_N \setminus \mathbb{Z}_0^d$, then we can again apply the same inequality and continue doing so until we reach as endpoint of an arrow an element y of the initial time layer $\mathbb{Z}_0^d$. Take now $\eta = \eta'$ except on $j \in \mathbb{Z}^d$. Of course, if $y = (0, k)$, then

$$\delta_y P^{\eta,\eta'} = \delta_{j,k} \delta(\eta_j, \eta_j').$$

Hence, if $x = (N, i) \in \mathbb{Z}_N^d$, then

$$\delta_x P^{\eta, \eta'} \leqq K_{i,j} \delta(\eta_j, \eta_j').$$

Fix $g \in D(X)$. The dependence of the time evolved function $P^N g$ on the spin at site j is now estimated by

$$\delta_j(P^N g) \leqq \sum_{i \in \mathbb{Z}^d} \delta_i g K_{i,j}. \tag{3-7}$$

We can repeat the derivation of (3-7) for different translates of the fundamental cell R (it does not matter which point in R we call the origin). Therefore, using the periodicity of the dependence matrix K,

$$\delta_j(P^N g) \leqq \|g\| \frac{1}{|R|} \sum_{i \in R} K_{i,j}, \tag{3-8}$$

and we obtain the $\| \cdot \|$-contraction of P^N by summing (3-8) over j:

$$\|P^N g\| \leqq \gamma \|g\|. \tag{3-9}$$

The rest of the proof is standard and we only sketch the main steps referring the reader to [L, LMS or M] for further details. We have from (3-9) that, for any $a' > a > 0, b \geqq 0$,

$$\| P^{a'N+b} f - P^{aN+b} f \| = \left\| \sum_{n=a}^{a'-1} (P^N - 1) P^{nN+b} f \right\| \leqq \sum_{n=a}^{a'-1} \|P^{nN+b} f\|$$

$$\leqq \|P^b f\| \frac{\gamma^a}{1-\gamma}. \tag{3-10}$$

Hence, $\lim_{n \to \infty} P^n f$ exists and must be constant. It follows that for all $f \in D(X)$

$$\lim_n P^n f = \int f(\eta) \rho(d\eta) \tag{3-11}$$

for any invariant measure ρ. But the space $D(X)$ is dense in $C(X)$. ∎

The High Noise Regime. The results above allow one to derive a characterization of the process in the so-called *high noise regime* and we therefore say that the criterion $C_{N,R}$ defines this region in the parameter space of coefficients λ_m appearing in the transition probabilities (2-4). One of the most important implications from having exponential decay in time is

Corollary 1 (Exponential Decay of the Space-Correlations). *Let f, g be any two functions depending on the configuration of spins in the finite regions A, respectively $B \subset \mathbb{Z}^d$, a distance $d(A, B) = vN$ apart. Condition $C_{N,R}$ implies that in the invariant measure ρ*

$$|\int f g(\eta) \rho(d\eta) - \int f(\eta) \rho(d\eta) \int g(\eta) \rho(d\eta)| \leqq 3 \|f\| \|g\| \frac{\gamma^{v/2r}}{1-\gamma}, \tag{3-12}$$

where the integer r is such that $p_0(\cdot | \eta)$ does not depend on any η_i for which $|i| \equiv d(0, i) \geqq r$.

Idea of Proof. At any fixed time, spins in the regions A and B communicate with each other only via events which happened at least at time $vN/2r$ in the past. But the dependence on this history is exponentially small (with factor $\gamma^{v/2r}$). The actual proof is very similar to the one of Theorem 2 in [LMS]. ∎

Condition $C_{N,R}$ can be viewed as prescribing a region in the parameter space $\{\lambda_1, \ldots, \lambda_M\}$ (or, for (2-4'), in the space of coupling coefficients $\{\lambda_A\}_{A \subset U}$), in which the corresponding process is ergodic. In that sense, the set of Markov processes – of finite range – satisfying $C_{N,R}$ is an open set. Hence, small perturbations of the process will also fall into the high noise regime.

Effectiveness. We ask here whether our criterion is also *necessary* to have fast convergence to the unique invariant measure.

Definition 3. *The process defined in (2-4)–(2-8) is uniformly exponentially ergodic (UEE) if there is a probability measure ρ and a constant $\lambda > 0$ such that for n large enough and all local functions f (having finite support)*

$$\| P^n f - \int f(\eta)\rho(d\eta) \| \leq c(f)e^{-\lambda n}, \tag{3-13}$$

for some constant $c(f) < \infty$.

Clearly, (3-5) says that the condition $C_{N,R}$ implies UEE with constants $c(f) \equiv 1/(1-\gamma)\sup_{0 \leq b < N} \||P^b f\|| $ and $\lambda \equiv -1/N \log \gamma$, but we also have a converse:

Theorem 2. *If the process is* UEE *with $c(f) \leq c\||f\||$ in (3-13), then for some N and R the condition $C_{N,R}$ holds.*

Proof. We can restrict ourselves to the case where U is a cube centered around the origin of sidelength $2r$ (r as in (3-12)). For L, N integers such that $L > 2Nr$, let $T \subset R \subset \mathbb{Z}_0^d$ be two cubes centered around the origin of sidelengths $L - 2Nr$ and L. We consider the trapezoid Λ with base $B(\Lambda) = R$ and top $T(\Lambda) = T_N = T + (N, 0)$, the set of points $(N, i) \in \Lambda$ with maximal time coordinate. To construct a covering of S_N with fundamental cell R, we take this large trapezoid Λ together with its translates $\Lambda(i) \equiv \Lambda + Li, i \in \mathbb{Z}^d$, and add unit trapezoids Λ_x for each point in the strip not covered by the large ones. The coupling P_x in the unit trapezoids can be chosen more or less arbitrary except for the fact that $P_x^{\sigma^1_{U(x)}, \sigma^2_{U(x)}}(\sigma_x^1 = \sigma_x^2) = 1$ whenever $\sigma^1_{U(x)} = \sigma^2_{U(x)}$. For the points $x \in I_\alpha$ *not* belonging to the top $T(\Lambda(i)), i \in \mathbb{Z}^d$, of any of the large trapezoids, we consider the smallest trapezoid $M(x) \subset \Lambda(i)$ which contains x and which has its base $B(x) \subset B_\alpha$, and put $k_{x,y}^\alpha = 1, y \in B(x), = 0$, otherwise. To find a good coupling in the large trapezoids we consider for any two initial configurations η^1 and η^2 the restrictions $q_\Lambda(\sigma_{T(\Lambda)}; \eta^1_{B(\Lambda)})$ and $q_\Lambda(\sigma_{T(\Lambda)}; \eta^2_{B(\Lambda)})$ on $T(\Lambda)$ of the conditional measures (2-9). By hypothesis, for N sufficiently large,

$$\sup_{f \in L^1(T(\Lambda))} |q_\Lambda(f)(\eta^1) - q_\Lambda(f)(\eta^2)| \leq 2c|T(\Lambda)|e^{-\lambda N}, \tag{3-14}$$

where $L^1(T(\Lambda)) = \{f \in C(\Omega_{T(\Lambda)}) : \delta_x f \leq 1, \text{ for all } x \in T(\Lambda)\}$. Equation (3-14) is the Vasserstein distance between the two conditional measures on $T(\Lambda)$ ([Du]) and therefore, there exists a coupling $P_\Lambda^{\eta^1, \eta^2}$ on Λ (which is the extension of a coupling on $T(\Lambda)$) for which

$$\sum_{x \in T(\Lambda)} \delta_x P_\Lambda^{\eta^1, \eta^2} \leq 2c|T(\Lambda)|e^{-\lambda N}. \tag{3-15}$$

Hence, by taking for $x \in T(\Lambda)$, $y \in B(\Lambda)$,

$$k_{x,y}^{\Lambda} \equiv \sup\{\delta_x P_{\Lambda}^{\eta^1, \eta^2}, \eta^1 = \eta^2 \text{ except on } y\},$$

we get

$$\sum_{j \in \mathbb{Z}^d} \sum_{i \in T} K_{i,j} = \sum_{y \in R} \sum_{x \in T(\Lambda)} k_{x,y}^{\Lambda} \leqq 2c \, |T(\Lambda)| e^{-\lambda N} (Nr)^d. \tag{3-16}$$

On the other hand, for all $(N, i) \in R_N \setminus T_N$, we have the bound $K_{i,j} \leqq W_{i,j}$ the number of oriented paths from (N, i) to $(0, j)$ via arrows l with $k_l \neq 0$. Clearly, $\sum_{j \in \mathbb{Z}^d} W_{ij} \leqq (2r)^{dN}$ and therefore

$$\gamma \leqq 2c \frac{(L - Nr)^d}{L^d} e^{-\lambda N} (Nr)^d + (2r)^{dN} \frac{L^d - (L - Nr)^d}{L^d}. \tag{3-17}$$

Choosing $L = 2(2r)^{dN} Nr$ and letting N become very large, we get $\gamma < 1$ as desired. ∎

The assumption in Theorem 2 that $c(f) \leqq c \, ||| f |||$ doesn't seem to be just technical but there are cases in which it is automatically satisfied. That happens for the so-called attractive Ising spin systems, i.e. monotone processes in which the spins can only take the two values ± 1, see [AH and KRS].

Corollary 2. *If the process is UEE and attractive with spin space $X_o = \{-1, +1\}$, then for some N and R the condition $C_{N,R}$ is verified.* ∎

Furthermore, inspection of the estimate in (3-17) shows that we don't really need an exponential convergence: we just have to beat the N^d factor. This leads us to the following

Corollary 3. *If for any two initial configurations $\eta^1 = \eta^2$ except in one point, and for all finite cubes $T \subset \mathbb{Z}^d$,*

$$\frac{1}{|T|} R_T(\mu_{n,T}^{\eta^1}, \mu_{n,T}^{\eta^2}) = o(n^{-d}), \tag{3-18}$$

where R_T is the Vasserstein distance for two measures on T, then, condition $C_{N,R}$ is satisfied for some choice of N and R and the process is UEE. ∎

Remark. In particular, if we consider an attractive Ising spin system, then we have here a discrete time version of the results in [H, AH and HS] on the possible rates of convergence to the stationary measure. More precisely, if for the function $f(\eta) \equiv \frac{1}{2}\eta_o$, $P^n f(+) - P^n f(-) = o(n^{-d})$ with $\pm$ the all plus or all minus configuration, then condition $C_{N,R}$ holds for some N, R and the process is UEE.

4. Applications

Applying criterion $C_{N,R}$ to specific models gives rise to at least three practical problems. The first is geometric in nature: finding a covering. In general, there is a lot of choice – as a rule, the larger N and R the better, but also, the harder it becomes to deal with the next two problems. The second is probabilistic: we have to find the couplings P_α and the corresponding numbers $k_{x,y}^\alpha$ satisfying (3-4). There is an immediate candidate for the coupling P_Λ on a trapezoid Λ suggested by

formula (2-9). We can take the product of couplings P_x on the unit trapezoids, i.e. if $\sigma^1(\sigma^2)$ is a configuration in $I(\Lambda)$ being equal to $\bar\sigma^1(\bar\sigma^2)$ on $B(\Lambda)$, then

$$P_\Lambda^{\bar\sigma^1,\bar\sigma^2}(\sigma^1,\sigma^2) = \prod_{x\in I(\Lambda)} P_x^{\sigma^1_{U(x)},\sigma^2_{U(x)}}(\sigma^1_x,\sigma^2_x). \tag{4-1}$$

If $X_o = \{-1,+1\}$, and the dynamics is attractive (that is, $p_0(1|\eta_U)$ is non-decreasing in η, see [LMS]), then using Holley's coupling $P_\Lambda^{\bar\sigma,\bar\sigma^y}$ (see [KRS] or [L]) for which

$$P_\Lambda^{\bar\sigma,\bar\sigma^y}(\sigma^1_\Lambda \geqq \sigma^2_\Lambda) = 1, \quad \text{if} \quad \bar\sigma > \bar\sigma^y,$$

we have

$$P_\Lambda^{\sigma,\sigma^y}(\sigma^1_x \neq \sigma^2_x) = P_\Lambda^{\bar\sigma,\bar\sigma^y}(1,-1) = q_\Lambda^{\bar\sigma}(1) - q_\Lambda^{\bar\sigma^y}(1),$$

and we can choose the numbers

$$k_{x,y}^\alpha \equiv \sup_\sigma |q_\alpha(\sigma_x = 1;\bar\sigma) - q_\alpha(\sigma_x = 1;\bar\sigma^y)|. \tag{4-2}$$

Finally, the third problem may involve some combinatorics: we have to compute the dependence matrix and (at least an upper bound to) the corresponding γ.

For the sake of simplicity, both examples below are for attractive dynamics with individual spin space $X_o = \{-1,+1\}$ so that we can use formulae (4-1) and (4-2). The third example is related to the problem of uniqueness of Gibbs measure and uses (4-1).

Examples.

Example 1: Stavskaya's Model. Consider the one-dimensional PCA on $\{-1,+1\}^{\mathbb{Z}}$ with neighborhood $U = \{0,-1\}$, defined by

$$\begin{aligned} p_0(1|\eta) &= 1 \quad \text{if} \quad \eta_o = \eta_{-1} = 1, \\ &= \varepsilon \quad \text{else,} \end{aligned} \tag{4-3}$$

where $0 \leq \varepsilon \leq 1$ is a parameter. There exists a critical ε_c for which the process is ergodic if $\varepsilon > \varepsilon_c$ and has another invariant state, besides the trap $\eta \equiv 1$, if $\varepsilon < \varepsilon_c$. The single site condition (1-2) gives the bound $\varepsilon_c \leqq \frac{1}{2}$. The simplest illustration of our method is to take $N = 2$, $R = \{0,1\}$, $\mathbb{L} = 2\mathbb{Z}$, $\Gamma_{N,R} = \{\Lambda_1,\Lambda_2\}$ with

$$\Lambda_1 = \{(2,0),(1,-1),(1,0),(0,-2),(0,-1),(0,0)\},$$

for the first generation, and

$$\Lambda_2 = \{(2,1),(1,0),(1,1)\}$$

for the second generation trapezoids. It is easy to check that this choice generates an admissible covering of the strip S_2, see Fig. 2.

We choose the numbers $k_{x,y}^\alpha$ according to (4-2),

$$\begin{aligned} k_{(2,0),(0,-2)}^1 &= k_{(2,0),(0,0)}^1 = (1-\varepsilon)^2, \\ k_{(2,0),(0,-1)}^1 &= (1-\varepsilon)^2(1+\varepsilon), \\ k_{(1,0),(0,-1)}^1 &= k_{(1,0),(0,0)}^1 = k_{(1,-1),(0,-2)}^1 = k_{(1,-1),(0,-1)}^1 = 1-\varepsilon, \\ k_{(2,1),(1,0)}^2 &= k_{(2,1),(1,1)}^2 = 1-\varepsilon, \end{aligned}$$

so that condition (3-4) is satisfied. We then get the dependence matrix K determined by $K_{0,-2} = K_{0,0} = K_{1,-1} = K_{1,1} = (1-\varepsilon)^2$, $K_{0,-1} = (1-\varepsilon)^2(1+\varepsilon)$, $K_{1,0} = 2(1-\varepsilon)^2$,
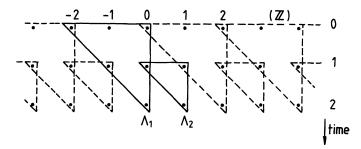
Fig. 2. Stavskaya's model

whose sum is equal to $(1-\varepsilon)^2(7+\varepsilon)$ and condition (3-3) becomes $\gamma = \frac{1}{2}(1-\varepsilon)^2(7+\varepsilon) <$ 1 from which the bound $\varepsilon_c \leqq 0.484$. It is straightforward to find better bounds using another covering but we will not pursue this issue here further given that for this model a very good bound $\varepsilon_c \leqq 0.323$ (numerically $\varepsilon_c \approx 0.31$) already exists, see [KMPSTV].

Example 2: Toom's Model. For Ising spins on the square lattice $\mathbb{Z}^2$, choose the transition probabilities

$$p_0(1|\eta) = 1 - \varepsilon \quad \text{if } \text{sign}(\eta_{(0,0)} + \eta_{(0,1)} + \eta_{(1,0)}) = 1,$$
$$= \varepsilon \qquad \text{else}, \tag{4-4}$$

$0 \leqq \varepsilon \leqq \frac{1}{2}$. It was proven by Toom, [T], that this two dimensional majority vote model with neighborhood $U = \{(0,0),(0,1),(1,0)\}$ has a phase transition in the sense that for ε small ($\varepsilon < \varepsilon_c$) the all $+$ and all $-$ state (which, obviously, are invariant states for $\varepsilon = 0$) are both stable (yielding two different invariant states). For $\varepsilon > \varepsilon_c$, there is a unique invariant state. The single site condition (1-2) gives $\varepsilon_c \leqq \frac{1}{3}$. To apply our criterion, we choose $N = 2$, $R = \{(0,0),(1,0),(0,1),(1,1)\}$ and generating trapezoids

$$\Lambda_1 = \{(2,0,0),(1,1,0),(1,0,1),(1,0,0),(0,2,0),$$
$$(0,1,1),(0,0,2),(0,1,0),(0,0,1),(0,0,0)\},$$
$$\Lambda_2 = \{(1,1,1),(0,1,1),(0,2,1),(0,1,2)\}$$

for the first generation, and

$$\Lambda_3 = \{(2,0,1),(1,0,1),(1,1,1),(1,0,2)\},$$
$$\Lambda_4 = \{(2,1,0),(1,1,0),(1,1,1),(1,2,0)\},$$
$$\Lambda_5 = \{(2,1,1),(1,1,1),(1,2,1),(1,1,2)\}$$

for the second generation.

The numbers $k_{x,y}^\alpha$ can be computed from (4-2), with $m \equiv 1 - 2\varepsilon$,

$$k_{(2,0,0),(0,2,0)}^1 = k_{(2,0,0),(0,0,2)}^1 = k_{(2,0,0),(0,0,0)}^1 = \frac{m^2}{2}(1+m^2),$$

$$k_{(2,0,0),(0,1,0)}^1 = k_{(2,0,0),(0,0,1)}^1 = m^2, k_{(2,0,0),(0,1,1)}^1 = \frac{m^2}{2}(3-m^2),$$

$$k_{(n,i,j),(n-1,i,j)}^\alpha = k_{(n,i,j),(n-1,i+1,j)}^\alpha = k_{(n,i,j),(n-1,i,j+1)}^\alpha = m \quad \text{for} \quad \alpha = 2,3,4,5$$

and satisfy (3-4). They give rise to a dependence matrix K for which $\gamma = 8m^2 + \frac{1}{4}m^4$. Hence, the bound $\varepsilon_c \leqq 0.324$.

Uniqueness of Gibbs State. An immediate application to ESM is to consider PCA for which all Gibbs states ρ_H with respect to some translation invariant finite range Hamiltonian H are invariant measures. This, in general, can be easily done by slightly changing the updating mechanism and to construct a PCA-like dynamics which is reversible (or, satisfies the condition of detailed balance) for the Gibbs measure ρ_H. One way to do this is to take a periodic (in the time direction) stacking of lattices, for example, a partition $\{S_q\}_{q=1}^{T}$ of $\mathbb{Z}^d$ into sublattices, and to apply successively a simultaneous updating on all spins of the sublattices S_q. Examples and details can be found in [LSM, LSM2, KMPSTV]. Equivalently, we can make the updating mechanism periodic in time, that is, at any given time n, only the spins in a given sublattice $S_{q(n)}$ with $q(n) = q(n + T)$ periodic, can change while the others spins remain unchanged with probability one. The partition $\{S_q\}$ of $\mathbb{Z}^d$ has to be such that, for all q and all $i \in S_q$ the conditional probabilities $Q_i(\eta_i = a; \eta_{\mathbb{Z}^d \setminus i})$, $a \in X_i$ effectively depend on the spin configuration in $\mathbb{Z}^d \setminus S_q$ only. It is then easy to check from the definition (1-4) that choosing at time n the transition probabilities,

$$p_{n,i}(\eta_i | \eta'_{U+i}) = Q_i(\eta_i; \eta_{\mathbb{Z}^d \setminus i} = \eta') \quad \text{if} \quad i \in S_{q(n)}$$
$$= \delta_{\eta_i, \eta'_i} \qquad \qquad \text{else,} \qquad (4\text{-}5)$$

leaves the corresponding Gibbs measures invariant (here, $\delta_{p,q}$ is one if $p = q$ and zero otherwise). The (appropriately rephrased) criterion $C_{N,R}$ then implies that there can only be one such Gibbs measure. As mentioned in the Introduction of this paper, [DS] have derived a constructive criterion C_V ($V \subset \mathbb{Z}^d$ finite) for uniqueness of Gibbs measures in terms of the self-consistent specifications (1-3). For every $j \in V^c$ and $\bar{\eta} \in X_{V^c}$, consider the pair of conditional distributions $Q_V(\cdot; \bar{\eta} \times \eta_j^1)$ and $Q_V(\cdot; \bar{\eta} \times \eta_j^2)$, where $\bar{\eta} \times \eta_j^\alpha = \bar{\eta}$ except at j where it is equal to $\eta_j^\alpha, \alpha = 1, 2$. Let $R_V(\bar{\eta} \times \eta_j^1, \bar{\eta} \times \eta_j^2)$ be the Vasserstein distance between these two conditional measures (see [DS]) and define

$$k_j^V \equiv \sup_{\bar{\eta}, \eta_j^1 \neq \eta_j^2} R_V(\bar{\eta} \times \eta_j^1, \bar{\eta} \times \eta_j^2). \qquad (4\text{-}6)$$

Condition C_V holds if

$$\frac{1}{|V|} \sum_{j \in V^c} k_j^V < 1. \qquad (4\text{-}7)$$

Comparing their condition with ours, it is particularly interesting to see if there are systems for which C_V fails for any arbitrary volume V, while $C_{N,R}$ works, thus implying uniqueness of Gibbs measures. Such an example in fact exists.

Example 3: The Čhech Model (see [Sh], [Sh2]). The single spin space is $X_o = \{0, 1, 2, 3\}$. The interaction is defined by a self-energy

$$h(a) = \begin{cases} 0 & a = 0, \\ 3^d & a = 1, 2, \\ 10^d & a = 3; \end{cases} \qquad (4\text{-}8a)$$

and a pair interaction

$$v(a, b) = v(b, a) = \begin{cases} 0 & a = b, \\ 1 & a = 1, 2, b = 0, 3, \\ 5^d & a = 0, b = 3 \quad \text{or} \quad a = 1, \quad b = 2. \end{cases} \tag{4-8b}$$

Let $A_L = \{(i_1, \ldots, i_d) \in \mathbb{Z}^d : |i_\alpha| \leq l, \alpha = 1, \ldots, d\}$ be a cube of side length $2L$, $L = 1, 2, \ldots$. The energy of the configuration $\eta_{A_L} \in X_{A_L}$ in the box A_L for given boundary condition $\eta_{A_L^c}$ on its complement, is

$$H_{A_L}(\eta_{A_L} | \eta_{A_L^c}) = \sum_{i \in A_L} h(\eta_i) + \sum_{\|i - j\| = 1} v(\eta_i, \eta_j), \tag{4-9}$$

where $\|i\| \equiv \max\{|i_\alpha|, \alpha = 1, \ldots, d\}$ for $i = (i_1, \ldots, i_d) \in \mathbb{Z}^d$. Consider the following two configurations on $A_L^c : \eta_{A_L^c} \equiv 3$ (all spins outside equal 3) and $\eta_{A_L^c}^a \equiv 3$ except at one given point j with $\|j\| = L + 1$, where it is equal to $a, a = 1, 2$. The characteristic feature of this Chech model can be discovered by comparing the groundstates in A_L for these two boundary conditions. In the first case (all spins outside A_L equal to 3), the energy (4-9) is minimized by the two configurations $\eta^b, b = 1, 2$,

$$\eta_i^b \equiv \begin{cases} 0 & \text{if} \quad \|i\| < L, \\ b & \text{if} \quad \|i\| = L. \end{cases} \tag{4-10}$$

In the second case only the configuration with $b = a$ minimizes the energy. In other words, modifying the boundary condition in just one point has a rather drastic effect: the low temperature conditional distribution changes along the whole boundary of A_L. Therefore, (4-6) is of order L^{d-1} and the left-hand side of (4-7) becomes proportional to L^{d-2} and is always larger than 1 if $d \geqslant 3$. Similar considerations can be made for other volumes V instead of A_L, so that condition C_V of (4-7) is never satisfied. There exist other types of Chech models for which the [DS] criterion works by choosing the appropriate volume V but for (4-9) the uniqueness of the low temperature Gibbs states had to be proven separately, see [Sh]. The last statement also follows from our criterion $C_{N,R}$ as we will now show.

The dynamics is of the type (4-5) with neighborhood $U = \{i \in \mathbb{Z}^d : \|i\| = 1\}$ and sublattices

$$S_q \equiv 2\mathbb{Z}^d + e_q, \quad q = 1, \ldots, T = 2^d, \tag{4.11}$$

where the $\{e_q\}_1^T \equiv \{i \in \mathbb{Z}^d : i_\alpha = 0, 1; \alpha = 1, \ldots, d\}$ in arbitrary order. Let $q(n) = n \bmod T$, $n > 0$. At time n all spins in the sublattice $S_{q(n)}$ are simultaneously updated, keeping the other fixed. The transition probabilities are obtained by combining (4-5), (4-9) and (1-3):

$$p_{n,i}^\beta(\eta_i | \eta_{U+i}') = \frac{\exp - \beta [h(\eta_i) + \sum\limits_{j : \|i - j\| = 1} v(\eta_i, \eta_j')]}{Z_i(\eta_{U+i}')} \quad \text{if} \quad i \in S_{q(n)}$$

$$= \delta_{\eta_i, \eta_i'} \qquad \qquad \text{else.} \tag{4-12}$$

The β-dependence of the dynamics is explicitly indicated.

Let $N = 2T$ and $R = A_L$, $\mathbb{L} = (2L + 1)\mathbb{Z}^d$. The covering of the strip S_N is generated by $\Gamma_{N,R}$ containing one "large" trapezoid Λ and lots of unit trapezoids:

the large trapezoid $\Lambda \subset S_N$ is the maximal (in the sense of inclusion) trapezoid with base $B(\Lambda) = R \subset \mathbb{Z}^d$; the unit trapezoids are used to fill the rest of the strip, i.e. for each point $x \in S_N$ which is not covered by $\Lambda + l$ for any $l \in \mathbb{L}$, we add the unit trapezoid Λ_x. To construct the couplings we proceed as in (4-1): we begin with the unit trapezoids. For each $x = (n, i) \in S_N$ we define the coupling $P_x^{\bar{\sigma}^1_{U(x)}, \bar{\sigma}^2_{U(x)}}(\sigma_x^1, \sigma_x^2)$ of the two conditional distributions $p_x(\sigma_x^1 | \bar{\sigma}^1_{U(x)})$, $p_x(\sigma_x^2 | \bar{\sigma}^2_{U(x)})$ (see (4-12)) to be

 i) arbitrary, if for some point $y \in U(x)$, $\bar{\sigma}^1_{U(x)} \neq \bar{\sigma}^2_{U(x)}$ or if $i \notin S_{q(n)}$, and

 ii) diagonal in all other cases: $P_x^{\bar{\sigma}^1_{U(x)}, \bar{\sigma}^2_{U(x)}}(\sigma_x^1 \neq \sigma_x^2) = 0$. For $x \in I(\Lambda)$, let $B(x) \subset B(\Lambda)$ be the base of the smallest trapezoid with the base in $\mathbb{Z}_0^d$ which contains x. Then, for any $\varepsilon > 0$, there is a $\beta_o = \beta_o(L, \varepsilon)$ such that for all $\beta > \beta_o$, for all $x = (N, i) \in \Lambda$ and any two initial data $\bar{\sigma}^1, \bar{\sigma}^2$,

$$P_\Lambda^{\bar{\sigma}^1, \bar{\sigma}^2}(\sigma_x^1 \neq \sigma_x^2) \leqq \varepsilon \sum_{y \in B(x)} \delta_{\bar{\sigma}^1_y, \bar{\sigma}^2_y}, \tag{4-13}$$

with $P_\Lambda^{\bar{\sigma}^1, \bar{\sigma}^2}$ the coupling (as in (4-1)) on the trapezoid Λ. The proof of (4-13) is easy and amounts to inspecting the one-point conditional Gibbs distributions $Q_i^\infty(\eta_i; \eta_{U+i})$ at $\beta = \infty$:

 i) $Q_i^\infty(\eta_i = 3; \eta_{U+i}) = 0$ for any η_{U+i}, and

 ii) $Q_i^\infty(\eta_i = 0; \eta_{U+i}) = 1$ provided that $n_j \neq 3$ for all $j \in U + i$. (4-14)

That means that in the middle level $M(\Lambda) \equiv \Lambda \cap \mathbb{Z}_T^d$ the probability

$$q_\Lambda^\infty\{\sigma \in \Omega_\Lambda : \sigma_x = 3 \text{ for some } x \in M(\Lambda)\} = 0, \tag{4-15}$$

which follows from (4-14i), and then, from (4-14ii),

$$q_\Lambda^\infty\{\sigma \in \Omega_\Lambda : \sigma_x = 0 \text{ for all } x = (N, i) \in \Lambda\} = 1. \tag{4-16}$$

Of course, the measure q_Λ^β depends continuously on β for $\beta \in [0, \infty]$ and hence, by construction of the coupling, (4-13). We conclude that (3-4) is satisfied for $x = (N, i) \in \Lambda$ with $k_{x,y}^\Lambda = \varepsilon$ if $y \in B(x)$, $= 0$ else. For the other points $x \in I(\Lambda)$, we simply put $k_{x,y}^\Lambda = 1$ for $y \in B(x)$, $= 0$ else. Finally, for all other (unit) trapezoids, we take $k_{x,y}^\alpha = 1$ for all $y \in U(x)$.

   We come now to the third problem, that of estimating the elements of the dependence matrix. By definition, for $(N, i) \in \Lambda$, $K_{i,j} = \varepsilon$ if $(0, j) \in B(N, i)$, $= 0$ else. For the other $i \in R$ we have the bound $K_{i,j} \leqq W_{i,j}$ if $j \in B(N, i)$, $= 0$ else; $W_{i,j}$ is the number of paths from (N, i) to $(0, j)$. Finally, (3-3) is verified since there are constants $c(d)$ and $c'(d)$ such that

$$\gamma \leqq \varepsilon c(d) + c'(d)L^{-1}. \tag{4.17}$$

Choosing L large enough, and then, ε small enough, we get $\gamma < 1$ as desired.

5. Discussion

We mention various aspects of our criterion:

1. Better estimates on the critical parameters can be obtained by applying our criterion using "bigger" trapezoids. In particular, $C_{N,R}$ is always better (in the sense

of "not worse") than the single site condition (1-2). The reason for this can be seen from formula (2-9): any "sub"-coupling on smaller trapezoids "admissibly covering" (in the sense of Definition 2) the bigger trapezoid Λ, can be used to construct the coupling for the conditional probabilities q_Λ. More precisely, let $\Gamma_{N,R}$ generate an admissible covering Π of the strip S_N with fundamental cell R and consider an admissible sub-covering Π' in which every trapezoid $\Lambda \in \Pi$ is covered by trapezoids $\Lambda' \in \Pi'$, $\Lambda' \subset \Lambda$ such that each point $x \in I(\Lambda)$ is an inner point of exactly one trapezoid Λ'. Then, we can construct a coupling $P_\Lambda^{\bar\sigma^1, \bar\sigma^2}$ for the conditional measures $q_\Lambda(\cdot; \bar\sigma^1)$ and $q_\Lambda(\cdot; \bar\sigma^2)$ from the corresponding couplings on the trapezoids Λ' (as was done to construct the coupling $P^{\eta, \eta'}$ in the proof of the Main Theorem). The simplest example is to consider the unit trapezoids $\{\Lambda_x\}_{x \in S_N \setminus \mathbb{Z}_0^d}$ as sub-covering. The γ (in (3-3)) it produces will just be the Nth power of (1-2). In Example 1, it would be $\gamma = 4(1 - \varepsilon)^2$; in Example 2, $\gamma = 9m^2$. We may expect that such a sub-coupling is increasing γ, as it should if the original construction was done in an efficient way. An unefficient way to apply our criterion would be to take $N = 1$ whatever we choose for fundamental cell R, since in this way we cannot really exploit the weak dependence on the initial conditions, beyond what is already contained in the single site condition (1-2).

2. The estimates we get from $C_{N,R}$ for a given PCA rule, depend on dimension and lattice. In a certain sense it is of course impossible to describe the *same* PCA rule on *different* lattices, but we will illustrate the meaning of our claim by considering Example 2 and comparing it with the Majority Vote model in one dimension. The latter is the PCA with transition probabilities identical to (4-2) but with neighborhood $U = \{-1, 0, 1\}$: the spin at site zero takes on the value of the majority vote in U, with probability $1 - 2\varepsilon$ and becomes ± 1 each with probability ε. While the rule looks similar, the dynamics it produces, are totally different. Toom's model satisfies the so-called "eroder condition" (see [LMS]) and its one dimensional counterpart certainly does not. In fact, it was recently proven by Gray that this one-dimensional Majority Vote model is ergodic for all $0 < \varepsilon \leq \frac{1}{2}$ (the samll ε version can be found in [G]). Still the single site estimate (1-2) gives the same bound as in the Toom model, $\varepsilon_c \leq \frac{1}{3}$, because there is no information on the lattice contained in it. Taking however a covering of the strip S_2 already shows a difference: at time $N = 2$ a spin depends on the configuration in 6 spins at time 0 for the Toom model and on 5 spins for the one-dimensional Majority Vote model. The estimates we get for the latter will therefore be better as can be checked by applying the analogous construction to Example 2, see Fig. 3. We find the corresponding $\gamma = \frac{15}{2}m^2 - \frac{1}{4}m^4$ implying the bound $\varepsilon_c \leq 0.317$. For larger N the difference becomes of course more outspoken.

3. As we saw in the previous section, $C_{N,R}$ is also useful for proving uniqueness of Gibbs states in ESM. Example 3 shows that $C_{N,R}$ does not imply the C_V condition of [DS]. It remains an open question whether the opposite holds true, see [AH] for a related question. However, comparing (1-2) with (4-5), it is obvious that the Dobrushin single site condition for the uniqueness of Gibbs measures is equivalent with the corresponding one site criterion for PCA. We conjecture that in general the $C_{N,R}$ criterion is less restrictive, i.e. if for some interaction the condition C_V holds for some finite $V \subset \mathbb{Z}^d$, then criterion $C_{N,R}$ is verified for a dynamics which leaves the corresponding Gibbs measures invariant.
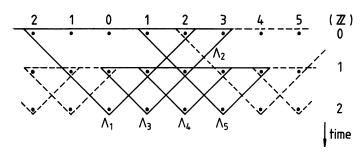
Fig. 3. Majority vote model

4. A more general setting of the problem is of course possible. First of all, there are the generalizations to compact metric state spaces, to the non-translation invariant case, to the non-finite range case, etc. which we consider doable. We are also optimistic about a continuous time formulation (the single site condition (1-2) is the $M < \varepsilon$ condition in [L]) and are encouraged by Theorem 2-12 in [AH]. An interesting generalization may also be found in the context of random fields possessing a *global* Markov property. Note that the PCA processes, as random fields on $\mathbb{Z}^{d+1}$, of course have this property, see also [GKLM]. The solution of the problem as presented here, indicates that the global Markov property may be the only natural condition for applying our method, just as the *local* Markov property (or, DLR-equation, cfr. (1-4)) is the crux in the methods of [DS] to establish uniqueness of Gibbs states.

References

[AH] Aizenman, M., Holley, R.: Rapid Convergence to equilibrium of stochastic Ising models in the Dobrushin–Shlosman regime In: Percolation Theory and Ergodic Theory of Infinite Particle Systems Kesten H. (ed.). IMA Vol. Math. Appl. vol. 8 pp 1–11. Berlin, Heidelberg, New York: Springer 1987

[D] Dobrushin, R. L.: Markov Processes with a large number of locally interacting components: Existence of a limit process and its ergodicity. Problems Inform. Transmission **7**, 149–164 (1971)

[D2] Dobrushin, R. L.: Description of a random field by means of its conditional probabilities and conditions of its regularity. Theory Probab. Appl. **13** 197–224 (1968)

[DS] Dobrushin, R. L., Shlosman, S. B.: Constructive criterion for the uniqueness of a Gibbs field. In: Statistical mechanics and dynamical systems. Fritz, J., Jaffe, A., Szasz, D. (eds.). pp 347–370. Boston: Birkhauser 1985

[DS2] Dobrushin, R. L., Shlosman, S. B.: Completely analytic interactions: constructive description. J. Stat. Phys. **46**, 983–1014 (1987)

[Du] Dudley, R. M.: Probability and Metrics. Aarhus Univ. Lecture Notes vol. **20**, 1976

[F] Feller: W.: An introduction to probability theory and its applications. New York: Wiley 1971

[Fe] Ferrari, P. A.: Ergodicity for Probabilistic Automata, To appear in Ann. Prob. (1989)

[Ge] Georgii,: H. O.: Gibbs measures and phase transitions. De Gruyter Studies in Mathematics, 1988

[GKLM] Goldstein, S., Kuik, R., Lebowitz, J. L., Maes, C.: From PCA's to Equilibrium systems and back. Commun. Math. Phys. **125** 71–79 (1989)

[G]   Gray, L.: The Behaviour of Processes with statistical mechanical properties. In:
         Percolation Theory and Ergodic Theory of Infinite Particle Systems. Kesten H.
         (ed.); IMA Vol. Math. Appl. vol. **8** Berlin, Heidelberg, New York: Springer 1987
[H]   Holley, R.: Possible rates of convergence in finite range, attractive spin systems.
         Contemp. Math. **41**, 215–234 (1985)
[HS]   Holley, R., Stroock, D. W.: Uniform and L^2 Convergence in One Dimensional
         Stochastic Ising Models. Commun. Math. Phys. **123**, 85–93 (1989)
[KRS]   Kirillov, A. B., Radulescu, D. C., Styer, D. F.: Vasserstein distances in two-state
         systems. J. Stat. Phys. **56**, 931–937 (1989)
[KMPSTV] Kurdumov, G. L., Mitjushin, L. G., Pirogov, S. A., Stavskaya, O. N., Toom, A. L.,
         Vasilyev, N. R.: Discrete Local Markov Systems, Preprint (1989)
[La]   Lamperti, J.: Stochastic processes: a survey of the mathematical theory. Berlin,
         Heidelberg, New York: Springer 1977
[LMS]   Lebowitz, J. L., Maes, C., Speer, E.: Statistical mechanics of probabilistic cellular
         automata. J. Stat. Phys. **59**, 117–170 (1990)
[LMS2]   Lebowitz, J. L., Maes, C., Speer, E.: Probabilistic cellular automata; statistical
         mechanical considerations. Proceedings of 1989 Complex Systems Summer School
         Santa Fe Institute, Santa Fe, New Mexico (1989)
[L]   Liggett, T. M.: Interacting Particle Systems. Berlin, Heidelberg, New York: Springer
         1985
[M]   Maes, C.: Some aspects of the stochastic time evolution of interacting spin systems.
         In: Leuven Notes in Mathematical and Theoretical Physics. Fannes, M., Verbeure,
         A. (eds.) vol. **1A**, pp. 91–132. Leuven, Belgium: Leuven University Press 1989
[Sh]   Shlosman, S. B.: Uniqueness and halfspace non-uniqueness and analyticity
         properties. Theor. Math. Phys. **66**, 284–293 (1986)
[Sh2]   Shlosman, S. B.: Unusual analytic Properties of Some Lattice Models: Complement
         of Łee-Yang Theory. Theor. Math. Phys. **69**, 1147–1150 (1986)
[S]   Steif, J.: $\bar{d}$-Convergence to Equilibrium and Space-time Bernoullicity for Spin
         Systems in the $M < \varepsilon$ Case. preprint, Rutgers University (1989)
[T]   Toom, A. L.: Stable and Attractive Trajectories in Multicomponent Random Fields.
         In: Multicomponent Random Fields. Dobrushin R. L., Sinai, Ya. G. (eds.),
         New York: Marcel Dekker 1980
[V]   Vasserstein, L. N.: Markov Processes over denumerable products of spaces,
         describing large sytems of automata. Problems Inform. Transmission **5**, 47–52 (1969)