ELSEVIER

Stochastic Processes and their Applications 96 (2001) 99–121

stochastic processes and their applications

www.elsevier.com/locate/spa

Convergence of locally and globally interacting Markov chains

Hans Föllmer, Ulrich Horst *

Institut für Mathematik—Bereich Stochastik, Humboldt-Universität zu Berlin, Unter den Linden 6, D-10099 Berlin, Germany

Received 15 March 2001; received in revised form 6 June 2001; accepted 13 June 2001

Abstract

We study the long run behaviour of interactive Markov chains on infinite product spaces. In view of microstructure models of financial markets, the interaction has both a local and a global component. The convergence of such Markov chains is analyzed on the microscopic level and on the macroscopic level of empirical fields. We give sufficient conditions for convergence on the macroscopic level. Using a perturbation of the Dobrushin–Vasserstein contraction technique we show that macroscopic convergence implies weak convergence of the underlying Markov chain. This extends the basic convergence theorem of Vasserstein for locally interacting Markov chains to the case where an additional global component appears in the interaction. © 2001 Elsevier Science B.V. All rights reserved.

MSC: 60K35; 60J10; 60J20

Keywords: Markov chains on infinite product spaces; Convergence of Markov chains; Contraction techniques; Gibbs measures

1. Introduction

We consider interactive Markov chains on a product space $S = C^{\mathbb{A}}$ where C is some finite state space and $\mathbb{A}$ is an infinite set of *sites* or *agents*. Thus, the state space of the Markov chain is the set of configurations $x = (x^a)_{a\in\mathbb{A}}$ which specify an individual state for each agent $a \in \mathbb{A}$. Suppose that the transition kernel is of the form

$$\Pi(x;\cdot) = \prod_{a\in\mathbb{A}} \pi^a(x;\cdot). \tag{1}$$

In such a situation, the behaviour of the agents is interactive insofar as the probability $\pi^a(x;c)$ that agent $a \in \mathbb{A}$ switches to the state $c \in C$ does not only depend on his own present state but may involve the states of other agents.

The convergence behaviour of Markov chains of the form (1) has been investigated in depth in the case where the interaction is purely *local*. This means that $\pi^a(x;\cdot)$ only depends on the states in some neighborhood $N(a)$. In this case Π may be viewed as

* Corresponding author. Tel.: +49-30-2093-1459; fax: +49-30-2093-1460.
E-mail address: horst@mathematik.hu-berlin.de (U. Horst).

PII: S0304-4149(01)00110-7

a Feller kernel on the compact state space S. Using Dobrushin's contraction technique and the Feller property, Vasserstein (1969) has shown that the Markov chain converges weakly to some unique equilibrium distribution μ if the interaction is not too strong.

In recent years there is an increasing interest in dynamical microstructure models of financial markets which involve interacting preferences and expectations of a large number of agents; see, e.g., Brock and Hommes (1997). In such a context, it becomes natural to introduce a global component into the interaction, i.e., to introduce some dependence on the average behaviour of the configuration $x \in S$ into the transition laws π^a. In Föllmer (1994) and Horst (2000) such Markov chains are used as a random environment for the evolution of stock prices, viewed as a sequence of temporary price equilibria. In order to analyze the asymptotic behaviour of such price processes, we need convergence results for the underlying Markov chain. This is the motivation for the present paper.

Our goal is to clarify some of the mathematical problems which arise in the presence of both a local and a global component in the interaction. We consider the case $\mathbb{A} = \mathbb{Z}^d$ where the average behaviour of a configuration $x \in S$ is described by the empirical distribution $\varrho(x)$ or, more completely by the empirical field $R(x)$. Due to the global dependence of the interaction the Feller property of Π will typically be lost. In order to prove convergence of the Markov chain $\{X_t\}_{t\in\mathbb{N}}$ governed by the kernel Π, we proceed in two steps. Due to a spatial law of large numbers for empirical fields, the macroscopic process $\{R(X_t)\}_{t\in\mathbb{N}}$ can be analyzed separately. Using contraction arguments with respect to a suitable metric, we obtain the convergence of the macroscopic process to some random field; this part is based on Horst (2000) and fills a gap in Föllmer (1979a). Our main result in Theorem 3.20 is based on a perturbation of the Dobrushin–Vasserstein contraction technique. We show that macroscopic convergence implies weak convergence of the underlying microscopic process $\{X_t\}_{t\in\mathbb{N}}$ to the same limiting random field. This may be viewed as an extension of Vasserstein's convergence theorem to the case where the interaction has both a local and a global component.

In the dynamical model (1) the individual transition laws π^a have an interactive structure, but the transition to a new configuration is made independently by the different agents. An interactive structure in the transition itself is captured by a model where the measure $\Pi(x;\cdot)$ is not a product measure but a Gibbs measure with respect to a system of conditional probabilities γ^x depending on the configuration x. Based on Horst (2000) and Horst (2001) we show how our convergence results can be extended to this general setting.

2. Locally and globally interacting Markov chains

Let C be some finite state space. We denote by $\mathbb{A}$ the d-dimensional integer lattice $\mathbb{Z}^d$ and by $S := C^{\mathbb{A}}$ the compact space of all configurations $x = (x^a)_{a\in\mathbb{A}}$ with $x^a \in C$. A probability measure μ on S will be called a *random field*. The space $\mathcal{M}(S)$ of all such random fields is compact with respect to the topology of weak convergence. Since the state space C is finite, the class $\mathcal{L}(S)$ of all *local* functions which depend

only on finitely many coordinates is dense in $\mathscr{C}(S)$ with respect to the topology of uniform convergence. Thus, a sequence $\{\mu_t\}_{t\in\mathbb{N}}$ of random fields converges weakly to $\mu\in\mathscr{M}(S)$ iff

$$\mu_t(f) := \int_S f \,\mathrm{d}\mu_t \overset{t\to\infty}{\to} \mu(f) \quad (f\in\mathscr{L}(S)). \tag{2}$$

Our aim is to analyze some aspects of the long run behaviour of interactive Markov chains on S with transition kernel $\Pi(x;\mathrm{d}y)$. Let us first assume that the kernel Π takes the product form

$$\Pi(x;\cdot) = \prod_{a\in\mathbb{A}} \pi^a(x;\cdot). \tag{3}$$

In such a model, the state of a single agent $a\in\mathbb{A}$ changes in reaction to the situation $x\in S$ according to the probability distribution $\pi^a(x;\cdot)$ on C. The individual transition probabilities $\pi^a(x;\cdot)$ have an interactive structure since they depend not only on the individual state x^a. Note, however, that the transition to a new configuration is made independently at different sites. In (10) below, we will admit an interactive structure in the transition itself. Such a situation is captured by a model where the measure $\Pi(x;\cdot)$ is not a product measure, but a Gibbs measure with respect to a system of conditional probabilities depending on the configuration x.

The convergence of interactive Markov chains of the form (3) has been investigated in depth in the case where the interaction is purely *local*, i.e., under the assumption that the individual transition law $\pi^a(x;\cdot)$ only depends on the local situation $(x^b)_{b\in N(a)}$ in some finite "neighborhood" $N(a)$; see, e.g., Föllmer (1979b), Lebowitz et al. (1990) or Vasserstein (1969). In such a situation, the stochastic kernel Π has the Feller property, i.e.,

$$\Pi f(\cdot) := \int_S f(x)\Pi(\cdot\,;\mathrm{d}x) \in \mathscr{C}(S)$$

whenever $f\in\mathscr{C}(S)$. This property is crucial for the basic convergence theorem in Vasserstein (1969): under suitable contraction bounds on the interaction between different sites Vasserstein (1969) establishes weak convergence of the Markov chain to some unique equilibrium distribution ν in the sense that

$$\lim_{t\to\infty} \mu\Pi^t(f) = \nu(f)$$

for all $f\in\mathscr{C}(S)$ and any initial distribution $\mu\in\mathscr{M}(S)$. Due to (2), weak convergence of the sequence $\{\mu\Pi^t\}_{t\in\mathbb{N}}$ may be viewed as a notion of *local convergence*.

The purpose of the present paper is to introduce a macroscopic component both into the interaction and into the notion of convergence. This means that for a given configuration $x=(x^a)_{a\in\mathbb{A}}\in S$, the influence of x at site $a\in\mathbb{A}$ is not only felt through the local situation $(x^b)_{b\in N(a)}$ in some neighborhood $N(a)$ of a but also through some global aspects of x. In the presence of a global component in the interaction, the Feller property of the transition kernel Π will typically be lost, and so we cannot apply the method of Vasserstein (1969) in order to study the asymptotic behaviour of the Markov chain on S.

In the following simple example where the transition behaviour at site $a\in\mathbb{A}$ depends both on the individual state x^a and on an empirical average $m(x)$ associated with x,

this problem is easily solved because we can study separately the convergence on the macroscopic and on the microscopic level.

Example 2.1. Let $C = \{0, 1\}$ and denote by S_1 the set of all configurations such that the empirical average associated with the configuration $x \in S_1$ exists along a suitable sequence of finite sets $\mathbb{A}_n \uparrow \mathbb{A}$:

$$S_1 := \left\{ x \in S\colon \exists m(x) := \lim_{n\to\infty} \frac{1}{|\mathbb{A}_n|} \sum_{a\in\mathbb{A}_n} x^a \right\}.$$

For $x \in S_1$ we assume that

$$\pi^a(x;\cdot) = \pi(x^a, m(x);\cdot) \tag{4}$$

where π is a transition probability from $C \times [0,1]$ to C, and

$$\Pi(x;\cdot) = \prod_{a\in\mathbb{A}} \pi(x^a, m(x);\cdot) \tag{5}$$

for any $x \in S_1$. It follows from the strong law of large numbers that

$$\lim_{n\to\infty} \frac{1}{|\mathbb{A}_n|} \sum_{a\in\mathbb{A}_n} y^a = \lim_{n\to\infty} \frac{1}{|\mathbb{A}_n|} \sum_{a\in\mathbb{A}_n} \pi(x^a, m(x); 1) \quad \Pi(x;\cdot)\text{-a.s.}$$

The product-measure $\Pi(x;\cdot)$ given by (5) is therefore concentrated on the set S_1, and the empirical average satisfies

$$m(y) = F(m(x)) := m(x)\pi(1, m(x); 1) + (1 - m(x))\pi(0, m(x); 1)$$

for $\Pi(x;\cdot)$-*a.e.* $y \in S_1$. Thus, the Markov chain $\{X_t\}_{t\in\mathbb{N}}$ with transition probability Π on S_1 induces almost surely a deterministic sequence of empirical averages $\{m(X_t)\}_{t\in\mathbb{N}}$. The dynamics of this "macroscopic process" is specified by the iteration of the function F acting on the interval $[0,1]$. For any starting point $x \in S_1$, the process $\{X_t\}_{t\in\mathbb{N}}$ may therefore be viewed as a Markov chain evolving in the time inhomogeneous but deterministic environment $\{m_t\}_{t\in\mathbb{N}}$ defined recursively by

$$m_0 = m(x) \quad \text{and} \quad m_t := F(m_{t-1}) \quad (t \geqslant 1).$$

Suppose now that the macroscopic process converges to some $m^* \in [0,1]$. In this case, it is easily seen that we obtain weak convergence of the Markov chain $\{X_t\}_{t\in\mathbb{N}}$ to the unique equilibrium μ^* of the Feller kernel

$$\Pi_{m^*}(x; \mathrm{d}y) := \prod_{a\in\mathbb{A}} \pi(x^a, m^*; \mathrm{d}y^a).$$

This convergence result is a special case of Theorem 3.20 below. The preceding argument illustrates the method of separating the analysis of macroscopic and microscopic convergence.

Let us now consider the case where the individual behaviour is influenced both by an empirical average and by the situation in some neighborhood. We fix $l > 0$ and define the neighborhood of an agent $a \in \mathbb{A}$ as

$$N(a) := \{b \in \mathbb{A}\colon |b - a| \leqslant l\}.$$

If the transition probability $\pi^a(x;\cdot)$ depends both on some average of x and on the values x^b $(b \in N(a))$ then the analysis of the convergence behaviour of the Markov chain becomes more involved. Only in very special cases such as the following example, we can still obtain a simple macroscopic equation for the deterministic evolution of the sequence of empirical averages $\{m(X_t)\}_{t\in\mathbb{N}}$.

Example 2.2 (Föllmer, 1994). As an illustration of the interplay between the long run behaviour on the level of configurations and the asymptotics of the sequence of empirical averages $\{m(X_t)\}_{t\in\mathbb{N}}$, we fix constants $\alpha, \beta, \gamma > 0$, and consider the following simple voter model with $C = \{0,1\}$: For $x \in S_1$, the individual transition law $\pi^a(x;\cdot)$ is described as the convex combination

$$\pi^a(x;1) = \alpha p(x^a) + \beta m^a(x) + \gamma m(x), \tag{6}$$

where $\alpha + \beta + \gamma = 1$. Here, $m^a(x)$ is the proportion of '1' in the neighborhood $N(a)$. It is easy to see that the sequence of empirical averages satisfies almost surely the deterministic dynamics

$$m(X_{t+1}) = \alpha\{m(X_t)\,p(1) + (1 - m(X_t))\,p(0)\} + (1-\alpha)m(X_t).$$

Thus, the macroscopic process $\{m(X_t)\}_{t\in\mathbb{N}}$ converges almost surely to

$$m^* := \frac{p(0)}{1 + p(0) - p(1)}.$$

It follows from Theorem 3.20 below that the microscopic process $\{X_t\}_{t\in\mathbb{N}}$ converges in law to the unique equilibrium of the Feller kernel

$$\Pi_{m^*}(x;\cdot) := \prod_{a\in\mathbb{A}} \pi^a(x, m^*;\cdot)$$

where the probability distribution $\pi^a(x, m^*;\cdot)$ on C takes the form

$$\pi^a(x, m^*; 1) = \alpha\, p(x^a) + \beta m^a(x) + \gamma m^*.$$

Thus, the long run behaviour of the microscopic process $\{X_t\}_{t\in\mathbb{N}}$ is determined by the unique limit of the macroscopic process $\{m(X_t)\}_{t\in\mathbb{N}}$.

The next example shows that we will typically not obtain a simple equation which describes the dynamics of the sequence of empirical averages $\{m(X_t)\}_{t\in\mathbb{N}}$.

Example 2.3. Consider the following generalization of the voter model (6). For $x \in S_1$, the individual transition probabilities can be described by a measurable mapping $g : C^{|N(a)|} \times [0,1] \to [0,1]$ in the sense that

$$\pi^a(x;1) = g(\{x^b\}_{b\in N(a)}, m(x)). \tag{7}$$

Typically, we cannot expect that there exist a function $F : [0,1] \to [0,1]$ such that $m(X_{t+1}) = F(m(X_t))$. Nevertheless, we will show that the macroscopic process $\{m(X_t)\}_{t\in\mathbb{N}}$ converges almost surely if the mapping g satisfies a suitable contraction condition in its second argument; see Example 3.11 below. Due to Theorem 3.20 below, this will imply weak convergence of the microscopic process $\{X_t\}_{t\in\mathbb{N}}$.

We are now going to specify the mathematical framework which allows us to analyze the long run behaviour of the Markov chain $\{X_t\}_{t\in\mathbb{N}}$ both on the macroscopic and on the microscopic level. To this end, we introduce the family of shift-transformations θ_a $(a\in\mathbb{A})$ on S defined by $(\theta_a x)(b)=x^{a+b}$.

Definition 2.4. (i) A probability measure $\mu\in\mathcal{M}(S)$ is called homogeneous, if μ is invariant under the shift maps θ_a. By

$$\mathcal{M}_h(S) := \{\mu\in\mathcal{M}(S):\ \mu=\mu\circ\theta_a \text{ for all } a\in\mathbb{A}\}$$

we denote the class of all homogeneous random fields μ on S.

(ii) A homogeneous probability measure $\mu\in\mathcal{M}_h(S)$ is called ergodic, if μ satisfies a 0-1-law on the σ-field of all shift invariant events. The class of all ergodic probability measures μ on S is denoted by $\mathcal{M}_e(S)$.

For a given $n\in\mathbb{N}$ we put

$$\mathbb{A}_n := [-n,n]^d\cap\mathbb{A}$$

and denote by S_e the set of all configuration $x\in S$ such that the *empirical field* $R(x)$, defined as the weak limit

$$R(x) := \lim_{n\to\infty}\frac{1}{|\mathbb{A}_n|}\sum_{a\in\mathbb{A}_n}\delta_{\theta_a x}(\cdot), \tag{8}$$

exists and belongs to $\mathcal{M}_e(S)$. The empirical field $R(x)$ carries all macroscopic information about the configuration $x=(x^a)_{a\in\mathbb{A}}\in S_e$. In particular, the *empirical distribution*

$$\varrho(x)=\lim_{n\to\infty}\frac{1}{|\mathbb{A}_n|}\sum_{a\in\mathbb{A}_n}\delta_{x^a}(\cdot)$$

is given as the one-dimensional marginal distribution of $R(x)$.

Consider the product kernel Π defined by the transition laws π^a in (7). Proposition 3.1 below shows that the measure $\Pi(x;\cdot)$ $(x\in S_e)$ is concentrated on the set S_e and that the empirical average satisfies

$$\begin{aligned} m(y) &= \lim_{n\to\infty}\frac{1}{|\mathbb{A}_n|}\sum_{a\in\mathbb{A}_n} y^a \\ &= \lim_{n\to\infty}\frac{1}{|\mathbb{A}_n|}\sum_{a\in\mathbb{A}_n}\pi(\theta_a x, m(x);1) \\ &= \int \pi(z,m(x);1)R(x)(\mathrm{d}z) \\ &:= G(R(x)) \end{aligned}$$

for $\Pi(x;\cdot)$-a.e. $y\in S_e$. Thus, we have to consider the full dynamics of the sequence of empirical fields $\{R(X_t)\}_{t\in\mathbb{N}}$ even if, as in Example 2.3, the behaviour of agent $a\in\mathbb{A}$ depends on $R(x)$ only on the empirical average $m(x)$. Our aim is to formulate conditions on the individual transition laws which guarantee convergence of the sequence of empirical fields $\{R(X_t)\}_{t\in\mathbb{N}}$ and to analyze the interplay between convergence of the Markov chain $\{X_t\}_{t\in\mathbb{N}}$ on the macroscopic level and on the microscopic level.

2.1. Macroscopic interaction: independent transitions

Let us now be more specific about the structure of the individual transition probabilities π^a. We assume that the interaction is *spatially homogeneous* and that the interactive influence of the present configuration x at site a is felt both through the local situation in the neighborhood $N(a)$ of a and through the average situation throughout the whole system. This average situation is described by the empirical distribution $\varrho(x)$ or, more completely, by the empirical field $R(x)$ associated with $x \in S_e$. Thus, we consider individual transition laws which take the form

$$\pi^a(x;\cdot) = \pi_{R(x)}(\theta_a x;\cdot)$$

where $\pi_\mu(x;\cdot)$ is a stochastic kernel from $S \times \mathcal{M}_h(S)$ to C.

Assumption 2.5. The probability laws $\{\pi_\mu(x;\cdot)\}_{x\in S}$ satisfy a spatial Markov property of order l in their dependence on the present configuration:

$$\pi_\mu(\theta_a x;\cdot) = \pi_\mu(\theta_a y;\cdot) \quad \text{if } \theta_a x = \theta_a y \text{ on } N(a).$$

Moreover, we assume that the mapping $\mu \mapsto \pi_\mu(x;\cdot)$ is continuous.

Let us now fix a homogeneous random field $\mu \in \mathcal{M}_h(S)$ and a configuration $x \in S$. It follows from our Assumption 2.5 that

$$\Pi_\mu(x;\cdot) := \prod_{a\in\mathbb{A}} \pi_\mu(\theta_a x;\cdot) \tag{9}$$

defines a Feller kernel on the configuration space S. In particular,

$$\Pi(x;\cdot) := \Pi_{R(x)}(x;\cdot) = \prod_{a\in\mathbb{A}} \pi_{R(x)}(\theta_a x;\cdot)$$

defines a stochastic kernel from S_e to S. In fact, we will see in Proposition 3.1 below that Π may be viewed as a stochastic kernel on the configuration space S_e. In contrast to the stochastic kernels Π_μ, the kernel Π typically does not have the Feller property, due to the macroscopic dependence on the present configuration x via the empirical field $R(x)$.

2.2. Macroscopic interaction: interactive transitions

Let us now extend the previous setting by introducing an interactive structure into the transition itself. This idea is captured by a model where $\Pi(x;\cdot)$ is not a product measure, but a *Gibbs measure* with respect to a system of conditional probabilities γ^x depending on the configuration x; see, e.g., Georgii (1989).

In order to make this more precise, we fix for any configuration $x \in S$ and for every homogeneous random field μ on S, a *local specification* $\gamma^{x,\mu} = (\gamma_a^{x,\mu})_{a\in\mathbb{A}}$; here $\gamma_a^{x,\mu}$ is a stochastic kernel from $C^{\mathbb{A}-\{a\}}$ to C which specifies the transition behaviour of agent $a \in \mathbb{A}$, given a *boundary condition* v on $\mathbb{A} - \{a\}$, i.e., the new states of the other agents.

We assume that the mapping $\mu \mapsto \gamma_0^{x,\mu}(\cdot\,; v)$ is continuous, and that $\gamma^{x,\mu}$ satisfies a Markov property of order l both in its dependence on the boundary condition and on the present configuration: For any fixed $x \in S$, we have

$$\gamma_a^{x,\mu}(\cdot\,; v) = \gamma_a^{x,\mu}(\cdot\,; w) \quad \text{if } v = w \text{ on } N(a) - \{a\}$$

and for each fixed boundary condition v on $\mathbb{A} - \{a\}$, we have

$$\gamma_a^{x,\mu}(\cdot\,; v) = \gamma_a^{y,\mu}(\cdot\,; v) \quad \text{if } x = y \text{ on } N(a).$$

If the transition to a new configuration is made independently by different agents, given the configuration x, the preceding conditions reduce to our Assumption 2.5. We also assume that the interaction is spatially homogeneous:

$$\gamma_a^{\theta_a x,\mu}(\cdot\,; \theta_a v) = \gamma_0^{x,\mu}(\cdot\,; v) \circ \theta_a \qquad (a \in \mathbb{A}).$$

Due to Dobrushin's fundamental uniqueness theorem, the specification $\gamma^{x,\mu}$ determines a unique random field $\Pi_\mu(x;\cdot)$ if we impose a suitable contraction condition on the specification; see, e.g., Georgii (1989), Theorem 8.7. Thus, the family of conditional probabilities $(\gamma^{x,R(x)})_{x \in S_e}$ defines a stochastic kernel

$$\Pi(x;\cdot) := \Pi_{R(x)}(x;\cdot) \tag{10}$$

from S_e to S; the product structure (3) is included as a special case. In fact, we will see that Π may be viewed as a stochastic kernel on the configuration space S_e.

3. Convergence theorems

We are now ready to study the dynamics of the interactive Markov chain $\{X_t\}_{t \in \mathbb{N}}$ on the state space S_e defined by the general transition kernel

$$\Pi(x;\cdot) = \Pi_{R(x)}(x;\cdot)$$

introduced in (10). In a first step, we use the following spatial law of large numbers for the random fields $\Pi_\mu(x;\cdot)$ in order to view Π as a transition kernel on the configuration space S_e. For the proof we refer to Horst (2000) or to Horst (2001); in the special product case (3) the argument is much simpler and can be found in Föllmer (1979a).

Proposition 3.1. *For all $x \in S_e$ and $\mu \in \mathcal{M}_h(S)$, the measure $\Pi_\mu(x;\cdot)$ is concentrated on the set S_e. For $\Pi_\mu(x;\cdot)$-a.e. $y \in S_e$, the empirical field $R(y)$ takes the form*

$$R(y)(\cdot) = \int_S \Pi_\mu(z;\cdot) R(x)(\mathrm{d}z).$$

Let $x \in S_e$. The preceding proposition shows that

$$R(y)(\cdot) = \int_{S_e} \Pi_{R(z)}(z;\cdot) R(x)(\mathrm{d}z) = \int_S \Pi_{R(x)}(z;\cdot) R(x)(\mathrm{d}z) \tag{11}$$

for $\Pi(x;\cdot)$-a.e. $y \in S_e$. In particular, we have

$$\Pi(x; S_e) = 1$$

for any $x \in S_e$, and so we will use S_e as the state space of the Markov chain $\{X_t\}_{t\in\mathbb{N}}$ with transition kernel Π. We denote by $\mathbb{P}_x$ the distribution of the chain $\{X_t\}_{t\in\mathbb{N}}$ with initial state $x \in S_e$. Since a configuration $x \in S_e$ induces an ergodic empirical field $R(x)$, the *microscopic process* $\{X_t\}_{t\in\mathbb{N}}$ induces $\mathbb{P}_x$-a.s. the *macroscopic process* $\{R(X_t)\}_{t\in\mathbb{N}}$ with state space $\mathscr{M}_e(S)$.

Let us now show that the spatial law of large numbers for ergodic empirical fields allows us to analyze the microscopic and the macroscopic process separately. In view of (11) the macroscopic process satisfies

$$R(X_{t+1}) = R(X_t)\Pi_{R(X_t)} \qquad \mathbb{P}_x\text{-a.s.}$$

i.e., the random field $R(X_{t+1})$ is $\mathbb{P}_x$-a.s. determined by the empirical field $R(X_t)$. In other words, we have

$$R(X_t) = R_t^x \qquad \mathbb{P}_x\text{-a.s.},$$

where we define the sequence of ergodic random fields $\{R_t^x\}_{t\in\mathbb{N}}$ by

$$R_0^x = R(x) \quad \text{and} \quad R_{t+1}^x := R_t^x \Pi_{R_t^x} = R_0^x \Pi_{R_0^x} \cdots \Pi_{R_t^x} \quad (t \in \mathbb{N}). \tag{12}$$

In this sense, for any initial state $x \in S_e$, the microscopic process may be viewed as a Markov chain evolving in a time inhomogeneous but deterministic environment $\{R_t^x\}_{t\in\mathbb{N}}$ which is generated by the macroscopic process. In particular, the law of the random variable X_{t+1} takes the form

$$\Pi^{t+1}(x;\cdot) = (\Pi_{R_0^x} \cdots \Pi_{R_t^x})(x;\cdot).$$

Our aim is now to study the asymptotics of the Markov chain $\{X_t\}_{t\in\mathbb{N}}$ both on the microscopic level and on the macroscopic level of empirical fields. Suppose that the microscopic process converges in law to some equilibrium ν_x in the sense that

$$\int_S f \,\mathrm{d}(\delta_x \Pi^t) \overset{t\to\infty}{\to} \int_S f \,\mathrm{d}\nu_x \quad (f \in \mathscr{C}(S)).$$

In this case, the sequence of empirical fields $\{R_t^x\}_{t\in\mathbb{N}} = \{R(x)\Pi^t\}_{t\in\mathbb{N}}$ converges weakly to the measure

$$\bar{\nu}_x(\cdot) := \int_S \nu_z(\cdot) R(x)(\mathrm{d}z).$$

Thus, Proposition 3.1 implies that we have at the same time convergence of macroscopic quantities of the form

$$\int f(z) R(X_t)(\mathrm{d}z) = \lim_{n\to\infty} \frac{1}{|\mathbb{A}_n|} \sum_{a\in\mathbb{A}_n} f(\theta_a X_t) \quad (f \in \mathscr{C}(S))$$

along $\mathbb{P}_x$-almost all paths of the microscopic process to $\int f(z)\bar{\nu}_x(\mathrm{d}z)$. In this sense, microscopic convergence implies macroscopic convergence. Sections 3.2, 3.3 and 3.4 may be viewed as a converse construction: Sections 3.2 and 3.3 provide a direct proof of macroscopic convergence. We will formulate conditions which guarantee that the macroscopic process $\{R(X_t)\}_{t\in\mathbb{N}}$ satisfies almost surely the contraction condition

$$d(R(X_{t+1}), R(X_t)) \leqslant \gamma d(R(X_t), R(X_{t-1})) \quad (\gamma < 1)$$

with respect to a suitable metric d on $\mathscr{M}(S)$, and this yields weak convergence of the sequence $\{R(X_t)\}_{t\in\mathbb{N}}$. The metric d will be introduced in the Section 3.1. In Section 3.4, we will show that macroscopic convergence implies microscopic convergence.

3.1. A metric for random fields

Let us denote by $\Delta_a(f)$ the oscillation of a function f on S at site $a\in\mathbb{A}$, i.e.,

$$\Delta_a(f) := \sup\{|f(x)-f(y)|: x=y \text{ off } a\}$$

and by

$$\Delta(f) := \sup\{|f(x)-f(y)|: x,y\in S\}$$

the oscillation of f on S. For any $f\in\mathscr{C}(S)$ we have

$$\Delta(f) \leqslant \sum_{a\in\mathbb{A}} \Delta_a(f).$$

We introduce a metric d on the class $\mathscr{M}(S)$ of all random field on S by

$$d(\mu,\nu) := \sup_{f\in\mathscr{C}(S)} \frac{|\mu(f)-\nu(f)|}{\sum_a 2^{\eta|a|}\Delta_a(f)} \quad (\mu,\nu\in\mathscr{M}(S)) \tag{13}$$

where η denotes a positive constant which will be specified later.

Remark 3.2. We have

$$d(\mu,\nu) \leqslant \sup_{f\in\mathscr{C}(S)} \frac{|\mu(f)-\nu(f)|}{\sum_a \Delta_a(f)} \leqslant \sup_{f\in\mathscr{C}(S)} \frac{|\mu(f)-\nu(f)|}{\Delta(f)} \leqslant \|\mu-\nu\|,$$

where $\|\mu-\nu\|$ denotes the total variation of the signed measure $\mu-\nu$ on S. The proof of the following proposition shows that

$$d(\mu,\nu) \leqslant d_V(\mu,\nu).$$

Here, d_V denotes the Vasserstein distance on $\mathscr{M}(S)$, i.e.,

$$d_V(\mu,\nu) := \sup\left\{\frac{|\mu(f)-\nu(f)|}{L(f)}: f\in\mathscr{C}(S)\right\}$$

and

$$L(f) := \sup_{x\neq y}\left\{\frac{|f(x)-f(y)|}{d_S(x,y)}\right\}$$

is the Lipschitz coefficient of the function f with respect to the metric

$$d_S(x,y) := \sum_{a\in\mathbb{A}} 2^{-\eta|a|}\mathbf{1}_{\{x^a\neq y^a\}}$$

on the configuration space S.

Proposition 3.3. *The metric d defined by* (13) *induces the weak topology on $\mathscr{M}(S)$. In particular, $(\mathscr{M}(S),d)$ is a compact metric space.*

Proof. In a first step, we are going to show that the metric d is dominated by the Vasserstein distance, i.e., we will verify that

$$d(\mu,\nu) \leqslant d_V(\mu,\nu). \tag{14}$$

To this end, let $f: S \to \mathbb{R}$ be a continuous function which satisfies

$$\sum_{a\in\mathbb{A}} 2^{\eta|a|}\Delta_a(f) < \infty.$$

In order to verify (14), it is enough to show that

$$L(f) \leqslant \sum_{a\in\mathbb{A}} 2^{\eta|a|}\Delta_a(f).$$

To this end, we fix $x, y \in S$ and put

$$J := \{a \in \mathbb{A} : x^a \neq y^a\}.$$

With no loss of generality, we may assume that $J = (j_n)_{n\in\mathbb{N}}$. Let $(x_n)_{n\in\mathbb{N}}$ be a sequence of configurations such that $x_0 = x$, such that $\lim_{n\to\infty} x_n = y$ and such that the following holds true for all $n, m \in \mathbb{N}$:

$$x_n^{j_n} \neq x_{n+1}^{j_n}, \qquad x_n^b = x_{n+1}^b \ (b \neq j_n), \qquad x_{n+m+1}^{j_n} = y^{j_n}.$$

Thus, we have that

$$\begin{aligned} |f(x) - f(y)| &\leqslant \sum_{a\in J} \Delta_a(f) \\ &\leqslant \sum_{a\in\mathbb{A}} \left(\sum_{b\in\mathbb{A}} 2^{\eta|b|}\Delta_b(f) \right) 2^{-\eta|a|} \mathbf{1}_{\{x^a\neq y^a\}} \\ &= d_S(x,y) \sum_{a\in\mathbb{A}} 2^{\eta|a|}\Delta_a(f). \end{aligned}$$

Dividing both sides of this inequality by $d_S(x,y)$, we see that

$$L(f) \leqslant \sum_{a} 2^{\eta|a|}\Delta_a(f).$$

This yields (14), and so $\lim_{n\to\infty} d(\mu_n,\mu) = 0$ whenever the sequence of random fields $\{\mu_n\}_{n\in\mathbb{N}}$ converges to μ in the weak topology.

Suppose now that $\lim_{n\to\infty} d(\mu_n,\mu) = 0$. In this case, we have

$$\lim_{n\to\infty} \mu_n(f) = \mu(f)$$

for any $f \in \mathscr{C}(S)$. This proves our assertion. □

3.2. Macroscopic convergence: independent transitions

Throughout this subsection, we assume that the stochastic kernel Π takes the product form (3), i.e., we assume that the transition to a new configuration is made independently at different sites, given the configuration $x \in S_e$.

Let us first formulate a uniform *Dobrushin–Vasserstein condition* on the individual transition probabilities in order to control the local interaction in the stochastic kernels Π_μ. To this end, we introduce a vector $r_a^\mu = (r_{a,i}^\mu)_{i\in\mathbb{A}}$ with components

$$r_{a,i}^\mu := \sup\{\tfrac{1}{2}\|\pi_\mu(x;\cdot) - \pi_\mu(y;\cdot)\| : x = y \text{ off } a - i\} \tag{15}$$

for any random field $\mu \in \mathscr{M}_h(S)$ and for every $a \in \mathbb{A}$. Note that $r_{a,i}^\mu = r_{a-i,0}^\mu$ by translation invariance.

Assumption 3.4. The vectors r_a^μ introduced in (15) satisfy

$$\alpha_0 := \sup_\mu \sum_a r_{a,0}^\mu < 1. \tag{16}$$

Remark 3.5. Under our Assumption 3.4, we may as well assume that the following "weighted" uniform Dobrushin–Vasserstein condition holds: for a small enough $\eta > 0$ we have

$$\alpha := \sup_\mu \sum_a 2^{\eta|a|} r_{a,0}^\mu < 1. \tag{17}$$

The equivalence of (16) and (17) follows from our Assumption 2.5 because the measures $\Pi_\mu(x;\cdot)$ have a product form, and because the mapping $\mu \mapsto r_{a,0}^\mu$ is continuous.

Remark 3.6. A vector $r = (r_a)_{a\in\mathbb{A}}$ is called an estimate for the random fields μ and ν on S if

$$|\mu(f) - \nu(f)| \leqslant \sum_{a\in\mathbb{A}} r_a \Delta_a(f) \tag{18}$$

for any $f \in \mathscr{C}(S)$. For two product measures $\mu = \prod_{a\in\mathbb{A}} \mu_a$ and $\nu = \prod_{a\in\mathbb{A}} \nu_a$ such an estimate is given by

$$r_a = \tfrac{1}{2}\|\mu_a - \nu_a\|, \tag{19}$$

cf., e.g., Simon (1993), Theorem V.2.2.

In view of (18) and (19) the product structure of the measures $\Pi_\mu(x;\cdot)$ implies

$$\Delta_a(\Pi_\mu f) \leqslant \sum_{i\in\mathbb{A}} r_{a,i}^\mu \Delta_i(f) \tag{20}$$

for any $f \in \mathscr{C}(S)$. Under Assumption 3.4 we obtain the estimate

$$\Delta(\Pi_\mu f) \leqslant \left(\sup_i \sum_a r_{a-i,0}^\mu\right) \sum_i \Delta_i(f) \leqslant \alpha_0 \sum_i \Delta_i(f).$$

For any sequence $\{\mu_t\}_{t\in\mathbb{N}}$ it follows by induction that

$$\Delta(\Pi_{\mu_0} \cdots \Pi_{\mu_t} f) \leqslant \sum_{a,i} r_{a,i}^{\mu_1} \Delta_i(\Pi_{\mu_1} \cdots \Pi_{\mu_t} f) \leqslant \alpha_0^{t+1} \sum_i \Delta_i(f) \tag{21}$$

and so

$$\lim_{t\to\infty} \Delta(\Pi_{\mu_0} \cdots \Pi_{\mu_t} f) = 0. \tag{22}$$

Remark 3.7. In the case where the transition kernel does not depend on μ, the preceding argument summarizes the proof of Vasserstein (1969) that the Markov chain Π converges to a unique equilibrium distribution. In our context, (22) shows that the microscopic process $\{X_t\}_{t\in\mathbb{N}}$ has local asymptotic loss of memory as soon as (16) holds true. In order to ensure weak convergence of the sequence $\{\Pi^t(x;\cdot)\}_{t\in\mathbb{N}}$, however, we need an additional contraction condition (see Assumption 25 below) which controls the dependence of the individual transition laws on the empirical fields.

Our Dobrushin-condition (16) allows us to establish the following contraction property of the transition kernels Π_μ.

Proposition 3.8. *Let $\nu,\tilde\nu\in\mathscr{M}(S)$ and $\mu\in\mathscr{M}_h(S)$. Under Assumption* 3.4 *we have that*

$$d(\nu\Pi_\mu,\tilde\nu\Pi_\mu)\leqslant\alpha d(\nu,\tilde\nu). \tag{23}$$

Proof. For any $\mu\in\mathscr{M}_h(S)$, let the vector $r_a^\mu=(r_{a,i}^\mu)_{i\in\mathbb{A}}$ be defined as in (15). Using (20) and (17), we obtain

$$\begin{aligned}\sum_a 2^{\eta|a|}\Delta_a(\Pi_\mu f)&\leqslant\sum_a\sum_i 2^{\eta|a|-\eta|i|}r_{a-i,0}^\mu 2^{\eta|i|}\Delta_i(f)\\&\leqslant\sup_i\left\{\sum_a 2^{\eta|a-i|}r_{a-i,0}^\mu\right\}\sum_i 2^{\eta|i|}\Delta_i(f)\\&=\sum_a 2^{\eta|a|}r_{a,0}^\mu\sum_i 2^{\eta|i|}\Delta_i(f)\\&\leqslant\alpha\sum_i 2^{\eta|i|}\Delta_i(f).\end{aligned}$$

In particular, we have for any $f\in\mathscr{C}(S)$ that

$$\frac{\sum_a 2^{\eta|a|}\Delta_a(\Pi_\mu f)}{\sum_a 2^{\eta|a|}\Delta_a(f)}\leqslant\alpha. \tag{24}$$

Since the transition probability Π_μ has the Feller property, we get

$$\begin{aligned}d(\nu\Pi_\mu,\tilde\nu\Pi_\mu)&=\sup_{f\in\mathscr{C}(S)}\frac{|\nu(\Pi_\mu f)-\tilde\nu(\Pi_\mu f)|}{\sum_a 2^{\eta|a|}\Delta_a(\Pi_\mu f)}\frac{\sum_a 2^{\eta|a|}\Delta_a(\Pi_\mu f)}{\sum_a 2^{\eta|a|}\Delta_a(f)}\\&\leqslant\alpha\sup_{f\in\mathscr{C}(S)}\frac{|\nu(\Pi_\mu f)-\tilde\nu(\Pi_\mu f)|}{\sum_a 2^{\eta|a|}\Delta_a(\Pi_\mu f)}\\&\leqslant\alpha\sup_{g\in\mathscr{C}(S)}\frac{|\nu(g)-\tilde\nu(g)|}{\sum_a 2^{\eta|a|}\Delta_a(g)}\\&=\alpha d(\nu,\tilde\nu)\end{aligned}$$

due to (24). This proves our assertion. □

Our goal is now to show that Assumption 3.4 combined with the following contraction condition implies weak convergence of the sequence of empirical fields $\{R_t^x\}_{t\in\mathbb{N}}$ $(x\in S_e)$ to a unique probability measure μ^* on S.

Assumption 3.9. There exists a constant $\beta<1-\alpha$ such that

$$\sup_x d(\Pi_\mu(x;\cdot),\Pi_\nu(x;\cdot))\leqslant \beta d(\mu,\nu) \tag{25}$$

for all $\mu,\nu\in\mathscr{M}_h(S)$.

Remark 3.10. Let us verify that our Assumption 3.9 holds true as soon as the individual transition laws satisfy

$$\sup_x \tfrac{1}{2}\|\pi_\mu(x;\cdot)-\pi_\nu(x;\cdot)\|\leqslant \beta d(\mu,\nu).$$

Indeed, since

$$\sup_x d(\Pi_\mu(x;\cdot),\Pi_\nu(x;\cdot))=\sup_x\sup_{f\in\mathscr{C}(S)}\frac{|\int f(y)(\Pi_\mu(x;\mathrm{d}y)-\Pi_\nu(x;\mathrm{d}y))|}{\sum_a 2^{\eta|a|}\Delta_a(f)},$$

and because the vector $r^{\mu,\nu}=(r_a^{\mu,\nu})_{a\in\mathbb{A}}$ with components

$$r_a^{\mu,\nu}:=\sup_x \tfrac{1}{2}\|\pi_\mu(x;\cdot)-\pi_\nu(x;\cdot)\| \tag{26}$$

is an estimate for the product measures $\Pi_\mu(x;\cdot)$ and $\Pi_\nu(x;\cdot)$, we obtain

$$\begin{aligned}\sup_x d(\Pi_\mu(x;\cdot),\Pi_\nu(x;\cdot))&\leqslant \sup_{f\in\mathscr{C}(S)}\frac{\sum_a \Delta_a(f)r_a^{\mu,\nu}}{\sum_a 2^{\eta|a|}\Delta_a(f)}\\&\leqslant \beta d(\mu,\nu)\sup_{f\in\mathscr{C}(S)}\frac{\sum_a \Delta_a(f)}{\sum_a 2^{\eta|a|}\Delta_a(f)}=\beta d(\mu,\nu).\end{aligned}$$

Example 3.11. Let us return to the individual transition laws introduced in (7). For any fixed $\mu\in\mathscr{M}_h(S)$ we can write

$$\pi_\mu(\theta_a x,\{+1\})=g((x^b)_{b\in N(a)},m(\mu)),$$

where $m(\mu):=\int_S x^0\,\mathrm{d}\mu$ for $\mu\in\mathscr{M}_h(S)$.

We assume that the mapping g satisfies a uniform Lipschitz condition in its second argument, i.e.,

$$|g(\cdot,m)-g(\cdot,\hat m)|\leqslant\beta|m-\hat m|.$$

For $f(x):=x^0$ we obtain

$$\begin{aligned}\sup_x \tfrac{1}{2}\|\pi_\mu(x;\cdot)-\pi_\nu(x;\cdot)\|&\leqslant\beta|m(\mu)-m(\nu)|\\&=\beta\frac{|\mu(f)-\nu(f)|}{\sum_a 2^{\eta|a|}\Delta_a(f)}\\&\leqslant\beta\, d(\mu,\nu)\end{aligned}$$

for all $\mu,\nu\in\mathscr{M}_h(S)$. In view of the previous remark, our Assumption 3.9 is satisfied whenever $\beta<1-\alpha$.

We are now going to prove the main theorem of this subsection.

Theorem 3.12. *If our Assumptions* 2.5, 3.4 *and* 3.9 *are satisfied, then there exists a unique homogeneous random field* ν *on* S *such that*

$$\mu R_t(\cdot) := \int_S R_t^x(\cdot)\mu(\mathrm{d}x) \xrightarrow{w} \nu(\cdot) \quad (t \to \infty) \tag{27}$$

for any initial distribution μ *on* S_e. *Here* $\xrightarrow{w}$ *denotes weak convergence of probability measures.*

Proof. Let us fix $x \in S_e$. We are going to show that the sequence of empirical fields $\{R_t^x\}_{t\in\mathbb{N}}$ defined recursively by (12) satisfies a contraction condition with respect to the metric d introduced in (13).

1. Due to Proposition 3.8, we know already that Assumption 3.4 implies

$$d(\nu\Pi_\mu, \tilde{\nu}\Pi_\mu) \leqslant \alpha d(\nu, \tilde{\nu}) \tag{28}$$

for any $\nu, \tilde{\nu}, \mu \in \mathscr{M}_h(S)$.

2. For $\mu, \nu \in \mathscr{M}_h(S)$, we can combine our Assumption 3.9 with (28) in order to obtain

$$d(\mu\Pi_\mu, \nu\Pi_\nu) \leqslant (\alpha + \beta)d(\mu, \nu). \tag{29}$$

Indeed, it follows from the definition of the metric d that

$$d(\nu\Pi_\mu, \nu\Pi_\nu) \leqslant \sup_x d(\Pi_\mu(x;\cdot), \Pi_\nu(x;\cdot))$$

and, due to (28), this implies

$$\begin{aligned} d(\mu\Pi_\mu, \nu\Pi_\nu) &\leqslant d(\mu\Pi_\mu, \nu\Pi_\mu) + d(\nu\Pi_\mu, \nu\Pi_\nu) \\ &\leqslant \alpha d(\mu, \nu) + \sup_x d(\Pi_\mu(x;\cdot), \Pi_\nu(x;\cdot)) \\ &\leqslant (\alpha + \beta)d(\mu, \nu) \quad (\alpha + \beta < 1). \end{aligned}$$

3. Let us now concentrate on the process $\{\mu R_t\}_{t\in\mathbb{N}}$. First, we fix $x \in S_e$ and analyze the case $\mu = \delta_x$. Since $R(x) = R_0^x$ and because

$$R_{t+1}^x = R_t^x \Pi^t = R_0^x \Pi_{R_0^x} \cdots \Pi_{R_t^x} = R_t^x \Pi_{R_t^x} \quad (t = 0, 1, 2, \ldots)$$

our estimate (29) yields the following contraction property for the sequence of ergodic random fields $\{R_t^x\}_{t\in\mathbb{N}}$:

$$\begin{aligned} d(R_{t+T}^x, R_t^x) &= d(R_{t+T-1}^x \Pi_{R_{t+T-1}^x}, R_{t-1}^x \Pi_{R_{t-1}^x}) \\ &\leqslant (\alpha + \beta)d(R_{t+T-1}^x, R_{t-1}^x) \\ &\leqslant (\alpha + \beta)^t d(R_T^x, R_0^x) \\ &\leqslant 2(\alpha + \beta)^t. \end{aligned}$$

Here, the last inequality follows from $d(\mu,\nu) \leqslant \|\mu-\nu\| \leqslant 2$; see Remark 3.2. In particular, we obtain that

$$\sup_T d(R^x_{t+T}, R^x_t) \leqslant 2(\alpha+\beta)^t$$

which shows that $\{R^x_t\}_{t\in\mathbb{N}}$ is a Cauchy sequence in the compact space $\mathscr{M}(S)$. Thus, the sequence $\{R^x_t\}_{t\in\mathbb{N}}$ converges weakly to some probability measure $\nu_x \in \mathscr{M}(S)$. Since $R^x_t \in \mathscr{M}_h(S)$ and because $\mathscr{M}_h(S)$ is a closed subset of $\mathscr{M}(S)$, the limit ν_x is a homogeneous random field. As the set $\mathscr{M}_e(S)$ is dense in $\mathscr{M}_h(S)$ but not closed, there is no reason to expect $\nu_x \in \mathscr{M}_e(S)$. It is now easily seen that, for any initial distribution μ on S_e, there exists a shift-invariant random field ν_μ on S, such that $\mu R_t \xrightarrow{w} \nu_\mu$ $(t\to\infty)$.

4. It remains to verify that $\nu_{\delta_x} = \nu_{\delta_y}$ for all $x, y \in S_e$. This, however, follows from

$$\begin{aligned} d(R^x_{t+1}, R^y_{t+1}) &= d(R^x_t \Pi_{R^x_t}, R^y_t \Pi_{R^y_t}) \\ &\leqslant (\alpha+\beta) d(R^x_t, R^y_t) \\ &\leqslant 2(\alpha+\beta)^{t+1} \overset{t\to\infty}{\to} 0. \end{aligned}$$

This proves our assertion. □

Let us now consider the case where the asymptotic behaviour of the macroscopic process depends on the initial configuration. To this end, we replace our Assumption 3.9 by the following weaker condition:

Assumption 3.13. For any $\mu \in \mathscr{M}_e(S)$, there exists constants $t(\mu) \in \mathbb{N}$ and $\beta < 1-\alpha$ such that

$$\sup_x d(\Pi_{\mu_s}(x;\cdot), \Pi_{\mu_s\Pi_{\mu_s}}(x;\cdot)) \leqslant \beta d(\mu_s, \mu_s\Pi_{\mu_s}) \tag{30}$$

for all $s \geqslant t(\mu)$. Here $\mu_0 = \mu$ and $\mu_{s+1} = \mu_s \Pi_{\mu_s}$. That is, we require (25) to hold true for all random fields ν which take the form $\nu = \mu_s \Pi_{\mu_s}$, $s \geqslant t(\mu)$.

As an example where Assumption 3.13 holds true whereas our Assumption 3.9 is violated, we consider the following variant of the voter model analyzed in Example 2.2.

Example 3.14. We put $C = \{0,1\}$, and assume that the individual transition probability takes the form

$$\pi^a(x;+1) = \alpha\, p(x^a) + \beta m^a(x) + \gamma f(m(x))$$

where α, β, γ are positive constants, where $m^a(x)$ denotes the proportion of "1" in the neighborhood of site a and where $f:[0,1]\to\mathbb{R}$ is a non-linear function. The special case $f(m) = \gamma m$ was analyzed in Example 2.2. In our present situation, the evolution of the sequence of empirical averages is almost surely described by the non-linear relation $m(X_{t+1}) = \hat{F}(m(X_t))$ $(t = 0,1,\ldots)$ where

$$\hat{F}(m) := \alpha\{mp(1) + (1-m)p(0)\} + \beta m + \gamma f(m).$$

It is easily seen that our Assumption 3.9 is violated whenever the mapping $\hat{F}$ has more than one fixed point.

Consider now the following situation: $p(1)=0,\ p(0)=\frac{1}{3},\ \alpha=\frac{1}{6},\ \beta=\frac{1}{4},\ \gamma=3$ and $f(m) := m^3(1-m^2)$. In this case, our uniform Dobrushin–Vasserstein condition is satisfied with $\alpha_0 \leqslant \frac{1}{2}$ and

$$\hat{F}(m)=\frac{1}{24}+\frac{5}{24}m+3m^3(1-m^2).$$

The mapping $\hat{F}$ has three fixed points: $m_0 \approx 0.07025$, $m_1 \approx 0.62885$, $m_2 \approx 0.75935$. Thus, our Assumption 3.9 does not hold. However, Assumption 3.13 is still satisfied. Indeed, an easy calculation shows that there exists a critical value

$$m_c := \hat{F}^{-1}(m_1) \neq m_1$$

such that the asymptotic behaviour of the sequence $\{m(X_t)\}_{t\in\mathbb{N}}$ depends in the following manner on the initial configuration:

$$\lim_{t\to\infty} m(X_t)=\begin{cases} m_0 & \text{if } m(x)\in[0,m_1)\cup(m_c,1] \\ m_1 & \text{if } m(x)\in\{m_1,m_c\} \\ m_2 & \text{otherwise} \end{cases} \qquad \mathbb{P}_x\text{-a.s.}$$

Since $f'(m_0)<\frac{1}{3}$ and $|f'(m_2)|<\frac{1}{3}$, we see that

$$|f(m(X_{t+1}))-f(m(X_t))| \leqslant \beta|m(X_{t+1})-m(X_t)| \quad \mathbb{P}_x\text{-a.s.}$$

where $\beta<\frac{1}{2}$ for all t sufficiently large. We can now proceed as in Example 3.11 in order to obtain

$$\sup_y d(\Pi_{R_t^x}(y;\cdot),\Pi_{R_t^x\Pi_{R_t^x}}(y;\cdot)) \leqslant \beta d(R_t^x, R_t^x\Pi_{R_t^x}),$$

where $\alpha_0+\beta<1$ for all $t\in\mathbb{N}$ large enough.

Let us now establish a generalization of Theorem 3.12.

Theorem 3.15. *Suppose that Assumptions* 2.5, 3.4 *and* 3.13 *are satisfied. In this case, the following holds true*:

(i) *For any $x\in S_e$, there exists a random field v_x such that $R_t^x \xrightarrow{w} v_x$ as $t\to\infty$.*

(ii) *For any initial configuration μ concentrated on S_e, we have*

$$\int R_t^x(\cdot)\mu(\mathrm{d}x)\xrightarrow{w} v_\mu(\cdot) := \int v_x(\cdot)\mu(\mathrm{d}x) \quad (t\to\infty).$$

Proof. Let us fix $x\in S_e$. Without loss of generality we may assume that $t(R(x))=1$. Using the same arguments as in the proof of Theorem 3.12 we get

$$d(R_{t+1}^x, R_t^x) \leqslant 2(\alpha+\beta)^t.$$

In particular, for any $\varepsilon>0$, there exists $t_0\in\mathbb{N}$ such that

$$\sup_T d(R_{t+T}^x, R_t^x) \leqslant 2\sum_{s\geqslant t_0}(\alpha+\beta)^s < \varepsilon$$

for all $t \geqslant t_0$. Thus, $\{R_t^x\}_{t\in\mathbb{N}}$ is again a Cauchy sequence with respect to the metric d, and so there exists a homogeneous probability measure ν_x on S such that $R_t^x \xrightarrow{w} \nu_x$ as $t \to \infty$. This yields our assertion. □

3.3. Macroscopic convergence: interactive transitions

Let us now return to the general setting of Section 2.2 and assume that the stochastic kernels Π_μ are determined by suitable families of local specifications $(\gamma^{x,\mu})_{x\in S}$.

Suppose that we have translation invariant estimates r_a^μ for the random fields $\Pi_\mu(x;\cdot)$ and $\Pi_\mu(y;\cdot)$ on S where $x = y$ off a.

Remark 3.16. Under suitable conditions on the specifications $\gamma^{x,\mu}$ there exists a constant $\lambda \geqslant 1$ such the vector r_a^μ with components

$$r_{a,i}^\mu = \frac{\lambda}{2} \sup\{\|\gamma_b^{x,\mu}(\cdot\,;v) - \gamma_b^{y,\mu}(\cdot\,;v)\|\colon x = y \text{ off } a, v \in S,\ b \in \mathbb{A}\}$$

defines a translation invariant estimate for the random fields $\Pi_\mu(x;\cdot)$ and $\Pi_\mu(y;\cdot)$ on S where $x = y$ off a; see Theorem V.2.2 in Simon (1993) or Theorem 8.20 in Georgii (1989) for details.

We assume that the estimates r_a^μ satisfy (17). Note, however, that in our present situation (16) and (17) are no longer equivalent: due to the interactive structure in the transition kernel Π_μ, the function $\Pi_\mu f$ does not belong to the class of local functions, even if $f \in \mathscr{L}(S)$.

We also assume that one of our Assumptions 3.9 or 3.13 is satisfied.

Remark 3.17. Under suitable conditions on the specifications $\gamma^{x,\mu}$ there exists a constant $\lambda \geqslant 1$ such the vector $r^{\mu,\nu}$ with components

$$r_a^{\mu,\nu} = \frac{\lambda}{2} \sup_{v,x} \|\gamma_0^{x,\mu}(\cdot\,;v) - \gamma_0^{x,\nu}(\cdot\,;v)\| \tag{31}$$

defines a translation invariant estimates for the random fields $\Pi_\mu(x;\cdot)$ and $\Pi_\nu(x;\cdot)$ on S. Suppose now that

$$\tfrac{1}{2} \sup_{v,x} \|\gamma_0^{x,\mu}(\cdot\,;v) - \gamma_0^{x,\nu}(\cdot\,;v)\| \leqslant \frac{\beta}{\lambda} d(\mu,\nu).$$

In this case, we obtain

$$\sup_x d(\Pi_\mu(x;\cdot), \Pi_\nu(x;\cdot)) \leqslant \beta d(\mu,\nu),$$

and so our Assumption 3.9 holds; cf. Remark 3.10.

An inspection of the proofs of Proposition 3.8 and Theorems 3.12 and 3.15 shows that all our arguments remain valid if the estimates r_a^μ satisfy (17) and if the dependence of the transition kernel Π_μ on the parameter μ satisfies the contraction condition specified in our Assumptions 3.9 and 3.13, respectively.

3.4. Microscopic convergence

In this subsection, we are going to prove that convergence on the macroscopic level of empirical fields implies local convergence on the microscopic level. Under suitable contraction and continuity assumptions we show that the microscopic and macroscopic limit coincide. Thus, we have at the same time macroscopic and microscopic convergence to the same random field μ on S.

Throughout this section, we assume that the sequence $\{R_t^x\}_{t\in\mathbb{N}}$ $(x\in S_e)$ converges in the weak topology to some random field ν_x on S. Recall that this convergence holds under Assumption 3.13 and under our Dobrushin condition (17). Moreover we assume that the behaviour of an individual agent depends continuously on the measure μ.

Assumption 3.18. Suppose that the measure $\Pi_\mu(x;\cdot)$ is a Gibbs measure with respect to a local specification $\gamma^{x,\mu}$. There exists a constant β^* such that

$$\sup_{x,v} \tfrac{1}{2}\|\gamma_0^{x,\mu}(\cdot\,;v)-\gamma_0^{x,v}(\cdot\,;v)\| \leqslant \beta^* d(\mu,v). \tag{32}$$

Remark 3.19. In the case where the measures $\Pi_\mu(x;\cdot)$ take the product form (3), the above assumption reduces to

$$\sup_{x} \tfrac{1}{2}\|\pi_\mu(x;\cdot)-\pi_\nu(x;\cdot)\| \leqslant \beta^* d(\mu,\nu). \tag{33}$$

Using a perturbation of the Dobrushin–Vasserstein contraction technique, we are now going to show that macroscopic convergence implies microscopic convergence.

Theorem 3.20. *Suppose that we have translation invariant estimates r_a^μ for the random fields $\Pi_\mu(x;\cdot)$ and $\Pi_\mu(y;\cdot)$ where $x=y$ off a and that our Assumptions* 3.4 *and* 3.18 *are satisfied. Let μ be an initial distribution which is concentrated on the set S_e and assume that the sequence of random fields $\{R_t^x\}_{t\in\mathbb{N}}$ converges for μ-a.e. $x\in S_e$ in the weak topology to some random field ν_x. Then the following holds true*:

(i) *The microscopic process $\{X_t\}_{t\in\mathbb{N}}$ converges in law to a probability measure $\bar\nu_\mu$. The random field $\bar\nu_\mu$ is the unique equilibrium of the Feller kernel Π_{ν_μ}, where $\nu_\mu := \int \nu_x \mu(\mathrm{d}x)$. That is*

$$\bar\nu_\mu = \bar\nu_\mu \Pi_{\nu_\mu}. \tag{34}$$

(ii) *The macroscopic and the microscopic limit coincide, i.e., $\bar\nu_\mu = \nu_\mu$. Thus, any limiting distribution is characterized by the fixed point property*

$$\nu_\mu = \nu_\mu \Pi_{\nu_\mu}. \tag{35}$$

Proof. Our proof extends an argument given in Föllmer (1979a) for the case of product kernels. For any initial distribution μ, we denote by $\mathbb{E}_\mu$ the expectation with respect to the law $\mathbb{P}_\mu$. We shall first consider the case $\mu=\delta_x$ and prove that our microscopic process converges in distribution to the unique equilibrium of the Feller kernel $\Pi_{\nu_{\delta x}}$.

1. Let us fix $x \in S_e$, a finite set $A \subset \mathbb{A}$ and some $B \subset C^A$. We are going to show that

$$\lim_{t\to\infty} \mathbb{E}_{\delta_x}[\Pi^t \mathbf{1}_B - \Pi^t_{\nu_x} \mathbf{1}_B] = 0. \tag{36}$$

Here, $\nu_x := \nu_{\delta_x}$. In Step 3 below, we use (36) and Vasserstein's convergence theorem in order to establish our assertion.

For $t, T \in \mathbb{N}$ we can write

$$\mathbb{E}_{\delta_x} \Pi^{t+T}(B) = \mathbb{E}_{\delta_x}[\Pi^t_{\nu_x} \mathbf{1}_B(X_T)] + R_{t,T},$$

where we put

$$R_{t,T} := \mathbb{E}_{\delta_x}[(\Pi_{R^x_T} \cdots \Pi_{R^x_{T+t-1}} - \Pi^t_{\nu_x})\mathbf{1}_B(X_T)].$$

In step 2 we show that $\lim_{T\to\infty} |R_{t,T}| = 0$ uniformly in $t \in \mathbb{N}$.

2. Note that

$$\begin{aligned} |R_{t,T}| &= \left| \sum_{k=1}^{t} \mathbb{E}_{\delta_x}[\Pi_{R^x_T} \cdots \Pi_{R^x_{T+k-2}} (\Pi_{R^x_{T+k-1}} \Pi^{t-k}_{\nu_x} - \Pi_{\nu_x} \Pi^{t-k}_{\nu_x}) \mathbf{1}_B(X_T)] \right| \\ &\leqslant \sum_{k=1}^{t} \mathbb{E}_{\delta_x} \left[\sup_y |(\Pi_{R^x_{T+k-1}} - \Pi_{\nu_x})(\Pi^{t-k}_{\nu_x} \mathbf{1}_B)(y)| \right]. \end{aligned}$$

Since the stochastic kernel Π_{ν_x} has the Feller property we can introduce continuous mappings $g_k \colon S \to \mathbb{R}$ $(k \in \mathbb{N})$ by

$$g_k(z) := \Pi^{t-k}_{\nu_x} \mathbf{1}_B(z)$$

and so

$$|R_{t,T}| \leqslant \sum_{k=1}^{t} \sup_y \left| \int g_k(z) \{\Pi_{R^x_{T+k-1}}(y; \mathrm{d}z) - \Pi_{\nu_x}(y; \mathrm{d}z)\} \right|.$$

For any $k \in \mathbb{N}$, let us define a vector $r^{R^x_k, \nu_x}$ by analogy with (31). In view of Remark 3.17 our continuity assumption (32) yields

$$\begin{aligned} &\sup_y \left| \int g_k(z) \{\Pi_{R^x_{T+k-1}}(y; \mathrm{d}z) - \Pi_{\nu_x}(y; \mathrm{d}z)\} \right| \\ &\quad \leqslant \lambda\beta^* \sum_i d(R^x_{T+k-1}, \nu_x) \Delta_i(g_k). \end{aligned}$$

This implies

$$\begin{aligned} |R_{t,T}| &\leqslant \lambda\beta^* \sum_k \left[\sum_i \Delta_i(g_k) d(R^x_{T+k-1}, \nu_x) \right] \\ &\leqslant \lambda\beta^* \left(\sup_{t \geqslant T} d(R^x_t, \nu_x) \right) \sum_{k,i} \Delta_i(g_k). \end{aligned}$$

Since the uniform Dobrushin–Vasserstein condition (16) is satisfied it follows from (21) that

$$\sum_{k,i} \Delta_i(g_k) \leqslant |A| \sum_{k=1}^{t} \alpha_0^{t-k} \leqslant |A| \frac{1}{1-\alpha_0}.$$

Thus, uniformly in $t \in \mathbb{N}$, we have

$$\lim_{T\to\infty} |R_{t,T}| \leqslant \frac{|A|\lambda\beta^*}{1-\alpha_0}\left(\lim_{T\to\infty}\sup_{t\geqslant T} d(R_t^x, v_x)\right) = 0 \tag{37}$$

as the sequence of empirical fields $\{R_t^x\}_{t\in\mathbb{N}}$ converges to v_x in the weak topology. This shows (36).

3. We shall now apply Vasserstein's convergence theorem in order to establish (i). Due to (37), we have that

$$\begin{aligned}&\lim_{t,T\to\infty} |\mathbb{E}_{\delta_x}[\Pi^{t+T}\mathbf{1}_B - \Pi_{v_x}^{t+T}\mathbf{1}_B]|\\ &\quad = \lim_{t,T\to\infty} |\mathbb{E}_{\delta_x}\Pi_{v_x}^t\mathbf{1}_B(X_T) + R_{t,T} - \mathbb{E}_{\delta_x}\Pi_{v_x}^t\mathbf{1}_B(X_T)| = 0.\end{aligned}$$

Due to (17), Vasserstein's theorem yields the existence of a unique random field $\bar{v}_x$ on S such that $\Pi_{v_x}^t(y;\cdot)\xrightarrow{w}\bar{v}_x(\cdot)$ as $t\to\infty$ $(y\in S)$. This shows (i) for $\mu=\delta_x$ since

$$\lim_{t\to\infty}\Pi^t(x;B) = \lim_{t,T\to\infty}\Pi^{t+T}(x;B) = \lim_{t,T\to\infty}\Pi_{v_x}^{t+T}(x;B) = \bar{v}_x(B).$$

4. Let μ be an initial distribution which is concentrated on the set S_e. From our preceding arguments it is easily seen that

$$\mu\Pi^t \xrightarrow{w}\bar{v}_\mu,$$

where $\bar{v}_\mu$ is the unique invariant measure of the kernel Π_{v_μ}.

5. Let us now verify the fixed-point property (35) and show that the limiting distribution both on the macroscopic and on the microscopic level coincide. Due to (17), the Feller kernel Π_{v_μ} admits a unique equilibrium, and so it is enough to show that $v_\mu = v_\mu\Pi_{v_\mu}$. To this end, it suffices to fix $x\in S_e$ and to consider the case $\mu=\delta_x$. In view of Remark 3.17 our Assumption 3.18 yields

$$\sup_x d(\Pi_\mu(x;\cdot),\Pi_\nu(x;\cdot)) \leqslant \lambda\beta^* d(\mu,\nu)$$

for some constant $\lambda\geqslant 1$. Thus, it follows from Proposition 3.8 and from the recursive definition of the sequence $\{R_t^x\}_{t\in\mathbb{N}}$ in (12) that

$$\begin{aligned}&d(v_x\Pi_{v_x}, v_x)\\ &\quad\leqslant d(v_x\Pi_{v_x}, v_x\Pi_{R_t^x}) + d(v_x\Pi_{R_t^x}, R_t^x\Pi_{R_t^x}) + d(R_t^x\Pi_{R_t^x}, v_x)\\ &\quad\leqslant \sup_y d(\Pi_{v_x}(y;\cdot),\Pi_{R_t^x}(y;\cdot)) + d(v_x, R_t^x) + d(R_{t+1}^x, v_x)\\ &\quad\leqslant \lambda\beta^* d(v_x, R_t^x) + d(v_x, R_t^x) + d(R_{t+1}^x, v_x).\end{aligned}$$

Since we assume that $\lim_{t\to\infty} d(R_t^x, v_x) = 0$, this yields

$$v_x = v_x\Pi_{v_x}.$$

Hence, it follows from (i) that $v_x = \bar{v}_x$, i.e., the limiting distribution on the microscopic and on the macroscopic level coincide.

This proves our assertion. □

Note that we have the following variant of Theorem 3.20.

Corollary 3.21. *Let (M, d_M) be a metric space and $\Phi : \mathcal{M}_h(S) \to M$ be a measurable mapping. Suppose we have a system of conditional probabilities $\gamma^{x,m}$ which determines a unique random field $\Pi_m(x;\cdot)$ on S for any $m \in M$. In this case, each $\gamma^{x,\Phi(R(x))}$ determines a unique measure $\Pi(x;\cdot) = \Pi_{\Phi(R(x))}(x;\cdot)$. We assume that the local specifications $\gamma^{x,m}$ are spatially homogeneous and that the dependence of the individual behaviour of agent $a \in \mathbb{A}$ on the macroscopic signal $\Phi(R(x))$ is continuous. It follows from the proof of Theorem* 3.20 *that convergence of the sequence $\{\Phi(R(X_t)\}_{t\in\mathbb{N}}$ to some Φ^* implies weak convergence of the Markov chain $\{X_t\}_{t\in\mathbb{N}}$ with transition kernel Π to the unique stationary measure μ^* of the Feller kernel Π_{Φ^*}.*

Example 3.22. Let us return to the dynamical model (6). Since the macroscopic process $\{m(X_t)\}_{t\in\mathbb{N}}$ converges almost surely to $m^* = p(0)/(1 + p(0) - p(1))$, the microscopic process $\{X_t\}_{t\in\mathbb{N}}$ converges in law to the unique equilibrium μ^* of the Feller kernel

$$\Pi_{m^*}(x;\cdot) = \prod_{a\in\mathbb{A}} \pi_{m^*}(\{x^b\}_{b\in N(a)};\cdot),$$

where

$$\pi_{m^*}(\{x^b\}_{b\in N(a)};1) = \alpha p(x^a) + \beta m_a(x) + \gamma m^*.$$

Acknowledgements

We would like to thank Alexander Schied for valuable discussions, and an anonymous referee for a number of comments which helped to improve the presentation of the results. Support of Deutsche Forschungsgemeinschaft via SFB 373, "Quantification and Simulation of Economic Processes", is gratefully acknowledged.

References

Brock, W.A., Hommes, C., 1997. A rational route to randomness. Econometrica 65 (5), 1059–1095.

Föllmer, H., 1979a. Macroscopic convergence of Markov chains on infinite product spaces. In: Colloquium on Random Fields, Ungarn. Math. Soc. Janos Bolyai, Esztergom, pp. 363–371.

Föllmer, H., 1979b. Tail Structure of Markov Chains on Infinite Product Spaces. Z. Wahrscheinlichkeitstheorie Verw. Geb. 50, 273–285.

Föllmer, H., 1994. Stock Price Fluctuation as a Diffusion in a Random Environment. Philos. Trans. Roy Soc. London, Ser. A 1684, 471–483.

Georgii, H., 1989. Gibbs Measures and Phase Transitions. Walter de Gruyter, Berlin.

Horst, U., 2000. Asymptotics of locally and globally interacting Markov chains arising in microstructure models of financial markets. Ph.D. thesis, Humboldt-Universität zu Berlin.

Horst, U., 2001. Asymptotics of locally interacting Markov chains with global signals. Working paper, Humboldt-Universität zu Berlin.

Lebowitz, J.L., Maes, C., Speer, E.R., 1990. Statistical mechanics of probabilistic cellular automata. J. Stastist. Phys. 59, 117–170.

Simon, B., 1993. The Statistical Mechanics of Lattice Gases. Princeton University Press, Princeton.

Vasserstein, L.N., 1969. Markov processes over denumerable product of spaces describing large systems of automata. Problemy Peredachi Informatsii 5, 64–72.