Asymptotics of locally interacting Markov chains with global signals

Ulrich Horst
Institut für Mathematik – Bereich Stochastik
Humboldt-Universität zu Berlin
Unter den Linden 6, D-10099 Berlin

3rd April 2001

Abstract

We study the long run behaviour of interactive Markov chains on infinite product spaces. The behaviour at a single site is influenced by the local situation in some neighborhood and by a random signal about the average situation throughout the whole system. The asymptotic behaviour 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. Combining a convergence result from the theory of random systems with complete connections with a perturbation of the Dobrushin-Vasserstein contraction technique we show that macroscopic convergence implies that the underlying Microscopic process has local asymptotic loss of memory.

Key Words: Markov chains on infinite product spaces, local asymptotic loss of memory, contraction techniques, Gibbs measures

AMS 1991 subject classification: 60K35, 60J10, 60J20

Abbreviated title: Markov chains with global signals

1 Introduction

We consider interactive Markov chains on a product space $E = C^{\mathbb{A}}$ 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 $a \in \mathbb{A}$. The convergence behaviour of Markov chains of the form

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

has been investigated in depth in the case where the interaction is purely *local*. This means that the probability $\pi^a(x; c)$ that agent $a \in \mathbb{A}$ switches to the state $c \in C$ only depends on the states in some neighborhood $N(a)$. In this case, Π may be viewed as a Feller kernel on the compact state space E. Using Dobrushin's contraction technique and the Feller property, Vasserstein (1969) has shown that such a Markov chain converges weakly to some unique equilibrium distribution μ if the interaction between different agents 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, the Markov chain Π describes the dynamics of all the individual agents' states governing, for instance, their expectations about the future evolution of asset prices. From an economic point of view, it is appropriate to assume that the price expectation of an agent $a \in \mathbb{A}$ for the following period does not only depend on the current states of his 'neighbors', but also on signals about the average expectation throughout the entire population. Thus, in the context of microstructure models, it becomes natural to introduce an additional dependence on the average behaviour of the configuration $x \in E$ into the interaction, i.e., into the transition laws π^a, and to study the run long behaviour of locally and globally interacting Markov chains.

Föllmer and Horst (2001) established an extension of Vasserstein's convergence theorem to the case where the interaction has both a local and a global component. They consider the case $\mathbb{A} = \mathbb{Z}^d$ where the average

behaviour of $x \in E$ is described by the associated empirical distribution $\varrho(x)$ or, more completely, by the empirical field $R(x)$. In such a situation, the Feller property of Π is typically lost. Using contraction arguments with respect to a suitable metric, Föllmer and Horst (2001) have shown that the microscopic process $\{x_t\}_{t\in\mathbb{N}}$ induces a sequence of empirical fields $\{R(x_t)\}_{t\in\mathbb{N}}$ which converges almost surely to some random field μ on E. Applying a perturbation of the Dobrushin-Vasserstein contraction technique, they proved that almost sure convergence of the *macroscopic process* $\{R(x_t)\}_{t\in\mathbb{N}}$ implies weak convergence of the underlying microscopic process $\{x_t\}_{t\in\mathbb{N}}$.

From an economic point of view, the model analyzed in Föllmer and Horst (2001) assumes that an individual agent $a \in \mathbb{A}$ has *complete information* about the average behaviour throughout the entire population. In view of many applications, however, such an assumption is rather restrictive. In microstructure models for financial markets, for instance, the empirical distribution $\varrho(x_t)$ may be regarded as the "mood of the market" in period t, and it seems more natural to assume that agents only have *incomplete information* about $\varrho(x_t)$. This means that the agents do not observe directly the average situation over the whole system $\mathbb{A}$, but receive a random signal whose law depends on $\varrho(x_t)$. In Horst (2000), Chapter 3, 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 process of empirical distributions $\{\varrho(x_t)\}_{t\in\mathbb{N}}$. This is the motivation for the present paper.

Our goal is to get some insight into the long run behaviour of locally and globally interacting Markov chains with transition kernel Π on an infinite product space of the form $E = C^{\mathbb{A}}$ where $\mathbb{A} = \mathbb{Z}^d$. We extend the model studied in Föllmer and Horst (2001) and consider Markov chains where the behaviour of an individual agent $a \in \mathbb{A}$ is influenced by the local situation in some neighborhood and by a random signal about the average situation throughout the whole population $\mathbb{A}$. We also admit an interactive structure in the transition itself. This means that $\Pi(x;\cdot)$ is a Gibbs measure with respect to a system of conditional probabilities depending on x; the product case (1) is included as a special case. The class of such interactive Markov

chains is introduced in Section 2.

In order to analyze the long run behaviour of the Markov chain $\{x_t\}_{t\in\mathbb{N}}$ governed by the kernel Π, we proceed in three steps. In Section 3, we prove a spatial law of large numbers for empirical fields. This allows us to analyze the asymptotics of the macroscopic process $\{R(x_t)\}_{t\in\mathbb{N}}$, and to extend a convergence result in Föllmer and Horst (2001) to the case of Gibbs measures. In Section 4, we prove that the macroscopic process may be viewed as the Markov chain associated with a certain random system with complete connections. Using a contraction argument with respect to a suitable metric, we obtain weak convergence of the macroscopic process to a unique equilibrium distribution. Combining a variant of the Dobrushin-Vasserstein contraction technique with a convergence result from the theory of random systems with complete connections, we show that convergence of the macroscopic process implies that the underlying microscopic process $\{x_t\}_{t\in\mathbb{N}}$ has local asymptotic loss of memory in the sense of Föllmer (1979b). In Theorem 4.19, we state conditions which ensure weak convergence of the microscopic process to a unique equilibrium distribution.

2 Locally Interacting Markov Chains with Global Signals

Let C be some finite state space. We denote by $\mathbb{A}$ the d-dimensional integer lattice $\mathbb{Z}^d$ and by $E := C^{\mathbb{A}}$ the compact space of all configurations $x = (x^a)_{a\in\mathbb{A}}$ with $x^a \in C$. A probability measure μ on E will be called a *random field*. The space $\mathcal{M}(E)$ of all such random fields is compact with respect to the topology of weak convergence. Since the state space C is finite, the class $\mathcal{L}(E)$ of all *local* functions which depend only on finitely many coordinates is dense in $\mathcal{C}(E)$ with respect to the topology of uniform convergence. Thus, a sequence $\{\mu_t\}_{t\in\mathbb{N}}$ of random fields converges weakly to $\mu \in \mathcal{M}(E)$ iff

$$\mu_t(f) := \int_E f d\mu_t \overset{t\to\infty}{\longrightarrow} \mu(f) \qquad (f \in \mathcal{L}(E)). \tag{2}$$

Our aim is to analyze some aspects of the long run behaviour of interactive Markov chains on E with transition kernel $\Pi(x; dy)$. 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 $x \in E$ 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 (17) 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., Vasserstein (1969) or Lebowitz, Maes, and Speer (1990). In such a situation, the stochastic kernel Π has the Feller property, i.e.,

$$\Pi f(\cdot) := \int_E f(x)\Pi(\cdot;dx) \in \mathcal{C}(E)$$

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

Föllmer and Horst (2001) introduced a *macroscopic* component both into the interaction and into the notion of convergence. In such a situation, the Feller property of Π will typically be lost. Föllmer and Horst (2001) analyzed the convergence behaviour Markov chains of the form (3) under

the assumption that the interactive influence of a given configuration $x = (x^b)_{b\in\mathbb{A}}$ on an individual agent $a \in \mathbb{A}$ is felt through the local situation $(x^b)_{b\in N(a)}$ in some neighborhood $N(a)$ and through the average situation throughout the whole system $\mathbb{A}$. The average situation of $x \in E$ is described by the associated empirical field $R(x)$, viewed as an ergodic random field on the configuration space E. Under suitable bounds on the local interaction between different agents and on the dependence of individual behaviour on the empirical field $R(x)$, they obtained convergence of the Markov chain $\{x_t\}_{t\in\mathbb{N}}$ governed by the transition kernel Π both on the microscopic level of configurations and on the macroscopic level of empirical fields.

In the present paper, we consider a randomized version of the model analyzed in Föllmer and Horst (2001). We study the long run behaviour of interactive Markov chains on infinite product spaces where the influence of a given configuration x at site $a \in \mathbb{A}$ is felt through the local situation in some neighborhood $N(a)$ and through a random signal about global properties of x. In many situations, such an approach provides an additional smoothing effect which allows us to prove convergence of the Markov chain $\{x_t\}_{t\in\mathbb{N}}$ on the macroscopic level without any condition which controls the dependence of individual behaviour on the signal about aggregate behaviour. In order to study the asymptotic behaviour of such Markov chains, we apply the method of separating the analysis of microscopic and macroscopic convergence introduced in Föllmer and Horst (2001). The following example where the probability that an agent $a \in \mathbb{A}$ switches to a state $c \in C$ depends both on his individual state x^a and on some random signal about the empirical average $m(x)$ associated with x illustrates this method.

Example 2.1 *Let $C = \{0, 1\}$ and denote by E_1 the set of all configurations such that the empirical average associated with the configuration $x \in E_1$ exists along a suitable sequence of finite sets $\mathbb{A}_n \uparrow \mathbb{A}$:*

$$E_1 := \left\{ x \in E : \exists\ m(x) := \lim_{n\to\infty} \frac{1}{|\mathbb{A}_n|} \sum_{a\in\mathbb{A}_n} x^a \right\}$$

Given a configuration $x \in E_1$, we assume that an individual agents reacts to his own state x^a and to a random signal $s \in [0, 1]$ about the empirical

average $m(x)$. The conditional law

$$Q(m(x);\cdot)$$

of the signal s, given the empirical average, $m(x)$ is described by a stochastic kernel Q on $[0,1]$. The situation analyzed in Föllmer and Horst (2001) corresponds to the case $Q(m;\cdot) = \delta_m(\cdot)$.

For $x \in E_1$, we assume that the probability that an agent switches to state $c \in C$ takes the form

$$\pi^a(x;c) = \int_0^1 \pi_s(x^a;c)Q(m(x);ds)$$

where $\pi_s(x^a;\cdot)$ is a transition kernel from $C \times [0,1]$ to C. For any fixed signal $s \in [0,1]$, the transition to a new configuration is therefore described by the product kernel

$$\Pi_s(x;\cdot) := \prod_{a\in\mathbb{A}} \pi_s(x^a;\cdot). \tag{4}$$

For $x \in E_1$, it follows from the strong law of large numbers that

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

Thus, the product-measure $\Pi_s(x;\cdot)$ given by (4) is concentrated on the set E_1 whenever $x \in E_1$, and the empirical average satisfies

$$m(y) = u(m(x),s) := m(x)\pi_s(1;1) + (1-m(x))\pi_s(0;1)$$

for $\Pi_s(x;\cdot)$-a.e. $y \in E_1$. Hence, the Markov chain $\{x_t\}_{t\in\mathbb{N}}$ with transition probability

$$\Pi(x;\cdot) := \int_0^1 \Pi_s(x;\cdot)Q(m(x);ds)$$

on E_1 induces almost surely the sequence of empirical averages $\{m(x_t)\}_{t\in\mathbb{N}}$. Conditioned on the environment $\{s_t\}_{t\in\mathbb{N}}$, this "macroscopic process" evolves almost surely in a deterministic manner. The dynamics of the macroscopic process can be described by a Markov chain $\{m_t\}_{t\in\mathbb{N}}$ on the state space $[0,1]$ whose transition operator U acts on the set bounded measurable functions $f : [0,1] \to \mathbb{R}$ according the formula

$$Uf(m) = \int f(u(m,s))Q(m;ds).$$

For any starting point $x \in E_1$, the microscopic process $\{x_t\}_{t\in\mathbb{N}}$ may therefore be viewed as a Markov chain evolving in the random environment $\{s_t\}_{t\in\mathbb{N}}$. The law the environment is governed by the initial empirical average $m(x)$. Below, we will formulate conditions on the individual transition probabilities π_s^a and on the stochastic kernel Q which ensure that the Markov chain $\{m_t\}_{t\in\mathbb{N}}$ converges in law to a unique equilibrium distribution. Combining techniques from the theory of random systems with complete connections with a variant of the Dobrushin-Vasserstein contraction technique, we will see that this implies that the Markov chain $\{x_t\}_{t\in\mathbb{N}}$ has local asymptotic loss of memory in the sense of Föllmer (1979b). This 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 a signal about the empirical average and by the situation in some neighborhood. We fix $l > 0$ and define the neighborhood of a *coalition* $A \subset \mathbb{A}$ as

$$N(A) := \{b \in \mathbb{A} : \exists\ a \in A \text{ such that } |b - a| \leq l\}.$$

If the transition probability $\pi_s^a(x;\cdot)$ depends on the values x^b in the neighborhood $N(a)$ of a, then the analysis of the convergence behaviour of the Markov chain becomes more involved. Only in very special cases such as in the following variant of Example 2.2 in Föllmer and Horst (2001), we can still obtain a simple macroscopic equation for the conditionally deterministic evolution of the sequence of empirical averages $\{m(x_t)\}_{t\in\mathbb{N}}$.

Example 2.2 *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 consider the following simple voter model with $C = \{0, 1\}$: For $x \in E_1$ and for a fixed signal $s \in [0, 1]$ about the empirical average $m(x)$, the individual transition law $\pi_s^a(x;\cdot)$ is described as the convex combination*

$$\pi_s^a(x; 1) = \alpha p(x^a) + \beta m^a(x) + \gamma s. \tag{5}$$

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

conditionally deterministic dynamics

$$m(x_{t+1}) = u(m(x_t), s_t) := \alpha\{mp(1) + (1-m)p(0)\} + \beta m(x_t) + \gamma s_t.$$

We assume that the conditional law $Q(m(x_t); \cdot)$ of the signal s_t given the empirical average $m(x_t)$ is described by a signal kernel Q on $[0,1]$. Thus, the macroscopic process $\{m(x_t)\}_{t\in\mathbb{N}}$ may be viewed as a Markov chain on the state space $[0,1]$ whose transition operator U is given by

$$Uf(m) = \int f(u(m,s))Q(m;ds).$$

In Theorem 4.12 below, we provide conditions which ensure that the macroscopic process converges in law to a unique equilibrium. Due to Theorem 4.15 this implies that the microscopic process $\{x_t\}_{t\in\mathbb{N}}$ has local asymptotic loss of memory in the sense of Föllmer (1979b).

The next example shows that the dynamics of the sequence $\{m(x_t)\}_{t\in\mathbb{N}}$ typically cannot be described by a Markov chain.

Example 2.3 *Consider the following generalization of the voter model (5). For $x \in E_1$ and $s \in [0,1]$, the individual transition probabilities can be described by a measurable mapping $g_s : C^{|N(a)|} \to [0,1]$ in the sense that*

$$\pi_s^a(x;1) = g_s\left(\{x^b\}_{b\in N(a)}\right). \tag{6}$$

Typically, we cannot expect that there exist a function $u : [0,1] \times [0,1] \to [0,1]$ such that $m(x_{t+1}) = u(m(x_t), s_t)$. Nevertheless, we will show that the macroscopic process $\{m(x_t)\}_{t\in\mathbb{N}}$ converges in law if the dependence of the mapping g on x^b $(b \in N(a))$ is not too strong; see Example 4.21 below.

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 E defined by $(\theta_a x)(b) = x^{a+b}$.

Definition 2.4 *(i) A probability measure $\mu \in \mathcal{M}(E)$ is called homogeneous, if μ is invariant under the shift maps $(\theta_a)_{a\in\mathbb{A}}$. By*

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

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

(ii) A homogeneous probability measure $\mu \in \mathcal{M}_h(E)$ 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 E is denoted by $\mathcal{M}_e(E)$.

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

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

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

$$R(x) := \lim_{n\to\infty} \frac{1}{|\mathbb{A}_n|} \sum_{a\in\mathbb{A}_n} \delta_{\theta_a x}(\cdot), \tag{7}$$

exists and belongs to $\mathcal{M}_e(E)$. The empirical field $R(x)$ carries all macroscopic information about the configuration $x = (x^a)_{a\in\mathbb{A}} \in E_e$. 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 $R(x)$.

Let us consider the product kernel Π_s on E governed by the individual transition laws π_s^a in (6). Proposition 3.1 below shows that the measure $\Pi_s(x;\cdot)$ $(x \in E_e)$ is concentrated on the set E_e and that the empirical average satisfies

$$\begin{aligned} m(y) &= \lim_{n\to\infty} \frac{1}{|\mathbb{A}_n|} \sum_{a\in\mathbb{A}_n} y^a \\ &= \lim_{n\to\infty} \frac{1}{|\mathbb{A}_n|} \sum_{a\in\mathbb{A}_n} \pi_s(\theta_a x; 1) \\ &= \int \pi_s(x;1) R(x)(dz) \\ &:= u(R(x), s) \end{aligned}$$

for $\Pi_s(x;\cdot)$-a.e. $y \in E_e$. 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 now 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

Consider the product kernel (3) and let us 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 $(x^b)_{b\in N(a)}$ in the neighborhood $N(a)$ of a and through some random signal about the average situation throughout the whole system which is described by the empirical field $R(x)$ associated with $x \in E_e$. We also assume that the conditional law

$$Q(R(x);\cdot) \tag{8}$$

of the signal s given the empirical field $R(x)$ is described by a stochastic kernel Q from $\mathcal{M}_h(E)$ to $\mathcal{S}$, where $(S,\mathcal{S})$ is an arbitrary measurable space, the *signal space*. The kernel Q will be called the *signal kernel*.

For a fixed signal $s \in S$, we consider individual transition laws which take the form

$$\pi_s^a(x;\cdot) = \pi_s(\theta_a x;\cdot) \tag{9}$$

where $\pi_s(x;\cdot)$ is a stochastic kernel from $E\times S$ to C.

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

$$\pi_s(\theta_a x;\cdot) = \pi_s(\theta_a y;\cdot) \quad \textit{if } \theta_a x = \theta_a y \textit{ on } N(a).$$

Let us now fix a signal $s \in S$ and a configuration $x \in E$. It follows from our Assumption 2.5 that

$$\Pi_s(x;\cdot) := \prod_{a\in\mathbb{A}} \pi_s(\theta_a x;\cdot) \tag{10}$$

defines a *Feller kernel* on the configuration space E which is spatially homogeneous in the sense that

$$\Pi_s f\circ\theta_a = \Pi_s(f\circ\theta_a)$$

for all $f \in \mathcal{C}(E)$ and all $a \in \mathbb{A}$. In particular, the individual transition laws π_s together with the signal kernel Q determine a stochastic kernel

$$\Pi(x;\cdot) := \int_S \Pi_s(\theta_a x;\cdot) Q(R(x);ds) \tag{11}$$

from E_e to E. In fact, we will see in Proposition 3.1 below that Π may be viewed as a stochastic kernel on the configuration space E_e. In contrast to Π_s, the kernel Π typically does not have the Feller property, due to the macroscopic dependence on the present configuration x via the random signal about 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 the random fields $\Pi_s(x;\cdot)$ are not *product* measures, but *Gibbs* measures with respect to a system of conditional probabilities $\gamma^{x,s} = \{\gamma_A^{x,s}\}_{A\in\mathcal{A}}$. Here,

$$\mathcal{A} := \{A \subset \mathbb{A} : |A| < \infty\}$$

denotes the class of all *local* subsets of $\mathbb{A}$, and $\gamma_A^{x,s}(\cdot;v)$ is a stochastic kernel form $E \times S \times C^{\mathbb{A}\backslash A}$ to C^A. For a given configuration $x \in E$ and a fixed signal $s \in S$, the kernel $\gamma_A^{x,s}$ specifies the joint behaviour of the coalition A, given a *boundary condition* on $\mathbb{A}\backslash A$, i.e., given the new states of all the agents $b \notin A$.

Let us now be more precise about the structure of the transition kernels Π_s. For any $s \in S$, we consider a local specification $\gamma^{x,s} = \{\gamma_A^{x,s}\}_{A\in\mathcal{A}}$ which is spatially homogeneous in the sense that

$$\gamma_A^{\theta_a x,s}(\cdot;\theta_a v) = \gamma_{-a+A}^{x,s}(\cdot;v) \circ \theta_a. \tag{12}$$

Assumption 2.6 *The local specifications $\gamma^{x,s}$ satisfy a Markov property of order l both in their dependence on the boundary condition v and on the present configuration x: For any fixed configuration x we have*

$$\gamma_A^{x,s}(\cdot;v) = \gamma_A^{x,s}(\cdot;w) \quad \textit{if } v = w \textit{ on } N(A)\backslash A.$$

For any fixed boundary condition v on A^c, we have

$$\gamma_A^{x,s}(\cdot;v) = \gamma_A^{y,s}(\cdot;v) \quad \textit{if } x = y \textit{ on } N(A).$$

Note that (12) and Assumption 2.6 reduce to (9) and Assumption 2.5, respectively, if the transition to a new configuration is made independently by different agents, given the configuration x.

Example 2.7 *Let us put $C = \{0, 1\}$ and assume that the local specifications $\gamma^{x,s}$ have the structure of an Ising model of statistical mechanics:*

$$\gamma_A^{x,s}(1; v) = \frac{\exp\left(T_1 \hat{m}^A(v) + T_2 m^A(x) + T_3 s\right)}{\exp\left(T_1 \hat{m}^A(v) + T_2 m^a(x) + T_3 s\right) + 1}. \tag{13}$$

Here, T_1, T_2, T_3 are positive constants and $\hat{m}^A(y)$ and $m^A(y)$ denotes the average situation of the configuration $y \in E$ in $N(A)\backslash A$ and in $N(A)$, respectively:

$$\hat{m}^A(y) := \frac{1}{|N(A)\backslash A|} \sum_{b \in N(A)\backslash A} y^b, \quad m^A(y) := \frac{1}{|N(A)|} \sum_{b \in N(A)} y^b.$$

Clearly, the local specifications in (13) are spatially homogeneous and satisfy a Markov property of order l.

We also assume that the interaction between different agents is not too strong. We specify this by means of a uniform Dobrushin contraction condition on the family of conditional probabilities $(\gamma^{x,s})_{x \in S}$.

Assumption 2.8 *The local specifications $\gamma^{x,s}$ satisfy the following uniform Dobrushin condition: Let $C(x, s) = (c_{a,b}(x, s))_{a,b \in \mathbb{A}}$ denote the Dobrushin interaction matrix for $\gamma^{x,s}$, i.e., put*

$$c_{a,b}(x, s) := \sup \left\{ \frac{1}{2} \|\gamma_b^{x,s}(\cdot; v) - \gamma_b^{x,s}(\cdot; w)\| : v = w \text{ off } a \right\}. \tag{14}$$

We have

$$\hat{c} := \sup_{x,s} \sup_b \sum_a c_{a,b}(x, s) < 1, \tag{15}$$

where $\|\cdot\|$ denotes the total variation norm of a signed measure.

Remark 2.9 *Since our specifications $\gamma^{x,s}$ are spatially homogeneous we have*

$$c_{a,b}(x, s) = c_{a-b,0}(\theta_{-b} x, s)$$

for all $a, b \in \mathbb{A}$, $x \in E$ and $s \in S$. Thus, (15) is equivalent to

$$\sup_{x,s} \sum_{a \in \mathbb{A}} c_{a,0}(x,s) < 1.$$

Moreover, it follows from Assumption 2.6 that

$$\sup_{x,s} c_{a,b}(x,s) = \sup_{x,s} c_{a-b,0}(x,s) = 0 \text{ for } |a-b| > l.$$

Example 2.10 *Consider the local specifications introduced in (13). It is well known that our uniform Dobrushin condition (15) is satisfied if T_1 is small enough.*

Due to Dobrushin's fundamental uniqueness theorem, our Assumption 2.8 excludes phase transitions. The random field specified by $\gamma^{x,s}$ is uniquely determined; see, e.g., Dobrushin (1968) or Georgii (1989), Theorem 8.7. Let us denote this random field by

$$\Pi_s(x; \cdot).$$

The family $(\gamma^{x,s})_{x \in E}$ defines a stochastic kernel Π_s on E. Due to our Assumption 2.6 and because of Proposition 7.11 and Theorem 8.23 (ii) in Georgii (1989), it is easy to show that the transition kernel Π_s has the *Feller property*. Due to (12), Π_s is spatially homogeneous, i.e.,

$$\Pi_s f \circ \theta_a = \Pi_s(f \circ \theta_a) \tag{16}$$

for all $f \in \mathcal{C}(E)$ and $a \in \mathbb{A}$. The local specifications $\gamma^{x,s}$ together with the signal kernel Q from $\mathcal{M}_h(E)$ to S introduced in (8) define a stochastic kernel

$$\Pi(x; \cdot) := \int_S \Pi_s(x; \cdot) Q(R(x); ds) \tag{17}$$

from E to E_e; the product kernel (11) is included as a special case. In fact, it follows form Proposition 3.1 below that Π may be viewed as a stochastic kernel on the configuration space E_e.

3 A Law of Large Numbers for Random Fields

This section is devoted to the proof of a spatial law of large numbers for ergodic empirical fields which will be the basis for our subsequent analysis. For the case of product kernel Π_s, the proof of the following proposition is much simpler and can be found in Föllmer (1979a).

Proposition 3.1 *("Law of large numbers"). Suppose that the local specifications $\gamma^{x,s}$ are spatially homogeneous and satisfy our Assumptions 2.6 and 2.8. Then the following holds true:*

(i) For all configurations $x \in E_e$ and for every signal $s \in S$, the measure $\Pi_s(x;\cdot)$ is concentrated on the set E_e. For $\Pi_s(x;\cdot)$-a.e. $y \in E_e$, the empirical field $R(y)$ takes the form

$$R(y)(\cdot) = \int_{E_e} \Pi_s(z;\cdot) R(x)(dz). \tag{18}$$

(ii) For any ergodic random field μ on E, we have $\mu\Pi_s \in \mathcal{M}_e(E)$.

Proof: In order to establish our assertion, we proceed in several steps.

1. Let $A \in \mathcal{A}$ and $f \in \mathcal{L}(E)$ be any $\mathcal{E}_A$-measurable function, where $\mathcal{E}_A$ denotes the σ-field generated by the projections $x \mapsto x^a$ $(a \in A)$. Since the stochastic kernels Π_s on E have the Feller property and are spatially homogeneous in the sense of (16), we have

$$\begin{aligned} \int_E \int_E f(y)\Pi_s(z;dy)R(x)(dz) &= \int_E (\Pi_s f)(z) R(x)(dz) \\ &= \lim_{n\to\infty} \frac{1}{|\mathbb{A}_n|} \sum_{a\in\mathbb{A}_n} \Pi_s(f\circ\theta_a)(x). \end{aligned}$$

We denote by $\mathbb{E}_{s,x}$ the expectation with respect to the measure $\Pi_s(x;\cdot)$, introduce the sets

$$L_n := \{a = (a_1,\ldots,a_d) \in \mathbb{A} : \max_i |a_i| = n\} \qquad (n \in \mathbb{N})$$

and put

$$Y_i := \sum_{a \in L_i} f \circ \theta_a \qquad (i \in \mathbb{N}).$$

Thus, for $x \in E_e$, we have

$$\lim_{n\to\infty} \frac{1}{|\mathbb{A}_n|} \sum_{i=0}^{n} \mathbb{E}_{s,x} Y_i = \int_E \Pi_s f(y) R(x)(dy).$$

In order to establish the existence of $R(y)$ and the identification (18), it is therefore enough to show that

$$\lim_{n\to\infty} \frac{1}{|\mathbb{A}_n|} \sum_{i=0}^{n} \{Y_i - \mathbb{E}_{s,x} Y_i\} = 0 \qquad \Pi_s(x;\cdot)\text{-a.s.} \tag{19}$$

2. We shall not prove (19) directly. Instead, we will first show that (19) holds true if we replace the expectation $\mathbb{E}_{s,x} Y_i$ by a suitable conditional expectation. It will then be verified that this conditional expectation can be chosen such that it is almost surely close enough to $\mathbb{E}_{s,x} Y_i$.

   To this end, let us introduce the random variables

$$M_n := \sum_{i=0}^{n} \{Y_i - \mathbb{E}_{s,x}[Y_i | Y_0, \ldots, Y_{i-1}]\} \quad (n \in \mathbb{N}),$$

   where $\mathbb{E}_{s,x}[Y_0 | Y_0, Y_{-1}] := \mathbb{E}_{s,x}[Y_0]$. Since the function $f : E \to \mathbb{R}$ is bounded, the sequence $\{M_n\}_{n\in\mathbb{N}}$ is a square integrable martingale with respect to the measure $\Pi_s(x;\cdot)$ and the filtration

$$\{\mathcal{E}_{n,0}\}_{n\in\mathbb{N}} := \{\sigma(Y_0, Y_1, \ldots, Y_n)\}_{n\in\mathbb{N}}.$$

   Note that $|L_n| \le 2d(2n+1)^{d-1}$ and that $|\mathbb{A}_n| = (2n+1)^d$. Thus, there exists a constant $c < \infty$ such that

$$\sum_{n\ge 1} \frac{\mathbb{E}_{s,x}[(M_n - M_{n-1})^2 | \mathcal{E}_{n-1,0}]}{|\mathbb{A}_n|^2} \le \sum_{n\ge 1} \frac{c}{(2n+1)^2} < \infty \quad \Pi_s(x;\cdot)\text{-a.s.}$$

   It follows from the strong law of large numbers for square integrable martingales that

$$\lim_{n\to\infty} \frac{1}{|\mathbb{A}_n|} \sum_{i=0}^{n} \{Y_i - \mathbb{E}_{s,x}[Y_i | Y_0, \ldots, Y_{i-1}]\} = 0 \qquad \Pi_s(x;\cdot)\text{-a.s.} \tag{20}$$

For $n, k \in \mathbb{N}$, let us put

$$Y_{n,0} := Y_n, \quad Y_{n,k} := \mathbb{E}_{s,x}[Y_n|\mathcal{E}_{n-1,k-1}], \quad \mathcal{E}_{n,k} := \sigma(Y_{0,k}, \ldots, Y_{n,k}),$$

where $\mathcal{E}_{-1,k} := \{\emptyset, \Omega\}$. Iterating (20) we obtain for any $k \in \mathbb{N}$ that

$$\begin{aligned} & \lim_{n\to\infty} \frac{1}{|\mathbb{A}_n|} \sum_{i=0}^{n} \{Y_{i,k} - \mathbb{E}_{s,x}[Y_i|\mathcal{E}_{i-1,k}]\} \\ = & \lim_{n\to\infty} \frac{1}{|\mathbb{A}_n|} \sum_{i=0}^{n} \{Y_{i,k} - Y_{i,k+1}\} = 0 \qquad \Pi_s(x;\cdot)\text{-a.s.} \end{aligned} \tag{21}$$

Observe that $\mathcal{E}_{n,k} \subset \mathcal{E}_{n+1,k}$ for all $k, n \in \mathbb{N}$. Thus, the random variable $Y_{i,k}$ is $\mathcal{E}_{n-1,k-1}$-measurable if $i \leq n$. This yields

$$\mathcal{E}_{n,k} \subset \mathcal{E}_{n-1,k-1} \subset \cdots \subset \mathcal{E}_{n-k,0} \qquad (n \geq k).$$

Due to (21), we have $\Pi_s(x;\cdot)$-a.s. that

$$\lim_{n\to\infty} \frac{1}{|\mathbb{A}_n|} \sum_{i=0}^{n} \{Y_i - \mathbb{E}_{s,x}[Y_i|\mathcal{E}_{i-1,k-1}]\} = \lim_{n\to\infty} \frac{1}{|\mathbb{A}_n|} \sum_{i=0}^{n} \{Y_i - Y_{i,k}\} = 0.$$

Hence, (19) holds true with $\mathbb{E}_{s,x}Y_i$ replaced by $\mathbb{E}_{s,x}[Y_i|\mathcal{E}_{i-1,k-1}]$. Our objective is now to show that we can find a large enough $k \in \mathbb{N}$ such that

$$\lim_{n\to\infty} \frac{1}{|\mathbb{A}_n|} \sum_{i=0}^{n} \{\mathbb{E}_{s,x}[Y_i|\mathcal{E}_{i-1,k-1}] - \mathbb{E}_{s,x}Y_i\} < \epsilon \qquad \Pi_s(x;\cdot)\text{-a.s.} \tag{22}$$

3. Let us fix $\epsilon > 0$. In order to achieve our goal, we are first going to verify that we can choose a constant $k_0 = k_0(\epsilon)$ such that

$$|\mathbb{E}_{s,x}[f \circ \theta_a|\mathcal{E}_{n-1,k-1}] - \mathbb{E}_{s,x}[f \circ \theta_a]| < \epsilon \qquad \Pi_s(x;\cdot)\text{-a.s.} \tag{23}$$

for all $n \geq k \geq k_0$. This will then allow us to establish (22).

To this end, we introduce, for $n \geq k$, the set

$$A^{n,k} := \bigcup_{i=0}^{n-k} \bigcup_{j \in L_i} \{j + A\}.$$

Since the mapping $f : E \to \mathbb{R}$ is $\mathcal{E}_A$-measurable, the random variable Y_i is $\mathcal{E}_{A^{n,k}}$-measurable whenever $i \leq n-k$, and so

$$\mathcal{E}_{n,k} \subset \mathcal{E}_{n-1,k-1} \subset \cdots \subset \mathcal{E}_{n-k,0} \subset \mathcal{E}_{A^{n,k}} \qquad (n \geq k). \tag{24}$$

Let us now denote by

$$\Pi^{v,A^{n,k}}(x;\cdot)$$

the conditional joint distribution of the random variables y^a $(a \in (A^{n,k})^c)$ with respect to $\mathcal{E}_{A^{n,k}}$ and $\Pi_s(x;\cdot)$, given the boundary condition v on $A^{n,k}$. Since the local specification $\gamma^x = \{\gamma_A^x\}_{A\in\mathcal{A}}$ satisfies our Assumption 2.8, we can apply Theorem 8.23 in Georgii (1989). The random field $\Pi^{v,A^{n,k}}(x;\cdot)$ is Markov of order l and is uniquely determined by its conditional distributions $\{\gamma_A^{v,x}\}_{A\in\mathcal{A}}$ which take the form

$$\gamma_A^{v,x}(\cdot;w) = \gamma_{A\setminus A^{n,k}}(\cdot;w^v). \tag{25}$$

Here, w^v is the boundary condition on $(A\setminus A^{n,k})^c$ which is equal to v on $A^{n,k}$ and equal to w on $A^c\setminus A^{n,k}$.

For any $a \in \mathbb{A}$, (24) implies that

$$\begin{aligned}\mathbb{E}_{s,x}[f\circ\theta_a|\mathcal{E}_{n-1,k-1}] &= \mathbb{E}_{s,x}\left[\mathbb{E}_{s,x}[f\circ\theta_a|\mathcal{E}_{A^{n,k}}]|\mathcal{E}_{n-1,k-1}\right]\\ &\leq \sup_v \int f\circ\theta_a(y)\Pi^{v,A^{n,k}}(x;dy) \quad \Pi_s(x;\cdot)\text{-a.s.}\end{aligned}$$

In particular, we have $\Pi_s(x;\cdot)$-a.s. that

$$\begin{aligned}&|\mathbb{E}_{s,x}[f\circ\theta_a|\mathcal{E}_{n-1,k-1}] - \mathbb{E}_{s,x}[f\circ\theta_a]|\\ \leq\ &\sup_v \left|\int f\circ\theta_a(y)\Pi^{v,A^{n,k}}(x;dy) - \int f\circ\theta_a(y)\Pi(x;dy)\right|.\end{aligned}$$

For any $n \in \mathbb{N}$ and for all $a \in L_n$, the distance

$$s(a+A, A^{n,k}) := \min\{|b-\widetilde{b}| \ :\ b \in a+A,\ \widetilde{b} \in A^{n,k}\}$$

of the sets $a+A$ and $A^{n,k}$ does only depend on k. We can therefore choose a small enough $\lambda \in \mathbb{R}$ and a sufficiently large $k_0 = k_0(\epsilon) \in \mathbb{N}$ satisfying

$$c(\lambda) := \sup_{x,s}\sum_b c_{b,0}(x,s)e^{\lambda|b|} < 1 \quad \text{and} \quad \exp(-\lambda s((a+A), A^{n,k})) < \frac{\epsilon}{2}$$

for all $n \geq k \geq k_0$ and $a \in L_n$.

For $n \geq k \geq k_0$, let us now choose a set $V \in \mathcal{A}$ such that, for all $a \in L_n$, the following holds true:

$$V \cap A^{n,k} = \emptyset, \quad a + A \subset V, \quad \exp(-\lambda s((a+A), V^c)) < \epsilon.$$

Thus, our uniform Dobrushin condition (15) together with Theorem 8.23 and with Remark 8.26 in Georgii (1989) yields the following estimate:

$$\begin{aligned}
& \sup_v \left| \int f \circ \theta_a(y) \Pi^{v, A^{n,k}}(x; dy) - \int f \circ \theta_a(y) \Pi(x; dy) \right| \\
= & \sup_v \left| \int f \circ \theta_a(y) (\Pi^{v, A^{n,k}}(x; \cdot) \gamma_V^{v,x})(dy) - \int f \circ \theta_a(y) \Pi(x; dy) \right| \\
\leq & \sup_w \left| \int f \circ \theta_a(y) (\gamma_V^x(dy; w) - \Pi(x; dy)) \right| \qquad (26) \\
\leq & \Delta(f) |A| \frac{1}{1 - c(\lambda)} \exp(-\lambda s((a+A), V^c)) \\
\leq & c(f) \epsilon
\end{aligned}$$

for some constant $c(f)$ depending on f. Here, (26) follows from (25).

4. We can now apply the preceding estimates in order to establish (22) and (19). Since $|L_n| \leq 2d(2n+1)^{d-1}$ it follows from (23) that there exists $c < \infty$ such that

$$|\mathbb{E}_{s,x}[Y_n | \mathcal{E}_{n-1,k-1}] - \mathbb{E}_{s,x} Y_n| \leq \epsilon c (2n+1)^{d-1} \qquad \Pi_s(x; \cdot)\text{-a.s.} \qquad (27)$$

for all $n \geq k \geq k_0$. Thus, as $|\mathbb{A}_n| = (2n+1)^d$, we deduce from (21) and from (27) that there exists a large enough $k \in \mathbb{N}$ which satisfies

$$\begin{aligned}
& \lim_{n \to \infty} \frac{1}{|\mathbb{A}_n|} \left| \sum_{i=0}^n \{Y_i - \mathbb{E}_{s,x} Y_i\} \right| \\
\leq & \lim_{n \to \infty} \frac{1}{|\mathbb{A}_n|} \left| \sum_{i=0}^n \{Y_i - \mathbb{E}_{s,x}[Y_i | \mathcal{E}_{i-1,k-1}]\} \right| \\
& + \lim_{n \to \infty} \frac{1}{|\mathbb{A}_n|} \left| \sum_{i=0}^n \{\mathbb{E}_{s,x}[Y_i | \mathcal{E}_{i-1,k-1}] - \mathbb{E}_{s,x} Y_i\} \right| \\
\leq & c\epsilon \qquad \Pi_s(x; \cdot)\text{-a.s.}
\end{aligned}$$

This shows (19) and, therefore, establishes the existence of the empirical field $R(y)$ for $\Pi_s(x;\cdot)$-a.e. $y \in E$ and the identification (18) as $\epsilon > 0$ is arbitrary.

5. Before we show that $R(y) \in \mathcal{M}_e(E)$, let us first establish (ii).

   To this end, we fix $\mu \in \mathcal{M}_e(E)$ and verify that $\mu\Pi_s$ is an ergodic random field on E. Due to (16), the probability measure $\mu\Pi_s$ is spatially homogeneous:

   $$(\mu\Pi_s)(\mathbf{1}_B \circ \theta_a) = \mu(\Pi_s(\mathbf{1}_B \circ \theta_a)) = \mu(\Pi_s(\mathbf{1}_B) \circ \theta_a) = \mu\Pi_s(\mathbf{1}_B)$$

   for all $a \in \mathbb{A}$ and $B \in \mathcal{E}$. Thus, Birkhoff's ergodic theorem implies that $\mu\Pi_s \in \mathcal{M}_e(E)$ whenever

   $$R(y) = \mu\Pi_s \qquad \mu\Pi_s\text{-a.s.}$$

   Since $\mu = R(x)$ for μ-a.e. $x \in E$ it follows from (18) that

   $$\mu\Pi_s(\{y : R(y) = \mu\Pi_s\}) = \int_E \Pi_s(x; \{y : R(y) = R(x)\Pi_s\})\mu(dx) = 1,$$

   and so $\mu\Pi_s \in \mathcal{M}_e(E)$. This shows (ii).

6. We can now easily show that $\Pi_s(x; E_e) = 1$ whenever $x \in E_e$. Indeed, for any $x \in E_e$, we have $R(x) \in \mathcal{M}_e(E)$, and so it follows from (ii) that $R(x)\Pi_s \in \mathcal{M}_e(E)$. Thus, $R(y) = R(x)\Pi_s \in \mathcal{M}_e(E)$ for $\Pi_s(x;\cdot)$-a.e. $y \in E$, due to (18) and therefore $\Pi(x; E_e) = 1$.

This completes the proof. □

Remark 3.2 *Consider the situation analyzed in Föllmer and Horst (2001), i.e., assume that* $S = \mathcal{M}_e(E)$ *and that* $Q(R;\cdot) = \delta_R(\cdot)$. *In this case, each specification* $\gamma^{x,R(x)}$ $(x \in E_e)$ *determines a unique random field* $\Pi(x;\cdot) = \Pi_{R(x)}(x;\cdot)$. *The preceding Proposition yields* $\Pi(x; E_e) = 1$ *and*

$$R(y)(\cdot) = \int_{E_e} \Pi_{R(z)}(x;\cdot)R(x)(dz) = \int_{E_e} \Pi_{R(x)}(x;\cdot)R(x)(dz)$$

for $\Pi(x;\cdot)$*-a.e.* $y \in E_e$. *This proves Theorem 3.1 (i) Föllmer and Horst (2001) for the case of Gibbs measures.*

4 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 E_e defined by the general transition kernel

$$\Pi(x;\cdot) = \int_S \Pi_s(x;\cdot)Q(R(x);ds)$$

introduced in (17). For any random field μ which is concentrated on the set E_e, we denote by $\mathbb{P}_\mu$ the distribution of the chain $\{x_t\}_{t\in\mathbb{N}}$ with initial distribution μ. Since a configuration $x \in E_e$ induces an ergodic empirical field $R(x)$, the *microscopic process* $\{x_t\}_{t\in\mathbb{N}}$ induces $\mathbb{P}_\mu$-a.s. the *macroscopic process* $\{R(x_t)\}_{t\in\mathbb{N}}$ with state space $\mathcal{M}_e(E)$.

4.1 Separating Macroscopic and Microscopic Convergence

Let us show that our spatial law of large numbers for ergodic empirical fields allows us to analyze the microscopic process and the macroscopic process separately. In a first step, we will verify that the dynamics of the macroscopic process can be described by a Markov chain on the state space $\mathcal{M}_h(E)$. In a second step, we are going to show that the microscopic process may be viewed as a Markov chain in a random environment where the distribution of the environment is governed by the initial empirical field $R(x)$.

Let us first analyze the structure of our macroscopic process. To this end, we introduce a mapping $u : \mathcal{M}_h(E) \times S \to \mathcal{M}_h(E)$ by

$$u(R,s) := R\Pi_s(\cdot) = \int \Pi_s(x;\cdot)R(dx). \tag{28}$$

It follows from our Proposition 3.1 that the macroscopic process satisfies

$$R(x_{t+1}) = u(R(x_t), s_t) \qquad \mathbb{P}_\mu\text{-a.s.} \tag{29}$$

Using the law of conditional iterated expectations and (29), it easy to prove the following theorem.

Theorem 4.1 *Under the measure $\mathbb{P}_x = \mathbb{P}_{\delta_x}$ $(x \in E_e)$, the macroscopic process is a Markov chain on the state space $\mathcal{M}_h(E)$ with initial state $R(x)$.*

Its transition operator U acts the class of all bounded measurable functions $f : \mathcal{M}_h(E) \to \mathbb{R}$ according to the formula

$$Uf(R) = \int f(u(R,s))Q(R;ds) \tag{30}$$

Let us now fix a signal sequence $\{s_t\}_{t\in\mathbb{N}}$ and put

$$S^{(t)} := \prod_{i=0}^{t} S \qquad \text{and} \qquad s_t^0 := (s_0, \ldots, s_t).$$

Iterating (29) we obtain

$$R(x_{t+1}) = u_t(R(x), s_t^0) \qquad \mathbb{P}_x\text{-a.s.} \tag{31}$$

where we define the mappings $u_t : \mathcal{M}_h(E) \times S^{(t)} \to \mathcal{M}_h(E)$ recursively by

$$u_0(R, s_0^0) := u(R, s_0) \quad \text{and} \quad u_t(R, s_t^0) := u(u_{t-1}(R, s_{t-1}^0), s_t) \qquad (t \geq 1).$$

Since, conditioned on the environment $\{s_t\}_{t\in\mathbb{N}}$, our macroscopic process evolves almost surely in a deterministic manner, we propose a *random system with complete connections* as a suitable mathematical framework for analyzing the convergence behaviour of the sequence $\{R(x_t)\}_{t\in\mathbb{N}}$. We refer the reader to the books of Iosefescu and Gregorescu (1993), Iosefescu and Theodorescu (1968) or Norman (1972) for a detailed discussion of random systems with complete connections.

Definition 4.2 *Let (M_1, d_{M_1}) be a metric space and $(M_2, \mathcal{M}_2)$ be a measurable space. Let Z denote a stochastic kernel from M_1 to $\mathcal{M}_2$ and let $v : M_1 \times M_2 \to M_1$ be a measurable mapping. Following Iosefescu and Theodorescu (1968), we call the quadruple*

$$\Sigma := ((M_1, d_{M_1}), (M_2, \mathcal{M}_2), Z, v)$$

a homogeneous random system with complete connections (RSCC).

(i) Given an initial value $\xi \in M_1$, a RSCC induces two stochastic processes $\{\xi_t\}_{t\in\mathbb{N}}$ and $\{\zeta_t\}_{t\in\mathbb{N}}$ on the canonical probability space $(\Omega, \mathcal{F}, \mathbb{P}_\xi)$ taking values in M_1 and in M_2, respectively, by

$$\xi_{t+1} = v(\xi_t, \zeta_t) \qquad (\xi_0 = \xi \quad \mathbb{P}_\xi\text{-a.s.})$$

and by

$$\mathbb{P}_\xi(\zeta_t \in \cdot | \xi_t, \zeta_{t-1}, \xi_{t-1}, \zeta_{t-2}, \ldots) = Z(\xi_t; \cdot).$$

These processes are called the associated Markov process and the signal sequence, respectively.

(ii) A random system with complete connections is called a distance-diminishing model, if the transformation $v : M_1 \times M_2 \to M_1$ satisfies the contraction condition

$$d_{M_1}(v(\xi, \zeta), v(\hat{\xi}, \zeta)) \leq \theta d_{M_1}(\xi, \hat{\xi})$$

for some constant $\theta < 1$.

In Section 4.2, we will state conditions on the local specifications $\gamma^{x,s}$ which guarantee that the mapping u in (28) satisfies the contraction condition

$$d(u(R, s), u(\widetilde{R}, s)) \leq \gamma d(R, \widetilde{R}) \qquad (\gamma < 1)$$

with respect to a suitable metric d which induces the weak topology on $\mathcal{M}_h(E)$. In this case, the random system with complete connections

$$\Sigma^* := ((\mathcal{M}_h(E), d), (S, \mathcal{S}), Q, u) \tag{32}$$

is distance-diminishing in the sense of Definition 4.2 (ii). Note that the transition operator of the Markov chain $\{\xi_t\}_{t\in\mathbb{N}}$ associated with Σ^* acts on the class of all bounded measurable functions $g : \mathcal{M}_h(E) \to \mathbb{R}$ according to the formula (30). Thus, for any $x \in E_e$, it follows form Theorem 4.1 that our macroscopic process $\{R(x_t)\}_{t\in\mathbb{N}}$ may be viewed as the Markov chain with initial state $R(x)$ associated with the distance-diminishing random system with complete connections Σ^*. This will allow us to apply a general convergence result in Norman (1972) in order to state conditions on the signal kernel Q which guarantee that the macroscopic process converges in law to a unique equilibrium distribution; see Theorem 4.12 below.

Remark 4.3 *Note that $R(x) \in \mathcal{M}_e(E)$ for all $x \in E_e$. Thus, our macroscopic process may as well be viewed as a Markov chain on the state space $\mathcal{M}_e(E)$. However, we want to apply Theorem 4.2 in Norman (1972) and the Riesz representation theorem. Thus, we have to regard the sequence $\{R(x_t)\}_{t\in\mathbb{N}}$ as a Markov chain on the* compact *metric space $(\mathcal{M}_h(E), d)$.*

Let us now concentrate on the dynamics of the microscopic process. In view of (31), the law of the random variable x_{t+1} is given by

$$\begin{aligned}\Pi^{t+1}(x;\cdot) &= \int_S \cdots \int_S \left[\Pi_{s_0} \cdots \Pi_{s_t}\right](x;\cdot) Q(u_{t-1}(R(x), s_{t-1}^0); ds_t) \cdots \\ &\quad \cdots Q(u_0(R(x), s_0); ds_1) Q(R(x); ds_0).\end{aligned}$$

In this sense, the microscopic process $\{x_t\}_{t\in\mathbb{N}}$ may be viewed as a Markov chain in the random medium $\{s_t\}_{t\in\mathbb{N}}$ where the law of the environment is determined by the empirical field $R(x)$. Combining a perturbation of the Dobrushin-Vasserstein contraction technique with a contraction method from the theory of random systems with complete connections, we shall prove in Theorem 4.19 that convergence in law of the macroscopic process implies that the microscopic process $\{x_t\}_{t\in\mathbb{N}}$ has local asymptotic loss of memory in the sense of Föllmer (1979b).

4.2 Macroscopic Convergence

In this section, we are going to state conditions on the local specifications $\gamma^{x,s}$ and on the signal kernel Q from $\mathcal{M}_h(E)$ to $\mathcal{S}$ which guarantee that the macroscopic process $\{R(x_t)\}_{t\in\mathbb{N}}$, viewed as a Markov chain on the state space $\mathcal{M}_h(E)$, converges in law to a unique equilibrium distribution.

Let us first formulate a weighted *Dobrushin-Vasserstein condition* on the specifications $\gamma^{x,s}$ in order to control the local interactions in the transition kernel Π_s. To this end, we introduce, for any pair $(x,s) \in E \times S$, the matrix $D(x,s) = (D(x,s)_{a,b})$ as the sum of the non-negative powers of the Dobrushin interaction matrix $C(x,s)$ defined in (14), i.e.,

$$D(x,s) := \sum_{n\geq 0} C^n(x,s).$$

We also introduce the vector $b(x,y,s)$ with components

$$b_a(x,y,s) := \frac{1}{2}\int_E \|\gamma_a^{x,s}(\cdot;v) - \gamma_a^{y,s}(\cdot;v)\| \Pi_s(x;dv) \qquad (a \in \mathbb{A}),$$

and, for $a \in \mathbb{A}$, $s \in S$, we define a vector r_a^s by

$$r_{a,b}^s := \sup\left\{\sum_{a\in\mathbb{A}} D_{a,b}(x,s) b_b(x,y,s) : x = y \text{ off } a\right\} \qquad (b \in \mathbb{A}). \tag{33}$$

Note that $r^s_{a,b} = r^s_{a-b,0}$ since our local specifications are translation invariant.

Assumption 4.4 *For a small enough $\eta > 0$, the vectors r^s_a introduced in (33) satisfy*

$$\alpha := \sup_s \sum_a 2^{\eta|a|} r^s_{a,0} < 1. \tag{34}$$

Example 4.5 *Let us return to the local specifications $\gamma^{x,s}$ introduced in (13). We assume that there exists a large enough constant β such that*

$$\sup_{x,s} c_{a,0}(x,s) \le e^{-\beta|a|},$$

i.e., we assume that T_1 is small enough. We also assume that T_2 is sufficiently small so that

$$\frac{1}{2} \sup \left\{ \|\gamma_0^{\theta_{-b}x,s}(\cdot;v) - \gamma_0^{\theta_{-b}y,s}(\cdot;v) \ : \ b \in \mathbb{A}, x = y \text{ off } a, v \in E\| \right\} \le \frac{\hat{\beta}}{2^{ld}+1}$$

for a suitable constant $\hat{\beta} > 0$. In this case, our Assumption 4.4 is satisfied. For details, we refer the reader to Proposition 2.17 in Horst (2000).

Remark 4.6 *Suppose that the transition to a new configuration is made independently by different agents. In this case, the vector r^s_a is given by*

$$r^s_{a,b} = \sup \left\{ \frac{1}{2} \|\pi_s(x;\cdot) - \pi_s(y;\cdot)\| : x = y \text{ off } a - b \right\} \qquad (b \in \mathbb{A}),$$

and our weighted uniform Dobrushin-Vasserstein condition (34) is equivalent to the uniform Dobrushin-Vasserstein condition

$$\alpha_0 := \sup_s \sum_a r^s_{a,0} < 1. \tag{35}$$

If the measures $\Pi_s(x;\cdot)$ have a product form, then the equivalence of (34) and (35) follows from our Assumption 2.5. Such an Assumption also appears in Föllmer and Horst (2001).

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

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

Remark 4.7 *A vector $r = (r_a)_{a\in\mathbb{A}}$ is called an estimate for the random fields μ and ν on E if*

$$|\mu(f) - \nu(f)| \leq \sum_{a\in\mathbb{A}} r_a \Delta_a(f) \tag{36}$$

for any $f \in \mathcal{C}(E)$. Let μ and ν be Gibbs measures on E with respect to the local specifications γ^μ and γ^ν, respectively, and denote by $D(\mu)$ the sum of the non-negative powers of the Dobrushin interaction matrix $C(\mu)$ associated with the random field μ. Then the vector $r = (r_a)_{a\in\mathbb{A}}$ with components

$$r_a = \sum_b D_{a,b}(\mu) b_b(\mu) \tag{37}$$

is such an estimate where the vector $b(\mu)$ is defined by

$$b_a(\mu) := \frac{1}{2} \int_E \|\gamma_a^\mu(\cdot; v) - \gamma_a^\nu(\cdot; v)\| \mu(dv);$$

cf., e.g., Föllmer (1982), Theorem 2.4 or Simon (1993), Theorem V.2.2.

In view of (36) and (37) we have

$$\Delta_a(\Pi_s f) \leq \sum_{b\in\mathbb{A}} r_{a,b}^s \Delta_b(f)$$

for any $f \in \mathcal{C}(E)$. Under Assumption 4.4 we obtain the estimate

$$\Delta(\Pi_s f) \leq \left(\sup_b \sum_a r_{a-b,0}^s \right) \sum_b \Delta_b(f) \leq \alpha \sum_b \Delta_b(f).$$

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

$$\Delta(\Pi_{s_0} \cdots \Pi_{s_t} f) \leq \sum_{a,b} r_{a,b}^{s_1} \Delta_b(\Pi_{s_1} \cdots \Pi_{s_t} f) \leq \alpha^{t+1} \sum_b \Delta_b(f),$$

and so

$$\lim_{t\to\infty} \Delta(\Pi_{s_0} \cdots \Pi_{s_t} f) = 0. \tag{38}$$

In this sense, our microscopic process has local asymptotic loss of memory for any fixed environment $\{s_t\}_{t\in\mathbb{N}}$ if our Assumption 4.4 holds true.

Remark 4.8 *In the case where the transition kernel does not depend on s, the preceding argument shows that the Markov chain Π converges to a unique equilibrium distribution. In our context, however, local asymptotic loss of memory does not necessarily yield the existence of a stationary measure for Π as this transition kernel typically does not have the Feller property.*

Let us introduce a metric d on the class $\mathcal{M}(E)$ of all random fields by

$$d(\mu,\nu) := \sup_{f\in\mathcal{C}(E)} \frac{|\mu(f)-\nu(f)|}{\sum_a 2^{\eta|a|}\Delta_a(f)} \qquad (\mu,\nu\in\mathcal{M}(E)).$$

Lemma 4.9 *(i) The metric d induces the weak topology on $\mathcal{M}(E)$. In particular, $(\mathcal{M}_h(E),d)$ is a compact metric space.*

(ii) Under our Assumption 4.4 the mapping $u:\mathcal{M}_h(E)\times E\to\mathcal{M}_h(E)$ defined in (28) satisfies

$$d(R\Pi_s,\widetilde{R}\Pi_s)\le\alpha d(R,\widetilde{R}) \tag{39}$$

uniformly in $s\in S$. In particular, the random system with complete connections Σ^* introduced in (32) is distance-diminishing in the sense of Definition 4.2 (ii).

Proof: The assertions follow from Propositions 3.3 and 3.8 in Föllmer and Horst (2001); see also Propositions 2.15 and 2.19 in Horst (2000). □

Suppose that $S=\mathcal{M}_h(E)$ and that there exists a constant $\beta<1-\alpha$ such that

$$\sup_x d(\Pi_\mu(x;\cdot),\Pi_\nu(x;\cdot))\le\beta d(\mu,\nu). \tag{40}$$

Under the assumption that $Q(R;\cdot)=\delta_R(\cdot)$ such a condition yields almost sure convergence of the macroscopic process $\{R(x_t)\}_{t\in\mathbb{N}}$ to a unique homogeneous random field μ^* on E. This is Theorem 3.12 in Föllmer and Horst (2001) for the case of Gibbs measures $\Pi_\mu(x;\cdot)$ instead of product measures. Our aim is now to establish convergence in law of macroscopic process without such a restrictive assumption. Instead, we assume that the signal kernel Q from $\mathcal{M}_h(E)$ to $\mathcal{S}$ satisfies the following two conditions.

Assumption 4.10 *(i) The signal kernel Q from $\mathcal{M}_h(E)$ to $\mathcal{S}$ satisfies a uniform Lipschitz condition: There exists a constant $L < \infty$ such that*

$$\sup_{B \in \mathcal{S}} |Q(\mu; B) - Q(\nu; B)| \leq L d(\mu, \nu).$$

(ii) The stochastic kernel Q has a lower bound: There exists a constant $\lambda > 0$ and a probability measure ν on $(S, \mathcal{S})$ such that

$$\inf_{\mu \in \mathcal{M}_h(E)} Q(\mu; \cdot) \geq \lambda \nu(\cdot).$$

Remark 4.11 *Note that our Assumption 4.10 (i) is satisfies whenever the probability distributions $Q(\mu; \cdot)$ have a density f_μ with respect to some measure ν satisfying*

$$|f_\mu(\cdot) - f_\nu(\cdot)| \leq L d(\mu, \nu).$$

We are now ready to state and prove the main result of this section.

Theorem 4.12 *Suppose that the local specifications $\gamma^{x,s}$ are spatially homogeneous and satisfy a Markov property of order l both in their dependence on x and on the boundary condition. If our Assumption 4.4 and 4.10 are satisfied, then the following holds true:*

1. *There exists a unique probability measure μ^* on the class $\mathcal{M}_h(E)$ of all homogeneous random fields on E such that the macroscopic process converges in distribution to μ^*.*

2. *The probability measure μ^* satisfies*

$$\mu^*(\mathcal{M}_e(E)) \in \{0, 1\}.$$

Proof: Let us denote by $(\{\xi_t\}_{t \in \mathbb{N}}, (\hat{\mathbb{P}}_\xi)_{\xi \in \mathcal{M}_h(E)})$ the Markov chain on $\mathcal{M}_h(E)$ associated with the random system Σ^*. Due to Lemma 4.9, $(\mathcal{M}_h(E), d)$ is a compact metric space, and Σ^* is distance diminishing in the sense of Definition 4.2 (ii). Thus, it follows from Theorem 4.2 in Norman (1972) that there exists a unique probability measure μ^* on $\mathcal{M}_h(E)$ such that

$$\lim_{t \to \infty} \int f(\xi_t) d\hat{\mathbb{P}}_\xi = \int f d\mu^*$$

for all $f \in \mathcal{C}(\mathcal{M}_h(E))$ and $\xi \in \mathcal{M}_h(E)$. This shows (i) since, for any starting point $x \in E_e$, our macroscopic process may be viewed as the Markov chain $\{\xi_t\}_{t\in\mathbb{N}}$ on $\mathcal{M}_h(E)$ with initial state $R(x)$. The second assertion follows from Proposition 3.1 (ii) as $u(R,s) \in \mathcal{M}_e(E)$ for all $s \in S$ if $R \in \mathcal{M}_e(E)$. □

Example 4.13 *Consider the local specifications $\gamma^{x,s}$ introduced in (13) and assume that the signal kernel Q satisfies our Assumption 4.10. If T_1 and T_2 are sufficiently small, then the macroscopic process $\{R(x_t)\}_{t\in\mathbb{N}}$ converges in law to a unique equilibrium distribution.*

The model analyzed in this paper may be viewed as a randomized version of the model studied in Föllmer and Horst (2001). Föllmer and Horst (2001) considered the case $Q(R;\cdot) = \delta_R(\cdot)$ and established almost sure convergence of the macroscopic process under the assumption that the local interaction in the kernels Π_s is not too strong and given that the dependence of the specifications on the parameter s is weak enough. In our present setting, the transition to a new state depends on the current empirical field $R(x_t)$ through some random variable s_t whose law depends on $R(x_t)$. As we have seen, this may provide an additional smoothing effect which allows us to establish a convergence result, namely convergence in law of the macroscopic process, without any condition which controls the dependence of the local specifications on the signal s, that is, without such a restrictive contraction condition like (40). We just have to control the local interaction in the kernels Π_s by means of a suitable Dobrushin-Vasserstein condition. Observe, however, that our Assumption 4.10 excludes the case $Q(R(x);\cdot) = \delta_{R(x)}(\cdot)$.

4.3 Microscopic Convergence

In this section, we analyze the asymptotics of the microscopic process $\{x_t\}_{t\in\mathbb{N}}$. We study the interplay between the long behaviour on the macroscopic level of empirical fields and the asymptotic behaviour on the microscopic level. We prove that convergence in law of the *macroscopic* process implies that the microscopic process has local asymptotic loss of memory in the sense of Föllmer (1979b). This means that the distribution of the states of any finite set of agents does, asymptotically, not depend on the starting point

of the microscopic process. In a second step, we prove weak convergence of the Markov chain Π under the additional assumption that the macroscopic process converges in law to a probability measure μ^* which is concentrated on the set $\mathcal{M}_e(E)$.

4.3.1 Asymptotic Loss of Memory

Throughout this section, we assume that the macroscopic process converges in law to a unique equilibrium μ^* on $\mathcal{M}_h(E)$. Our goal is to show that the underlying microscopic process has local asymptotic loss of memory. To this end, we combine a convergence result from the theory of random systems with complete connections with a variant of the Dobrushin-Vasserstein contraction technique. The next result follows from Theorem 2.1.65 in Iosefescu and Theodorescu (1968).

Lemma 4.14 *Suppose that the Markov chain* $(\{\xi_t\}_{t\in\mathbb{N}}, (\hat{\mathbb{P}}_\xi)_{\xi\in\mathcal{M}_h(E)})$ *associated with the random system with complete connections* Σ^* *converges in law to a unique equilibrium distribution* μ^* *on* $\mathcal{M}_h(E)$ *and put* $\hat{\mathcal{F}}_{T,t} = \sigma(s_i : T \le i \le T+t)$. *If the signal kernel* Q *satisfies our Assumption 4.10 (i), then the signal sequence* $\{s_t\}_{t\in\mathbb{N}}$ *associated with* Σ^* *is uniformly ergodic in the strong sense, i.e.,*

$$\lim_{t\to\infty} \sup_T \|\hat{\mathbb{P}}_\xi - \mathbb{P}^*\|_{\hat{\mathcal{F}}_{T,t}} = 0.$$

Here, $\mathbb{P}^*(\cdot) := \int_{\mathcal{M}_h(E)} \mathbb{P}_\xi(\cdot)\mu^*(d\xi)$, *and* $\|\cdot\|_{\hat{\mathcal{F}}_{T,t}}$ *denotes the total variation norm of signed measures on* $\hat{\mathcal{F}}_{T,t}$.

Thus, if the macroscopic process converges in law, then the microscopic process evolves asymptotically in a random environment whose law does not depend on the initial configuration. Moreover, for any *fixed* environment $\{s_t\}_{t\in\mathbb{N}}$, we have

$$\lim_{t\to\infty} \Delta(\Pi_{s_0}\cdots\Pi_{s_t} f) = 0$$

for any $f \in \mathcal{C}(E)$ if our uniform Dobrushin-Vasserstein condition is satisfied; see (38). This allows us to prove the following theorem.

Theorem 4.15 *Suppose that the local specifications $\gamma^{x,s}$ are spatially homogeneous and satisfy our Assumptions 2.6, 2.8 and the weighted Dobrushin-Vasserstein condition 4.4. If the stochastic kernel Q from $\mathcal{M}_h(E)$ to $\mathcal{S}$ satisfies Assumption 4.10, then the microscopic process has local asymptotic loss of memory in the sense of Föllmer (1979b), i.e., we have*

$$\lim_{t\to\infty} \sup_{x,y} \left|\Pi^t(x;B) - \Pi^t(x;B)\right| = 0 \tag{41}$$

for each $A \in \mathcal{A}$ and $B \in \mathcal{E}_A$. Here $\mathcal{E}_A$ denotes the σ-field generated by the projections $x \mapsto x^a$ $(a \in A)$.

In order to prepare the proof of Theorem 4.15, we introduce some additional notation. For $t, T \in \mathbb{N}$ we put

$$s_t^T := (s_T, \ldots, s_{T+t}).$$

For $x \in E_e$ and $t, n, m \in \mathbb{N}$ we denote by

$$Q^{n,m}(x;\cdot) \qquad \text{and} \qquad Q_t^{n,m}(x;\cdot) \tag{42}$$

the law of the random variable s_m^n and the conditional law of the random variables s_m^n, given the signal vector s_t^0, respectively, under $\mathbb{P}_x$. For a given signal vector s_t^T, we consider the transition kernels $\Pi_{s_t^T}$ on E defined by

$$\Pi_{s_t^T}(x;\cdot) := \left[\Pi_{s_T} \cdots \Pi_{s_{T+t}}\right](x;\cdot). \tag{43}$$

Let us first prove (41) under the additional assumption that the stochastic kernels Π_s take the product form (10). In this case, the mapping $s_t^0 \mapsto \sup_x \Pi_{s_t^0}(x;B)$ is measurable for any $B \in \mathcal{E}_A$ $(A \in \mathcal{A})$ as we just have to determine the supremum of finitely many measurable functions.

Proof of Theorem 4.15: Independent Transitions

Let us fix $A \in \mathcal{A}$ and $B \in \mathcal{E}_A$. In terms of the notation introduced in (42) and (43) we have for any initial configuration $x \in E_e$ that

$$\begin{aligned} &\Pi^{T+t+1}(x;B) \\ = &\int \cdots \int \Pi_{s_t^T}(y;B) Q_{T-1}^{T,t}(x;ds_t^T) \Pi_{s_{T-1}^0}(x;dy) Q^{0,T-1}(x;ds_{T-1}^0), \end{aligned}$$

due to Proposition 3.1. Let us now fix $t \in \mathbb{N}$. The quantity $\Pi^{t+T+1}(x; B)$ is bounded above by

$$\int \cdots \int \{\sup_z \Pi_{s_t^T}(z; B)\} Q_{T-1}^{T,t}(x; ds_t^T) Q^{0,T-1}(x; ds_{T-1}^0)$$

and bounded below by

$$\int \cdots \int \{\inf_z \Pi_{s_t^T}(z; B)\} Q_{T-1}^{T,t}(x; ds_t^T) Q^{0,T-1}(x; ds_{T-1}^0).$$

This yields the following estimate:

$$\begin{aligned} & \sup_{x,y} |\Pi^{T+t+1}(x; B) - \Pi^{T+t+1}(y; B)| \\ \leq \; & \sup_{x,y} \left| \int \{\sup_z \Pi_{s_t^T}(z; B)\} \left[Q^{T,t}(x; ds_t^T) - Q^{T,t}(y; ds_t^T) \right] \right| \qquad (44) \\ & + \sup_y \int \{\sup_z \Pi_{s_t^T}(z; B) - \inf_z \Pi_{s_t^T}(z; B)\} Q^{T,t}(y; ds_t^T). \qquad (45) \end{aligned}$$

We are going to analyze the quantities (44) and (45) separately.

1. Since $Q^{T,t}(x; \cdot) - Q^{T,t}(y; \cdot)$ is a signed measure on $(\prod_{i=0}^t S, \otimes_{i=0}^t \mathcal{S})$ with total mass zero, we can estimate (44) by

   $$\sup_{x,y} \|Q^{T,t}(x; \cdot) - Q^{T,t}(y; \cdot)\|.$$

   Due to Theorem 4.12 and Lemma 4.14, the signal sequence associated with the random system Σ^* is uniformly ergodic in the strong sense. Thus, there exist measures Q_t^* on $(\prod_{i=0}^t S, \otimes_{i=0}^t \mathcal{S})$ such that

   $$\lim_{T \to \infty} \sup_{x,t} \|Q^{T,t}(x; \cdot) - Q_t^*(\cdot)\| = 0. \qquad (46)$$

   In particular, we have that

   $$\lim_{T \to \infty} \sup_{t,x,y} \left| \int \sup_z \Pi_{s_t^T}(z; B) \left[Q^{T,t}(x; ds_t^T) - Q^{T,t}(y; ds_t^T) \right] \right| = 0.$$

2. Let us now analyze the integral in (45). It follows from (46) that

$$\begin{aligned}\lim_{T\to\infty}\sup_y &\left|\int\{\sup_z \Pi_{s_t^T}(z;B)-\inf_z \Pi_{s_t^T}(z;B)\}Q^{T,t}(y;ds_t^T)\right| \\ = &\left|\int\{\sup_z \Pi_{s_t^0}(z;B)-\inf_z \Pi_{s_t^0}(z;B)\}Q_t^*(ds_t^0)\right| \qquad (47)\end{aligned}$$

For any fixed environment $\{s_t\}_{t\in\mathbb{N}}$, our uniform Dobrushin-Vasserstein-Condition (Assumption 4.4) implies that

$$\sup_{x,y}|\Pi_{s_t^0}(x;B)-\Pi_{s_t^0}(y;B)| \quad\le\quad c\alpha^{t+1} \qquad (48)$$

for some constants $c=c(|A|)<\infty$ and $\alpha<1$ which do neither depend on $B\in\mathcal{E}_A$ nor on the environment $\{s_t\}_{t\in\mathbb{N}}$; cf. (38). Thus, for any given $\epsilon>0$, there exists large enough $t_0, T_0\in\mathbb{N}$ such that

$$\sup_{x,y,s_t^0}\left|\Pi_{s_t^0}(x;B)-\Pi_{s_t^0}(y;B)\right|<\epsilon \qquad (t\ge t_0)$$

and such that

$$\sup_{x,y,t}\|Q^{T,t}(x;\cdot)-Q_t^*(\cdot)\|<\epsilon \qquad (T\ge T_0).$$

Hence, for all $t\ge t_0$ and $T\ge T_0$ we have the following estimate:

$$(45)\le\sup_{x,y}\|Q^{T,t}(x;\cdot)-Q_t^*(\cdot)\|+\sup_{s_t^0}\Delta(\Pi_{s_t^0}(\cdot;B))\le 2\epsilon. \qquad (49)$$

Thus, for any $t\ge t_0$ and for all $T\ge T_0$, we have that

$$\sup_{x,y\in S_0}|\Pi^{t+T+1}(x;B)-\Pi^{t+T+1}(y;B)|<3\epsilon.$$

Since $\epsilon>0$ is arbitrary, the assertion follows. □

Let us now consider the general transition kernel Π introduced in (17). For any two configurations $v,x\in E$ and for all local sets $A\in\mathcal{A}$, we denote by $\gamma_A^{x,s}(\cdot;v)$ the probability measure $\gamma_A^{x,s}(\cdot;(v^a)_{a\in A^c})$ on C^A with boundary condition $(v^a)_{a\in A^c}$ on A^c. Moreover, we put

$$\Pi_{s_t^0,v}^{A_t^0}(x;\cdot):=\int_E\cdots\int_E\gamma_{A_0}^{x,s_0}(dx_1;v)\cdots\gamma_{A_{t-1}}^{x_{t-1},s_{t-1}}(dx_t;v)\gamma_{A_t}^{x_t,s_t}(\cdot;v),$$

where $A_0 \supset A_1 \supset \ldots$ are local sets.

In order to avoid a problem of measurability, it will be convenient to use the following approximation result which follows immediately from our Assumptions 2.6 and 2.8 together with Theorem 8.23 and Remark 8.26 in Georgii (1989).

Lemma 4.16 *For any local set $A \in \mathcal{A}$, for each $t \in \mathbb{N}$ and for all $\epsilon > 0$, there exits local sets $A_0 \supset A_1 \supset \cdots \supset A_t$ such that*

$$\sup_{B \in \mathcal{E}_A} |\Pi^{A_t^0}_{s_t^0, v}(x; B) - \Pi_{s_t^0}(x; B)| < \epsilon$$

uniformly in $x \in E$, in $s_t^0 \in S^{(t)}$ and in the boundary condition $v \in E$.

Remark 4.17 *Observe that, for any given configuration $v \in E$, for all $t \in \mathbb{N}, A \in \mathcal{A}, B \in \mathcal{E}_A$ and for any fixed local sets $A_0, A_1, \ldots, A_t$, the mapping*

$$s_t^0 \mapsto \sup_x \Pi^{A_t^0}_{s_t^0, v}(x; B)$$

is measurable.

We are now ready to prove Theorem 4.15.

Proof of Theorem 4.15: Interactive Transitions

Let us fix $A \in \mathcal{A}$, $B \in \mathcal{E}_A$, $t \in \mathbb{N}$ and $\epsilon > 0$. Due to Lemma 4.16, we can choose local sets $A_0 \supset \cdots \supset A_t$ such that

$$\sup_{v, x, B, s_t^T} |\Pi_{s_t^T}(x; B) - \Pi^{A_t^0}_{s_t^T, v}(x; B)| < \epsilon.$$

Thus, uniformly in $B \in \mathcal{E}_A$, the quantity $\Pi^{t+T+1}(x; B)$ is bounded above by

$$\int \cdots \int \{\sup_z \Pi^{A_t^0}_{s_t^T, v}(z; B)\} Q^{T,t}_{T-1}(x; ds_t^T) Q^{0,T-1}(x; ds^0_{T-1}) + \epsilon$$

and bounded below by

$$\int \cdots \int \{\inf_z \Pi^{A_t^0}_{s_t^T, v}(z; B)\} Q^{T,t}_{T-1}(x; ds_t^T) Q^{0,T-1}(x; ds^0_{T-1}) - \epsilon.$$

By analogy with the case of product kernels we have the following estimate:

$$\begin{aligned}
& \sup_{x,y} |\Pi^{T+t+1}(x;B) - \Pi^{T+t+1}(y;B)| \\
\leq\ & \sup_{x,y} \left| \int \cdots \int \{\sup_z \Pi^{A_t^0}_{s_t^T,v}(z;B)\} Q^{T,t}_{T-1}(x;ds_t^T) Q^{0,T-1}(x;ds^0_{T-1}) \right. \\
& \left. - \int \cdots \int \{\inf_z \Pi^{A_t^0}_{s_t^T,v}(z;B)\} Q^{T,t}_{T-1}(y;ds_t^T) Q^{0,T-1}(y;ds^0_{T-1}) \right| + 2\epsilon \\
\leq\ & \sup_{x,y} \left| \int \{\sup_z \Pi^{A_t^0}_{s_t^T,v}(z;B)\} \left[Q^{T,t}(x;ds_t^T) - Q^{T,t}(y;ds_t^T) \right] \right| \\
& + \sup_y \int \{\sup_z \Pi^{A_t^0}_{s_t^T,v}(z;B) - \inf_z \Pi^{A_t^0}_{s_t^T,v}(z;B)\} Q^{T,t}(y;ds_t^T) + 2\epsilon.
\end{aligned}$$

Using a tedious but straightforward 3-ϵ-argument one can now easily show that

$$\lim_{T\to\infty} \sup_{y,t} \int \{\sup_z \Pi^{A_t^0}_{s_t^T,v}(z;B) - \inf_z \Pi^{A_t^0}_{s_t^T,v}(z;B)\} = 0.$$

This yields our assertion by analogy with the case of product kernels. □

4.4 Existence and Uniqueness of Invariant Measures

Let us now state conditions which guarantee the existence of a stationary distribution for the microscopic process.

Lemma 4.18 *Let μ^* be any stationary distribution for the Markov chain associated with the random system with complete connections Σ^* defined by (32) which is concentrated on the set $\mathcal{M}_e(E)$. Then the microscopic process is stationary under the law $\mathbb{P}_{\nu^*}$, where the random field ν^* on E is given by*

$$\nu^*(\cdot) := \int_{\mathcal{M}_e(E)} \mu(\cdot)\mu^*(d\mu).$$

Proof: In order to prove our assertion, it is enough to show that

$$\mathbb{P}_{\nu^*}[x_0 \in B] = \mathbb{P}_{\nu^*}[x_1 \in B]$$

for any $B \in \mathcal{E}_A$ ($A \in \mathcal{A}$). Since $R(x) = \nu$ for ν-a.e. $x \in E$ whenever $\nu \in \mathcal{M}_e(E)$, we have

$$\mathbb{E}_{\nu^*}[R(x_0)] = \nu^*.$$

Since μ^* is concentrated on the set $\mathcal{M}_e(E)$, the macroscopic process is stationary under $\mathbb{P}_{\nu^*}$. Thus, we have

$$\mathbb{E}_{\nu^*}[R(x_0)] = \mathbb{E}_{\nu^*}[R(x_1)],$$

and so it is enough to show that

$$\mathbb{P}_{\nu^*}[x_1 \in B] = \mathbb{E}_{\nu^*}[R(x_1)(B)].$$

Due to Proposition 3.1 (ii), we have $\mu\Pi_s \in \mathcal{M}_e(E)$ if $\mu \in \mathcal{M}_e(E)$. In particular, $\mu\Pi_s = R(x)$ for $\mu\Pi_s$-a.e. $x \in E$, and so

$$\begin{aligned}
\mathbb{E}_{\nu^*}[R(x_1)(B)] &= \int_{E_e}\int_S\int_{E_e} R(x_1)(B)\Pi_{s_0}(x_0;dx_1)Q(R(x_0);ds_0)\nu^*(dx_0) \\
&= \int_{\mathcal{M}_e(E)}\int_S\int_E R(x_1)(B)\mu\Pi_{s_0}(dx_1)Q(\mu;ds_0)\mu^*(d\nu) \\
&= \int_{\mathcal{M}_e(E)}\int_S \mu\Pi_{s_0}(B)Q(\mu;ds_0)\mu^*(d\mu) \\
&= \mathbb{P}_{\nu^*}[x_1 \in B].
\end{aligned}$$

Thus, ν^* is an invariant measure for the stochastic kernel Π, and so the microscopic process are stationary under the law $\mathbb{P}_{\nu^*}$. □

We are now ready to prove the main result of this paper.

Theorem 4.19 *Suppose that the local specifications $\gamma^{x,s}$ are spatially homogeneous and satisfy our Assumptions 2.6 and 2.8. Assume moreover, that the signal kernel Q from $\mathcal{M}_h(E)$ to $\mathcal{S}$ satisfies Assumption 4.10. If the unique stationary distribution μ^* for the Markov chain associated with the random system with complete connections Σ^* satisfies $\mu^*(\mathcal{M}_e(E)) = 1$, then there exists a unique stationary probability measure μ for the microscopic process. For any starting point $x \in E_e$, the sequence $\{\Pi^t(x;\cdot)\}_{t\in\mathbb{N}}$ converges weakly to μ.*

Proof: It follows from Theorem 4.12 that the macroscopic process converges in law to the unique equilibrium distribution μ of the Markov chain associated with the random system with complete connections Σ^*. Thus, we deduce from Lemma 4.18 that there exists a stationary distribution for

our microscopic process. Uniqueness of the stationary distribution and weak convergence of the sequence $\{\Pi^t(x;\cdot)\}_{t\in\mathbb{N}}$ follows from Theorem 4.15. $\square$

Example 4.20 *Consider the mean-field type interaction in Example 2.1, and assume that the interaction if spatially homogeneous. In this case, it is easy to show that the unique invariant measure of the Markov chain associated with the random system Σ^* is concentrated on the set $\mathcal{M}_e(E)$ whenever the dependence of the probability law $\pi_s(x^a;\cdot)$ on x^a is not too strong.*

Example 4.21 *Let us return to the individual transition laws π_s defined in (6). Assume that the signal kernel Q from $\mathcal{M}_h(E)$ to $[0,1]$ satisfies Assumption 4.10. We also assume that the uniform Dobrushin condition (35) is satisfied, i.e., that the dependence of the probability distribution $\pi_s(x;\cdot)$ on the configuration x is not too strong. In this case, Markov chain $\{x_t\}_{t\in\mathbb{N}}$ with transition kernel*

$$\Pi(x;\cdot) = \prod_{a\in\mathbb{A}} \pi_s(\theta_a x;\cdot) Q(R(x);ds)$$

has local asymptotic loss of memory. If, moreover, the unique stationary measure of the random system Σ^* is concentrated on the set $\mathcal{M}_e(E)$, the process $\{x_t\}_{t\in\mathbb{N}}$ converges in distribution to unique random field μ. It remains an open problem, however, to give sufficient conditions for $\mu^*(\mathcal{M}_e(E)) = 1$.

Acknowledgement

This paper is based on Chapter 3 of my Ph.D. thesis. I would like to thank Peter Bank, Hans Föllmer and Alexander Schied for valuable discussions. Support of Deutsche Forschungsgemeinschaft via SFB 373, "Quantification and Simulation of Economic Processes", is gratefully acknowledged.

References

Brock, W. A., and C. Hommes (1997): "A Rational Route to Randomness," *Econometrica*, 65 (5), 1059–1095.

Dobrushin, R. L. (1968): "Description of Random Fields by Means of Conditional Probabilities and the Conditions Governing its Regularity," *Theor. Probab. Appl.*, 15, 458–486.

Föllmer, H. (1979a): "Macroscopic Convergence of Markov Chains on Infinite Product Spaces," in *Colloquium on Random Fields*, pp. 363–371, Esztergom, Ungarn. Math. Soc. Janos Bolyai.

——— (1979b): "Tail Structure of Markov Chains on Infinite Product Spaces," *Z. Wahrscheinlichkeitstheorie Verw. Geb.*, 50, 273–285.

——— (1982): "A Covariance Estimate for Gibbs Measures," *Journal of Functional Analysis*, 46 (3), 387–395.

Föllmer, H., and U. Horst (2001): "Convergence of Locally and Globally Interaction Markov Chains," Preprint, Humboldt-Universität zu Berlin; submitted to *Stochastic Processes Appl.*

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.

Iosefescu, M., and S. Gregorescu (1993): *Dependence With Complete Connections And Its Applications*. Cambridge University Press, Cambridge.

Iosefescu, M., and M. Theodorescu (1968): *Random Processes and Learning*. Springer-Verlag, Berlin.

Kirman, A. (1998): "On the Transitory Nature of Gurus," Working paper, EHESS et Université de Marseille III.

Lebowitz, J. L., C. Maes, and E. R. Speer (1990): "Statistical Mechanics of Probabilistic Cellular Automata," *J. Stast. Phys.*, 59, 117–170.

Norman, F. M. (1972): *Markov Processes and Learning Models*. Academic Press, New York.

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 Peredaci Informacii*, 5, 64–72.

Ulrich Horst
Institut für Mathematik – Bereich Stochastik
Humboldt-Universität zu Berlin
Unter den Linden 6
D-10099 Berlin
email:horst@mathematik.hu-berlin.de
Phone: +49 (0) 30 2093 1459 Fax: +49 (0) 30 2093 5848