Probab. Theory Relat. Fields 102, 165–197 (1995)

Probability Theory and Related Fields

Finite and infinite systems of interacting diffusions

J.T. Cox[1,*], Andreas Greven[2,], Tokuzo Shiga[3,***]**

[1] Mathematics Department, Syracuse University, Syracuse, New York 13244, USA
(e-mail jtcox@mailbox.syr.edu)
[2] Mathematisches Institut, Universität Erlangen-Nürnberg, Bismarckstrasse 1 1/2, D-91054 Erlangen, Germany
(e-mail greven@namu01.gwdg.de)
[3] Department of Applied Physics, Tokyo Institute of Technology, Oh-okayama, Meguro, Tokyo 152, Japan
(e-mail shiga@ap.titech.ac.jp)

Received: 20 December 1993 / In revised form: 15 December 1994

Summary. We study the problem of relating the long time behavior of finite and infinite systems of locally interacting components. We consider in detail a class of linearly interacting diffusions $x(t) = \{x_i(t), i \in \mathbb{Z}^d\}$, in the regime where there is a one-parameter family of nontrivial invariant measures. For these systems there are naturally defined corresponding finite systems, $x^N(t) = \{x_i^N(t), i \in \Lambda_N\}$, with $\Lambda_N = (-N, N]^d \cap \mathbb{Z}^d$. Our main result gives a comparison between the laws of $x(t_N)$ and $x^N(t_N)$ for times $t_N \to \infty$ as $N \to \infty$. The comparison involves certain mixtures of the invariant measures for the infinite system.

Mathematics Subject Classification (1991): 60K35

1 Introduction

The purpose of this paper is to study the approximation of large finite systems of interacting components by corresponding infinite systems, and vice versa, as considered in [CG1] and extended in [CG2] and [DG1]. We give a new class of examples of the phenomena found in these papers, which includes some models of interest in mathematical biology. The models considered in the cited papers all had rather special properties which made them mathematically tractable: duality, "family" independence, or "mean field" independence. This is not the case with the class of models treated here. Consequently, we show that the "finite systems

* Partly supported by the U.S. Army Research Office through the Mathematical Sciences Institute of Cornell University, by the National Science Foundation, and by the National Security Agency
** Research supported in part by the DFG
*** Partly supported by S.R.63540155 of Japan Ministry of Education

scheme" put forth in these earlier papers is not merely an artifact of such special properties, but describes a type of behavior which, in our opinion, holds more broadly. We carry this point further in [CG4] by proving an abstract theorem relating the behavior of finite and infinite systems of interacting components. See [CG1] for a list of references on the study of finite versus infinite systems. We start by defining infinite systems of interacting diffusions.

Infinite systems of interacting diffusions

Let $I \subset \mathbb{R}$ be an interval. The *infinite system* $x(t) = \{x_i(t), i \in \mathbb{Z}^d\} \in I^{\mathbb{Z}^d}$ is a Markov process defined through the following system of stochastic differential equations:

$$\text{(1.1a)} \quad \begin{aligned} &dx_i(t) = \left[\textstyle\sum_{j\in\mathbb{Z}^d} a(i,j)x_j(t) - x_i(t)\right]dt + \sqrt{g(x_i(t))}\,dw_i(t), \quad i \in \mathbb{Z}^d\,, \\ &x(0) \in \mathbb{E}\,. \end{aligned}$$

The ingredients in the above system are as follows:

(1.1b) A matrix $a(i,j)$ which is irreducible and satisfies

$$a(i,j) \geq 0, \quad a(i,j) = a(0,j-i) \quad \forall i,j \in \mathbb{Z}^d, \quad \sum_{j\in\mathbb{Z}^d} a(0,j) = 1.$$

(1.1c) A function $g : I \to \mathbb{R}^+$ which vanishes at finite endpoints of I, is locally Lipschitz, and satisfies

$$\limsup_{|\theta|\to\infty} \frac{g(\theta)}{\theta^2} < \infty.$$

(1.1d) A collection $\{w_i(t), i \in \mathbb{Z}^d\}$ of independent one-dimensional Brownian motions.

(1.1e) The state space $\mathbb{E} \subset I^{\mathbb{Z}^d}$, which is defined via a suitable norm. Let $\{\gamma_i, i \in \mathbb{Z}^d\}$ be a strictly positive, summable sequence such that for some finite constant Γ,

$$\sum_{i\in\mathbb{Z}^d} \gamma_i\, a(i,j) \leq \Gamma\gamma_j, \quad j \in \mathbb{Z}^d.$$

For $x \in I^{\mathbb{Z}^d}$ let $\|x\|_1 = \sum_{i\in\mathbb{Z}^d} \gamma_i|x_i|$, and define $\mathbb{E} = \{x \in I^{\mathbb{Z}^d} : \|x\|_1 < \infty\}$. We endow $\mathbb{E}$ with the topology of component-wise convergence.

Remark. Since $a(i,j)$ is irreducible and translation invariant, if $x \in \mathbb{E}$ then $\sigma_i x \in \mathbb{E}$ for all $i \in \mathbb{Z}^d$, where σ_i is the shift by i, $(\sigma_i x)_j = x_{i+j}$. Thus one can consider shift invariant probability measures on $\mathbb{E}$. We note that if μ is a shift invariant probability measure on $I^{\mathbb{Z}^d}$ with $E^\mu|x_0| < \infty$, then μ is automatically supported in $\mathbb{E}$, and hence is a shift invariant measure on $\mathbb{E}$. Of course, $\mathbb{E} = I^{\mathbb{Z}^d}$ if I is bounded.

Although Theorem 3.2 of [SS] is more restrictive than the present situtation, a modification of its proof, with truncation by stopping times, yields the following. For every $x(0) \in \mathbb{E}$ there exists a unique strong solution $x(t)$ of (1.1) such that

$$P(x(t) \text{ is an } \mathbb{E}\text{–valued continuous function of } t \geq 0) = 1.$$

The solution defines a Markov process $(P^x, x(t))$ taking values in $\mathbb{E}$, and a Markov semigroup $S(t)$ acting on $C_b(\mathbb{E})$, the collection of all bounded continuous functions on $\mathbb{E}$, such that

$$S(t)f - f = \int_0^t S(s)\mathfrak{A} f\, ds, \quad f \in C_0^2(\mathbb{E}).$$

Here $C_0^2(\mathbb{E})$ is the set of all C^2 functions on $\mathbb{E}$ which depend on only finitely many coordinates, and have the property that $\mathfrak{A}f$ is bounded, where

$$\mathfrak{A} f(x) = \frac{1}{2}\sum_{i\in\mathbb{Z}^d} g(x)\frac{\partial^2 f}{\partial x_i^2} + \sum_{i\in\mathbb{Z}^d}[\sum_{j\in\mathbb{Z}^d}(a(i,j) - \delta(i,j))x_j]\frac{\partial f}{\partial x_i},$$

with $\delta(i,j) = 1$ if $i = j$, and 0 otherwise. If we define $\|x\|_2^2 = \sum \gamma_i x_i^2$, then the set $\mathbb{E}' = \{x \in \mathbb{E} : \|x\|_2 < \infty\}$ has the property that for $x(0) \in \mathbb{E}'$,

$$P(x(t) \text{ is an } \mathbb{E}'\text{–valued } \|\cdot\|_2 \text{ -continuous function of } t \geq 0) = 1,$$

and the induced semigroup is Feller. The following special cases of (1.1) have been studied in the literature.

Example 1. $I = [0, 1]$: $g(\theta) = c\theta(1 - \theta)$ (Wright-Fisher stepping stone model [S1]), and $g(\theta) = c\theta^2(1 - \theta)^2$ (Ohta-Kimura model [NS]).

Example 2. $I = [0, \infty)$: $g(\theta) = c\theta$ (branching diffusion or "super random walk" [D2]).

Example 3. $I = (-\infty, \infty)$: $g(\theta) = c$ (critical Ornstein-Uhlenbeck process [D1]).

Example 4. $I = [0, \infty)$: $g(\theta) = c\theta^2$ (scalar field in a non-stationary random potential [S4]).

We refer the reader to the indicated papers for more information on these examples.

As shown in [CG3] and [SS], the symmetrized kernel

$$\widehat{a}(i,j) = \frac{a(i,j) + a(j,i)}{2}$$

plays a fundamental role in describing the ergodic behavior of the interacting diffusion $x(t)$. Let $\widehat{a}_t(i,j)$ be the continuous time kernel $\widehat{a}_t(i,j) = e^{-t}\sum_{n=0}^{\infty} t^n\, \widehat{a}^{(n)}(i,j)/n!$, and let $\widehat{A}(i,j) = \int_0^\infty \widehat{a}_{2s}(i,j)\,ds$. In order to use second moment techniques *we assume from now on the condition*

$$\limsup_{|\theta|\to\infty} \frac{g(\theta)}{|\theta|^2} < \frac{1}{\widehat{A}(0,0)} \quad . \tag{1.2}$$

We need some additional notation to review the basic ergodic theory of the infinite systems. Let $\mathcal{T}(\mathbb{E})$ denote the collection of probability measures on $\mathbb{E}$ which are shift invariant, and let $\mathcal{T}_p(\mathbb{E})$ be all $\mu \in \mathcal{T}(\mathbb{E})$ such that $\langle \mu, |x_0|^p \rangle < \infty$, where in general $\langle \mu, f \rangle = \int f d\mu$. $\mathcal{L}$ denotes law, and δ_θ denotes the unit point mass at the element $x_i \equiv \theta$. Let $\to_p$ denote convergence in probability, let $\overset{\text{fdd}}{\Longrightarrow}$ denote convergence of finite dimensional distributions, and let $\Rightarrow$ denote weak convergence. In particular, for continuous processes $Z_n(\cdot)$ and $Z(\cdot)$, $Z_n(\cdot) \Rightarrow Z(\cdot)$ means that the probability laws induced by $Z_n(\cdot)$ on the continuous path space converges weakly to that induced by $Z(\cdot)$ as $n \to \infty$. For any probability measure μ on $\mathbb{E}$ we will write $\mu S(t)$ for $\mathcal{L}(x(t))$ when $\mathcal{L}(x(0)) = \mu$. $\mathcal{I}$ denotes the set of invariant measures for $x(t)$, i.e., all probability measures μ on $\mathbb{E}$ such that $\mu S(t) = \mu$ for all $t \geq 0$.

The behavior of $x(t)$ depends on whether or not $\widehat{a}(i,j)$ is recurrent or transient. In the transient case we have the following, taken from [CG3] and [S4].

Theorem 0. *Assume $\widehat{a}(i,j)$ is transient.*

(a) *For $\theta \in I$ and $\mathcal{L}(x(0)) = \delta_\theta$, the weak limit $\nu_\theta = \lim_{t\to\infty} \mathcal{L}(x(t))$ exists, is an element of $\mathcal{T}_2(\mathbb{E})$, is associated and mixing, and satisfies $\langle \nu_\theta, x_0 \rangle = \theta$.*

(b) *For $\theta \in I$, if $\mathcal{L}(x(0)) \in \mathcal{T}_1(\mathbb{E})$ is shift-ergodic, with $Ex_0(0) = \theta$, then $\mathcal{L}(x(t)) \Rightarrow \nu_\theta$ as $t \to \infty$.*

(c) *The set of extreme points of $\mathcal{I} \cap \mathcal{T}_1(\mathbb{E})$ is exactly $\{\nu_\theta, \theta \in I\}$.*

Remark. When necessary we will write ν_θ^g for ν_θ to indicate the dependence of ν_θ on g.

The picture for recurrent $\widehat{a}(i,j)$ is not nearly as complete, except for some special cases and the general compact I case, in which the phenomenon of *clustering* occurs. We refer the reader to the papers [S1], [S2], [NS] and [CG3] for more details on this.

Finite systems of interacting diffusions

For $N = 1, 2, \ldots$ let $\Lambda_N = (-N, N]^d \cap \mathbb{Z}^d$ be viewed as a *torus*, and let $\mathbb{E}_N = I^{\Lambda_N}$. We define an $\mathbb{E}_N$–valued Markov process $x^N(t) = \{x_i^N(t), i \in \Lambda_N\}$ via the system

$$\begin{aligned} dx_i^N(t) &= \left[\sum_{j\in\Lambda_N} a^N(i,j) x_j^N(t) - x_i^N(t) \right] dt + \sqrt{g(x_i^N(t))}\, dw_i(t), \quad i \in \Lambda_N, \\ x^N(0) &\in \mathbb{E}_N, \end{aligned} \tag{1.3}$$

where $a^N(i,j) = \sum_{k\in\mathbb{Z}^d} a(i, j + 2Nk)$, $i,j \in \Lambda_N$. We let $S^N(t)$ denote the corresponding transition semigroup operators.

We view the processes $x^N(t)$ as *finite versions* of the infinite system $x(t)$, since it is easily seen that given $x(0) \in \mathbb{E}$, if $x_i^N(0) = x_i(0)$ for all $i \in \Lambda_N$ and all N, then for fixed $t > 0$,

$$\mathcal{L}(x^N(t)) \Rightarrow \mathcal{L}(x(t)) \text{ as } N \to \infty.$$

To be precise about the meaning of this convergence, we introduce the periodic extension operators $\pi_N : \mathbb{E}_N \to \mathbb{E}$, $(\pi_N x^N)_j = x_i^N$ where $i \in \Lambda_N$, $i = j \bmod (2N)$. We also let π_N denote the induced operator mapping probability measures on $\mathbb{E}_N$ to probability measures on $\mathbb{E}$. If μ^N is a probability measure on $\mathbb{E}_N$, $N = 1, 2, \ldots$, and μ is a probability measure on $\mathbb{E}$, we write $\mu^N \Rightarrow \mu$ as $N \to \infty$ to mean $\pi_N \mu^N \Rightarrow \mu$ as $N \to \infty$.

If $\widehat{a}(i,j)$ is transient, the long-term behavior of the finite systems differs drastically from the long-term behavior of the corresponding infinite system (as given in Theorem 0a). For instance, if $I \subset \mathbb{R}^+$ and $x_i^N(0) \equiv \theta$, then for fixed N, $\sum_{i \in \Lambda_N} x_i^N(t)$ is a nonnegative martingale, and must converge a.s. as $t \to \infty$. From this fact it is easy to see that as $t \to \infty$,

$$\mathcal{L}(x^N(t)) \Rightarrow (1-\theta)\delta_0 + \theta\delta_1 \quad \text{(Example 1)},$$
$$\mathcal{L}(x^N(t)) \Rightarrow \delta_0 \quad \text{(Example 2)}.$$

Using Gaussian techniques it can be shown that (suitably interpreted) as $t \to \infty$,

$$\mathcal{L}(x^N(t_N)) \Rightarrow \frac{1}{2}\delta_{-\infty} + \frac{1}{2}\delta_{+\infty} \quad \text{(Example 3)}.$$

To obtain a more precise picture of the asymptotic behavior of the finite systems we will compare the behavior of $x(t)$ and $x^N(t)$ as *both* N and t tend to infinity, using the framework of the *finite systems scheme* of [CG1].

Ingredients of the finite systems scheme

In order to state our results we define the following objects.

(1.4a) The time scale $\beta_N = (2N)^d$.

(1.4b) The empirical densities $\Theta^N(t) = |\Lambda_N|^{-1} \sum_{i \in \Lambda_N} x_i^N(t)$.

(1.4c) The rescaled process of empirical densities $Z_N(t) = \Theta^N(t\beta_N)$.

(1.4d) The diffusion $Z(t)$ on I, defined for the case $\widehat{a}(i,j)$ transient, by

$$dZ(t) = \sqrt{g^*(Z(s))}\, dw(s), \quad Z(0) = \rho,$$

where $w(t)$ is a Brownian motion on $\mathbb{R}$ and g^* is the function $g^*(\theta) = E^{\nu_\theta^g} g(x_0)$. (The fact that this stochastic differential equation has a unique weak solution will follow from Lemma 2.12 of Sect. 2 below.) The probability transition function of $Z(t)$ will be denoted $Q_t(\rho, d\theta)$.

(1.4e) The empirical measures of the finite systems

$$U_N(t) = |\Lambda_N|^{-1} \sum_{i \in \Lambda_N} \delta_{\sigma_i^N x^N(t)},$$

where σ_i^N is the shift by i on Λ_N, $(\sigma_i^N x)_j = x_k$, $k = (i+j) \bmod (2N)$.

Main results

We treat only the case $\widehat{a}(i,j)$ is transient in this paper, the recurrent case will be contained in [CGS]. Theorem 1 analyzes the behavior of the finite systems from a *global* point of view with the empirical densities $U_N(t)$. Theorem 2 takes a more *local* viewpoint, and is the direct analogue of results proved for branching random walk, the voter model and the contact process in [CG1]. Our proofs require certain random walk estimates (see Proposition 2.1 below) which hold for all genuinely d-dimensional random walk kernels, $d \geq 3$. The estimates hold also for transient random walk kernels in $d \leq 2$ which possess certain regularity properties. For simplicity we will consider only the $d \geq 3$ case.

Theorem 1. *Assume $d \geq 3$. Suppose that* $\sup_N E\langle U_N(0), |x_0|^p\rangle < \infty$ *for some $p > 2$, and for some random variable Z_0, $\mathcal{L}(\Theta_N(0)) \Rightarrow \mathcal{L}(Z_0)$ as $N \to \infty$. Then as $N \to \infty$,*

$$Z_N(\cdot) \Rightarrow Z(\cdot), \quad Z(0) = Z_0\,, \tag{1.5}$$

and

$$\{U_N(t\beta_N); t > 0\} \overset{\text{fdd}}{\Longrightarrow} \{\nu_{Z(t)}; t > 0\}. \tag{1.6}$$

Theorem 2. *Assume $d \geq 3$. Suppose that for some $p > 2$, $\mathcal{L}(x^N(0)) \in \mathcal{T}_p(\mathbb{E}_N)$,* $\sup_N E|x_0^N(0)|^p < \infty$, *and for some $\rho \in I$, $\Theta^N(0) \to_p \rho$ as $N \to \infty$. Let $t_N \uparrow \infty$ and $t_N/\beta_N \to s \in [0,\infty)$, and in the case $t_N/N^2 \not\to \infty$, assume also that $\mathcal{L}(x^N(0)) \Rightarrow$ some ergodic element of $\mathcal{T}_2(\mathbb{E})$. Then*

$$\mathcal{L}(\Theta^N(t_N)) \Rightarrow \mathcal{L}(Z(s)), \quad Z(0) = \rho\,, \tag{1.7}$$

and

$$\mathcal{L}(x^N(t_N)) \Rightarrow \int_I Q_s(\rho, d\theta)\nu_\theta. \tag{1.8}$$

Remark. Another way to look at the phenomena described in Theorems 1 and 2 is to consider time averages. Let $\phi : \mathbb{E} \to \mathbb{R}^+$ be a continuous function depending on finitely many coordinates. Under the assumptions of Theorem 2, if $l_N \to \infty$, $l_N = o(\beta_N)$, then one can prove

$$l_N^{-1} \int_{t\beta_N}^{t\beta_N + l_N} \phi(x^N(s))\, ds \Rightarrow \int_I Q_t(\rho, d\theta)\langle \nu_\theta, \phi\rangle.$$

This shows that the phenomena described in the theorems are observable in the statistical sense.

Remark. It is possible to take $p = 2$ in Theorems 1 and 2 above at the cost of strengthening (1.2) to

$$\limsup_{|\theta|\to\infty} \frac{g(\theta)}{|\theta|^2} = 0. \tag{1.2$'$}$$

In [CGS] we will study properties of the mapping $g \to g^*$, and discuss the special role of $g(x) = cx(1-x)$, $g(x) = cx$, and $g(x) = c$ in this context. The remainder of this paper is devoted to the proofs of Theorems 1 and 2. Section 2 is a lengthy section containing numerous technical preliminaries. Theorem 1 is proved in Sect. 3, Theorem 2 is proved in Sect. 4.

2 Technical preparations

We collect here various technical results that we will need. These include random walk estimates, moment formulae, and a formulation of the basic coupling, which is our most important tool. We apply the coupling to derive some regularity properties of the invariant measures ν_θ^g.

Random walk estimates

We first state a proposition containing some results on the asymptotic behavior of the random walk kernels $a_t^N(i,j)$ and $\widehat{a}_t^N(i,j)$. The results were proved in [C] and [CG1] for simple symmetric random walk.

Lemma 2.1.
(a) *If* $t_N/N^2 \to \infty$ *as* $N \to \infty$, *then* $\sup_{t\geq t_N} \sup_{i,j\in\Lambda_N} (2N)^d \, |a_t^N(i,j) - (2N)^{-d}| \to 0$.

(b) *If* $d \geq 3$, *and* $\lambda > 0$, *then* $\lim_{N\to\infty} \int_0^\infty e^{-\lambda t/(2N)^d}\, \widehat{a}_{2t}^N(i,j)\,dt = \lambda^{-1} + \widehat{A}(i,j)$.

(c) *If* $d \geq 3$, *and* $T(N)/\beta_N \to s \in (0,\infty)$ *as* $N \to \infty$, *then*

$$\lim_{N\to\infty} \int_0^{T(N)} \widehat{a}_{2t}^N(i,j)\,dt = \widehat{A}(i,j) + s.$$

Proof. The fact that

$$\sup_{t\geq t_N} \sup_{j\in\Lambda_N} (2N)^d \, |a_t^N(0,j) - (2N)^{-d}| \to 0 \tag{2.1}$$

holds if $a(i,j)$ is the transition kernel of simple symmetric random walk in $\mathbb{Z}^d$, $d \geq 2$, was established in [C] (see the Proposition on page 1341 there). The proof used a refined local limit expansion, and is easily modified to show that (2.1) holds for $d \geq 1$ for any irreducible random walk kernel $a(i,j)$ of finite range (i.e., for some $K < \infty$, $a(i,j) = 0$ if $|i-j| > K$). Assuming now that $a(i,j)$ is not of finite range, we choose $K < \infty$ large enough so that with $\gamma = \sum_{|k| \le K} a(0,k)$, $0 < \gamma < 1$, the kernel

$$p(i,j) = \gamma^{-1} a(i,j) 1_{\{|i-j| \le K\}}$$

is irreducible. Letting

$$q(i,j) = (1-\gamma)^{-1} a(i,j) 1_{\{|i-j| > K\}},$$

we have $a(i,j) = \gamma p(i,j) + (1-\gamma) q(i,j)$, and on the torus Λ_N, $a^N(i,j) = \gamma p^N(i,j) + (1-\gamma) q^N(i,j)$. It follows that

$$a_t(i,j) = p_{\gamma t}\, q_{(1-\gamma)t}(i,j) = \sum_{k \in \mathbb{Z}^d} p_{\gamma t}(i,k) q_{(1-\gamma)t}(k,j). \tag{2.2}$$

and

$$a_t^N(i,j) = p_{\gamma t}^N\, q_{(1-\gamma)t}^N(i,j) = \sum_{k \in \Lambda_N} p_{\gamma t}^N(i,k) q_{(1-\gamma)t}^N(k,j). \tag{2.3}$$

Now put $\epsilon_N = \sup_{t \ge \gamma t_N} \sup_{j \in \Lambda_N} (2N)^d \, |p_t^N(0,j) - (2N)^{-d}|$. By (2.1), $\epsilon_N \to 0$ as $N \to \infty$, and therefore, uniformly in $t \ge t_N$, $i,j \in \Lambda_N$,

$$\begin{aligned} (2N)^d \, |a_t^N(i,j) - (2N)^{-d}| &= (2N)^d \, \Big| \sum_{k \in \Lambda_N} p_{\gamma t}^N(i,k) q_{(1-\gamma)t}^N(k,j) - (2N)^{-d} \Big| \\ &\le (2N)^d \sum_{k \in \Lambda_N} |p_{\gamma t}^N(i,k) - (2N)^{-d}| q_{(1-\gamma)t}^N(k,j) \\ &\le \sum_{k \in \Lambda_N} \epsilon_N \, q_{(1-\gamma)t}^N(k,j) = \epsilon_N, \end{aligned}$$

which proves part (a).

For (b), we apply a result of [MW], which says that if $\hat{\phi}(u) = \sum_{j \in \mathbb{Z}^d} \widehat{a}(0,j) \times e^{i(u \cdot j)}$, $u \in \mathbb{R}^d$, where $i^2 = -1$, then for $\lambda > 0$

$$\int_0^\infty e^{-\lambda t} \, \widehat{a}_t^N(0,j) \, dt = (2N)^{-d} \sum_{k \in \Lambda_N} \frac{\exp(-i(2\pi j \cdot k)/2N)}{1 + \lambda - \hat{\phi}(\pi k/N)}.$$

By separating out the $k = 0$ term we obtain

$$\begin{aligned} \int_0^\infty e^{-\lambda t/(2N)^d} \, \widehat{a}_t^N(0,j) \, dt &= \lambda^{-1} + (2N)^{-d} \sum_{\substack{k \in \Lambda_N \\ k \neq 0}} \frac{\exp(-i(2\pi j \cdot k)/2N)}{1 + \lambda/(2N)^d - \hat{\phi}(\pi k/N)} \\ &\to \lambda^{-1} + \int_{(-1,1)^d} \frac{\exp(i 2\pi (u \cdot j))}{1 - \hat{\phi}(\pi 2\pi u)} \, du \end{aligned}$$

as $N \to \infty$. It is here that the assumption $d \ge 3$ has been used to simplify passage to the limit. Since $\hat{\phi}$ is real, P7.5 of [Sp] implies that for some positive constant C, $1 - \hat{\phi}(u) \ge C|u|^2$ on $(-\pi, \pi)^d$. For $d \ge 3$, this is enough to apply the dominated convergence theorem. (For transient $\widehat{a}(i,j)$ in $d \le 2$ some additional assumption on the regularity of $\hat{\phi}$ near zero is needed for this step.) Finally, the last integral above equals $\widehat{A}(0,j)$.

For (c), let $t_N/N^2 \to \infty$, $t_N/N^d \to 0$ as$N \to \infty$. By the estimate of part (a),

$$\int_{t_N}^{\infty} e^{-\lambda t/(2N)^d} \widehat{a}_{2t}^N(0,j)\,dt = \int_{t_N}^{\infty} \frac{1+o(1)}{(2N)^d} e^{-\lambda t/(2N)^d}\,dt$$
$$= \frac{1+o(1)}{\lambda} \exp(-\lambda t_N/(2N)^d),$$

which tends to λ^{-1} as $N \to \infty$. Therefore by (b), it must be the case that

$$\int_0^{t_N} e^{-\lambda t/(2N)^d} \widehat{a}_{2t}^N(0,j)\,dt \to \widehat{A}(0,j). \tag{2.4}$$

Now for $0 \le t \le t_N$, $1 \ge \exp(-\lambda t/(2N)^d) \ge \exp(-\lambda t_N/(2N)^d) \to 1$ as $N \to \infty$. This observation and (2.4) prove

$$\int_0^{t_N} \widehat{a}_{2s}^N(0,j) \to \widehat{A}(0,j).$$

Furthermore, if $T(N)/(2N)^d \to s \in (0,\infty)$, the estimates of part (a) easily give

$$\int_{t_N}^{T(N)} \widehat{a}_{2t}^N(0,j)\,dt \to s$$

and we are done. □

Moments

Explicit expressions for the first and second moments of the coordinates $x_i(t)$ can be readily obtained from (1.1) and the Itô calculus (see [NS],[CG3], and [S4]). If $\sup_i E|x_i(0)| < \infty$, respectively $\sup_i E|x_i(0)|^2 < \infty$, then

$$Ex_i(t) = \sum_{k\in\mathbb{Z}^d} a_t(i,k)Ex_k(0),$$

$$Ex_i(t)x_j(t) = \sum_{k,l\in\mathbb{Z}^d} a_t(i,k)a_t(j,l)Ex_k(0)x_l(0) \tag{2.5}$$
$$+ \int_0^t \sum_{k\in\mathbb{Z}^d} a_{t-s}(i,k)a_{t-s}(j,k)Eg(x_k(s))ds.$$

If $\mathcal{L}(x(0)) \in \mathcal{T}_1(\mathbb{E})$, respectively $\mathcal{T}_2(\mathbb{E})$, then

$$Ex_i(t) = Ex_0(0)$$

(2.6)

$$Ex_i(t)x_j(t) = \sum_{k,l\in\mathbb{Z}^d} a_t(i,k)a_t(j,l)Ex_k(0)x_l(0) + \int_0^t \widehat{a}_{2(t-s)}(i,j)Eg(x_0(s))ds.$$

Of course, similar formulae hold for $x^N(t)$. The following provides us with some needed uniformity.

Lemma 2.2. *Let $\widehat{a}(i,j)$ be transient.*

(a) *If μ is a probability measure on $\mathbb{E}$ such that $C = \sup_i \langle \mu, x_i^2 \rangle < \infty$, then there is a finite constant M which depends on μ only through C such that* $\sup_{t\geq 0} \sup_{i\in\mathbb{Z}^d} E^\mu |x_i(t)|^2 \leq M$.

(b) *There is a $p > 2$ such that if $\mu \in \mathcal{T}_p(\mathbb{E})$, then $\sup_{t\geq 0} E^\mu |x_0(t)|^p < \infty$.*

(c) *Assume $d \geq 3$. If $\mu^N \in \mathcal{T}_2(\mathbb{E}_N)$, and $C = \sup_N \langle \mu^N, x_0^2 \rangle < \infty$, then for any $T < \infty$ there is a finite constant M which depends on the μ^N only through C such that*

$$\sup_N \sup_{0\leq t\leq T\beta_N} E^{\mu^N} (x_0^N(t))^2 \leq M.$$

(d) *Assume $d \geq 3$. There is a $p > 2$ such that if $\mu^N \in \mathcal{T}_p(\mathbb{E}_N)$, and $\sup_N \langle \mu^N, |x_0^N|^p \rangle < \infty$, then for any $T < \infty$,*

$$\sup_N \sup_{0\leq t\leq T\beta_N} E^{\mu^N} |x_0^N(t)|^p < \infty.$$

Proof of Lemma 2.2. For $\mu \in \mathcal{T}_2(\mathbb{E})$, (a) is Lemma 2.1 of [S4]. More generally, by (1.2) there are finite constants b, c, with $0 < b < \widehat{A}(0,0)^{-1}$, such that $g(\theta) \leq c + b\theta^2$ for all $\theta \in I$. Letting $f_i(t) = E^\mu x_i^2(t)$ and $\bar{f}(t) = \sup\{f_i(s) : 0 \leq s \leq t, i \in \mathbb{Z}^d\}$, it is routine to see that $\bar{f}(t) < \infty$ for $t > 0$. For $u \leq t$, (2.5) implies that

$$\begin{aligned} f_i(u) &= \sum_{k,l\in\mathbb{Z}^d} a_u(i,k)a_u(i,l)E^\mu x_k x_l + \int_0^u \sum_k a_{u-s}(i,k)a_{u-s}(i,k)E g(x_k(s))\, ds \\ &\leq C + \int_0^u \sum_k a_{u-s}(i,k)a_{u-s}(i,k)[c + bEx_k^2(s)]\, ds \\ &\leq C + c\widehat{A}(0,0) + b\widehat{A}(0,0)\bar{f}(t). \end{aligned}$$

Taking the supremeum over $u \leq t$ and rearranging gives

$$\bar{f}(t) \leq \frac{C + c\widehat{A}(0,0)}{1 - b\widehat{A}(0,0)},$$

which proves (a).

For (b) define

$$K = E^\mu |x_0|^p, \quad f(t) = E^\mu |x_0(t)|^p, \quad \bar{f}(t) = \sup\{|f(s)|;\ 0 \leq s \leq t\}.$$

It follows from (1.1) that

$$x_i(t) = \sum_{j\in\mathbb{Z}^d} a_t(i,j)x_j(0) + \int_0^t \sum_{j\in\mathbb{Z}^d} a_{t-s}(i,j)\sqrt{g(x_j(s))}dw_j(s), \quad i \in \mathbb{Z}^d.$$

This fact is easily established by applying Itô's formula to the stochastic process

$$\sum_{j\in\mathbb{Z}^d} a_{t-s}(i,j)x_j(s), \quad 0 \le s \le t, \ (t \text{ fixed})$$

and then integrating both sides. Next, it follows from (4.1) of [B] that for a continuous L_p-martingale M_t, if $p \ge 2$, for all $0 \le u \le t$,

$$E|M_u|^p \le (p-1)^p E([M]_t)^{p/2}.$$

Now for fixed t, letting $M_u = \int_0^u \sum_{j\in\mathbb{Z}^d} a_{t-s}(i,j)\sqrt{g(x_j(s))}dw_j(s)$, $\{M_u, 0 \le u \le t\}$ is a martingale, and hence

$$E|M_t|^p \le (p-1)^p E\left[\int_0^t \sum_{j\in\mathbb{Z}^d} a_{t-s}(i,j)^2 g(x_j(s))ds\right]^{p/2}.$$

Using these observations we have

$$\begin{aligned} E|x_i(t) - \sum_{j\in\mathbb{Z}^d} a_t(i,j)x_j(0)|^p &= E|\int_0^t \sum_{j\in\mathbb{Z}^d} a_{t-s}(i,j)\sqrt{g(x_j(s))}dw_j(s)|^p \\ &\le (p-1)^p E\left[\int_0^t \sum_{j\in\mathbb{Z}^d} a_{t-s}(i,j)^2 g(x_j(s))ds\right]^{p/2}. \end{aligned}$$

If we let $\alpha = 2 - 4/p$, and $2/p + 1/q = 1$, then $\alpha q = 2 = (2-\alpha)p/2$. Thus the right-hand side above equals

$$\begin{aligned} &(p-1)^p E\left[\int_0^t \sum_{j\in\mathbb{Z}^d} a_{t-s}(i,j)^\alpha a_{t-s}(i,j)^{2-\alpha} g(x_j(s))ds\right]^{p/2} \\ &\le (p-1)^p \left[\int_0^t \sum_{j\in\mathbb{Z}^d} a_{t-s}(i,j)^2\, ds\right]^{p/2q} \int_0^t \sum_{j\in\mathbb{Z}^d} a_{t-s}(i,j)^2 E[g(x_j(s))]^{p/2} ds \\ &= (p-1)^p \left[\int_0^t \widehat{a}_{2s}(0,0)ds\right]^{\frac{p}{2}-1} \int_0^t \widehat{a}_{2s}(0,0)E[g(x_0(t-s))]^{p/2}\, ds \end{aligned}$$

where we have used Hölder's inequality. By assumption (1.2) there are finite constants b, c $0 < b < \widehat{A}(0,0)^{-1}$ such that for $p \in (2,3]$, $|g(\theta)|^{p/2} \le b^{p/2}|\theta|^p + c$ for all $\theta \in I$. Finally, we need the elementary inequality: for $p > 2$ there exists $C_p < \infty$ such that

$$(\alpha+\beta)^p \le (p-1)\alpha^p + C_p\beta^p \text{ for } \alpha, \beta > 0.$$

By combining the previous results we obtain, for $u \le t$,

$$E|x_0(u)|^p \le KC_p + (p-1)^{p+1}\left[\int_0^u \widehat{a}_{2s}(0,0)ds\right]^{\frac{p}{2}-1}$$
$$\int_0^u \widehat{a}_{2s}(0,0)[b^{p/2}E|x_0(u-s)|^p + c]\,ds$$
$$\le KC_p + (p-1)^{p+1}\widehat{A}(0,0)^{\frac{p}{2}-1}[b^{p/2}\widehat{A}(0,0)\bar{f}(t) + c\widehat{A}(0,0)].$$

Taking the supremum over $u \le t$ gives

$$\bar{f}(t) \le KC_p + (p-1)^{p+1}\widehat{A}(0,0)^{\frac{p}{2}-1}[b^{p/2}\widehat{A}(0,0)\bar{f}(t) + c\widehat{A}(0,0)].$$

Now if $p > 2$ is sufficiently close to 2, then for all t,

$$\bar{f}(t) \le \frac{KC_p + c(p-1)^{p+1}\widehat{A}(0,0)}{1-(p-1)^{p+1}\widehat{A}(0,0)^{\frac{p}{2}-1}b^{p/2}\widehat{A}(0,0)} < \infty.$$

We turn to the proof of (d), which is somewhat more involved. With

$$K = \sup_N\langle\mu^N, |x_0|^p\rangle, \quad f^N(t) = E^{\mu^N}(|x_0^N(t)|^p, \quad \bar{f}^N(t) = \sup\{|f^N(s)|; 0 \le s \le t\},$$

we proceed as before to obtain

$$E|x_0^N(t)|^p \le KC_p + (p-1)^{p+1}\left[\int_0^t \widehat{a}_{2s}^N(0,0)ds\right]^{\frac{p}{2}-1}.$$
$$\int_0^t \widehat{a}_{2s}^N(0,0)[b^{p/2}E|x_0^N(t-s)|^p + c]\,ds.$$

Letting $J^N(t) = \int_0^t \widehat{a}_{2s}^N(0,0)ds$, we have for $u \le t$,

$$f^N(u) \le KC_p + (p-1)^{p+1}J^N(u)^{\frac{p}{2}-1}\int_0^u \widehat{a}_{2s}^N(0,0)[b^{p/2}\bar{f}^N(u-s) + c]ds.$$

Taking the supremum over $u \le t$ gives

$$\bar{f}^N(t) \le KC_p + (p-1)^{p+1}J^N(t)^{\frac{p}{2}-1}\int_0^u \widehat{a}_{2s}^N(0,0)(b^{p/2}\bar{f}^N(t-s) + c)ds.$$

For $r > 0$, define

$$I^N(r) = \int_0^\infty e^{-rt}\widehat{a}_{2t}^N(0,0)\,dt \quad \text{and} \quad F^N(r) = \int_0^{2T\beta_N} e^{-rt}\bar{f}^N(t)dt.$$

By integrating the previous inequality we find that

$$F^N(r) \le KC_p/r + (p-1)^{p+1}J^N(2T\beta_N)^{\frac{p}{2}-1}\{I^N(r)(b^{p/2}F^N(r) + c/r)\}.$$

A little rearrangement yields

$$\frac{r}{\beta_N}F^N(\frac{r}{\beta_N}) \le \frac{KC_p + (p-1)^{p+1}J^N(2T\beta_N)^{\frac{p}{2}-1}cI^N(r/\beta_N)}{1-b^{p/2}(p-1)^{p+1}J^N(2T\beta_N)^{\frac{p}{2}-1}I^N(r/\beta_N)},$$

provided the denominator is positive. To see that this is possible for appropriate choices of p and r, we note that by Lemma 2.1, $J^N(2T\beta_N) \to \widehat{A}(0,0) + 2T$, $I^N(r/\beta_N) \to r^{-1} + \widehat{A}(0,0)$, and also

$$I^N(r/\beta_N) \geq \int_0^{2T\beta_N} e^{-rs/\beta_N}\widehat{a}^N_{2s}(0,0)ds \geq e^{-2Tr}J^N(2Tr\beta_N).$$

Therefore,

$$\limsup_{N\to\infty} b^{p/2}(p-1)^{p+1}J^N(2T\beta_N)^{\frac{p}{2}-1}I^N(r/\beta_N)$$
$$\leq b^{p/2}(p-1)^{p+1}(e^{2Tr}(\widehat{A}(0,0)+r^{-1}))^{\frac{p}{2}-1}(\widehat{A}(0,0)+r^{-1}).$$

Recalling that $b\widehat{A}(0,0) < 1$, we can first choose $r < \infty$ and then $p > 2$ such that the right-hand side above is strictly less than one. Thus for some $M < \infty$, $\sup_N F^N(r/\beta_N)/\beta_N \leq M$. Consequently,

$$\int_0^{2T} e^{-rt}\bar{f}^N(t\beta_N)dt \leq M.$$

Using the monotonicity of $\bar{f}^N$, we get $\sup_N \bar{f}^N(T\beta_N) \leq e^{2rT}M/T$, and we are done.

The proof of (c) is similar, but shorter, so we omit it. □

Coupling

Our primary tool is the coupling of [CG3] and [S4], and we will need several forms of it. The first is a coupling of two versions of the infinite system, $x(t)$ and $y(t)$. Given the $\mathbb{E}\times\mathbb{E}$–valued pair $(x(0),y(0))$, the bivariate process $(x(t),y(t))$ is defined by

$$\begin{aligned} dx_i(t) &= \Big[\textstyle\sum_{j\in\mathbb{Z}^d} a(i,j)x_j(t) - x_i(t)\Big]\,dt + \sqrt{g(x_i(t))}\,dw_i(t), \quad i\in\mathbb{Z}^d, \\ dy_i(t) &= \Big[\textstyle\sum_{j\in\mathbb{Z}^d} a(i,j)y_j(t) - y_i(t)\Big]\,dt + \sqrt{g(y_i(t))}\,dw_i(t), \quad i\in\mathbb{Z}^d, \end{aligned} \tag{2.7}$$

where *one* set of Brownian motions is used for *both* coordinates. We note that the existence and uniqueness results of [SS] hold for (2.7). For $\theta\in I$, let $\mathcal{T}_2^\theta(\mathbb{E})$ be the set of all $\mu\in\mathcal{T}_2(\mathbb{E})$ such that

$$\sum_{i\in\mathbb{Z}^d} a_t(0,j)x_j \to \theta \text{ in } L_2(\mu) \text{ as } t\to\infty.$$

We will show in the subsection on L_2-theory that $\mathcal{T}_2^\theta(\mathbb{E})$ contains all $\mu\in\mathcal{T}_2(\mathbb{E})$ which are shift-ergodic and have $\langle\mu,x_0\rangle = \theta$. The following was proved in [CG3] and [S4].

Proposition 2.3. *Let $(x(t), y(t))$ be the bivariate process defined by (2.7).*

(a) *If $\mathcal{L}(x(0), y(0)) \in \mathcal{T}_1(\mathbb{E} \times \mathbb{E})$, then $E|x_i(t) - y_i(t)|$ is* nonincreasing *in t.*

(b) *If in addition, $\mathcal{L}(x(0)), \mathcal{L}(y(0)) \in \mathcal{T}_2^\theta(\mathbb{E})$ for some $\theta \in I$, and $\widehat{a}(i,j)$ is transient, then $E|x_i(t) - y_i(t)| \to 0$ as $t \to \infty$.*

(c) *If $\widehat{a}(i,j)$ is transient, the measures ν_θ of Theorem 0 belong to $\mathcal{T}_2^\theta(\mathbb{E})$.*

(d) *If $\widehat{a}(i,j)$ is transient, and $\mu \in \mathcal{T}_2^\theta(\mathbb{E})$, then $\mu S(t) \Rightarrow \nu_\theta$ as $t \to \infty$.*

Versions of the finite systems may be coupled in a like manner. Given the $\mathbb{E}_N \times \mathbb{E}_N$–valued pair $(x^N(0), y^N(0))$, the bivariate process $(x^N(t), y^N(t))$ is defined by

$$
\begin{aligned}
dx_i^N(t) &= \left[\textstyle\sum_{j\in\Lambda_N} a^N(i,j)x_j^N(t) - x_i^N(t)\right] dt + \sqrt{g(x_i^N(t))}\,dw_i(t), \quad i \in \Lambda_N, \\
&\qquad (2.8) \\
dy_i^N(t) &= \left[\textstyle\sum_{j\in\Lambda_N} a^N(i,j)y_j^N(t) - y_i^N(t)\right] dt + \sqrt{g(y_i^N(t))}\,dw_i(t), \quad i \in \Lambda_N.
\end{aligned}
$$

As in Proposition 2.3(a), if $\mathcal{L}(x^N(0), y^N(0)) \in \mathcal{T}_1(\mathbb{E}_N \times \mathbb{E}_N)$, then $E|x_i^N(t) - y_i^N(t)|$ is *nonincreasing* in t.

We also need to couple finite and infinite systems. Given $(x(0), x^N(0)) \in \mathbb{E} \times \mathbb{E}_N$, construct the bivariate process $(x(t), x^N(t))$ via

$$
\begin{aligned}
dx_i(t) &= \left[\textstyle\sum_{j\in\mathbb{Z}^d} a(i,j)x_j(t) - x_i(t)\right] dt + \sqrt{g(x_i(t))}\,dw_i(t), \quad i \in \mathbb{Z}^d, \\
&\qquad (2.9) \\
dx_i^N(t) &= \left[\textstyle\sum_{j\in\Lambda_N} a^N(i,j)x_j^N(t) - x_i^N(t)\right] dt + \sqrt{g(x_i^N(t))}\,dw_i(t), \quad i \in \Lambda_N.
\end{aligned}
$$

Comparison estimates

We now use this coupling to show that at fixed times t, *uniformly* over a certain class of initial states, the laws of $x(t)$ and $x^N(t)$ are close for large N.

Proposition 2.4. *Let $(x(t), x^N(t))$ be the bivariate process defined by (2.9). For $C < \infty$ let $\mathcal{V}_C^N$ be the collection of all probability measures $\bar{\mu}$ on $\mathbb{E} \times \mathbb{E}_N$ such that $\bar{\mu}(x_i = x_i^N, i \in \Lambda_N) = 1$ and $\sup_i \langle \bar{\mu}, x_i^2 \rangle \le C$.*

(a) *For fixed $i \in \mathbb{Z}^d$ and $t > 0$, $\sup_{\bar{\mu}\in\mathcal{V}_C^N} E^{\bar{\mu}}|x_i(t) - x_i^N(t)| \to 0$ as $N \to \infty$. In particular, there exists a sequence $\{\ell_N\}$ depending only on C, $\ell_N \uparrow \infty$, such that*

$$\sup_{\bar{\mu}\in\mathcal{V}_C^N} E^{\bar{\mu}}|x_i(\ell_N) - x_i^N(\ell_N)| \to 0 \text{ as } N \to \infty.$$

(b) *For fixed $t > 0$, $\sup_{\bar{\mu}\in\mathcal{V}_C^N} |\Lambda_N|^{-1} \sum_{i\in\Lambda_N} E^{\bar{\mu}}|x_i(t) - x_i^N(t)| \to 0$ as $N \to \infty$. In particular, there exists a sequence $\{\ell_N\}$ depending only on C, $\ell_N \uparrow \infty$, such that*

$$\sup_{\bar{\mu}\in\mathcal{V}_C^N} |\Lambda_N|^{-1} \sum_{i\in\Lambda_N} E^{\bar{\mu}}|x_i(\ell_N) - x_i^N(\ell_N)| \to 0 \text{ as } N \to \infty.$$

Proof. Let $\Delta_i^N(t) = x_i^N(t) - x_i(t)$ for $i \in \Lambda_N$. As in [CG3], Itô's lemma can be applied to $|x_i^N - x_i|$, and we obtain (with sgn the signum function)

$$\begin{aligned}\frac{d}{dt}E|\Delta_i^N(t)| &= E\left[\operatorname{sgn}(\Delta_i^N(t))\Big(\sum_{j\in\Lambda_N} a^N(i,j)x_j^N(t) - \sum_{j\in\mathbb{Z}^d} a(i,j)x_j(t)\Big)\right] - E|\Delta_i^N(t)|,\\ &\le \sum_{j\in\Lambda_N} a^N(i,j)E[|\Delta_j^N(t)| - |\Delta_i^N(t)|] + \sum_{j\in\Lambda_N} |a^N(i,j) - a(i,j)|E|x_j(t)|\\ &+ \sum_{j\notin\Lambda_N} a(i,j)E|x_j(t)|.\end{aligned}$$

By Lemma 2.2(a) there exists $M < \infty$ depending only on C such that $E|x_i(t)| \le M$ for all t. Letting $\epsilon_i^N = \sum_{j\notin\Lambda_N} a(i,j)$, we have

$$\frac{d}{dt}E|\Delta_i^N(t)| \le \sum_{j\in\Lambda_N} a^N(i,j)E[|\Delta_j^N(t)| - |\Delta_i^N(t)|] + 2M\epsilon_i^N \quad i \in \Lambda_N.$$

From this it follows that

$$\text{(2.10)} \qquad E|\Delta_i^N(t)| \le \sum_{j\in\Lambda_N} a_t^N(i,j)E|\Delta_j^N(0)| + 2M\int_0^t \sum_{j\in\Lambda_N} a_{t-s}^N(i,j)\epsilon_j^N\, ds.$$

By assumption, $\Delta_i^N(0) = 0$ for all $i \in \Lambda_N$. Furthermore, writing $\Lambda_{N/2}$ for $\Lambda_{[N/2]}$, where $[N/2]$ is the greatest integer $\le N/2$, $j + \Lambda_{N/2} \subset \Lambda_N$ for $j \in \Lambda_{N/2}$. Therefore,

$$\epsilon_0^{N/2} = \sum_{k\notin\Lambda_{N/2}} a(0,k) = \sum_{k\notin j+\Lambda_{N/2}} a(j,k) \ge \sum_{k\notin\Lambda_N} a(j,k) = \epsilon_j^N.$$

Thus,

$$\begin{aligned}E|\Delta_i^N(t)| &\le 2M\int_0^t \sum_{j\in\Lambda_N} a_{t-s}^N(i,j)\epsilon_j^N\, ds\\ &\le 2M\int_0^t \sum_{j\notin\Lambda_{N/2}} a_s^N(i,j)\, ds + 2Mt \sup_{j\in\Lambda_{N/2}} \epsilon_j^N\\ &\le 2M\int_0^t \sum_{j\notin\Lambda_{N/2}} a_s(i,j)\, ds + 2Mt\epsilon_0^{N/2}\end{aligned}$$

and this last expression tends to 0 as $N \to \infty$ for fixed t, proving (a).

For (b), fix $t > 0$ and $\delta > 0$ and choose $K < \infty$ such that $\epsilon_0^K < \delta/2Mt$. Observe that if $i \in \Lambda_N$, and $i + \Lambda_K \subset \Lambda_N$, then translation invariance of the kernel $a(i,j)$ gives

$$\epsilon_i^N \le \sum_{j\notin i+\Lambda_K} a(i,j) = \epsilon_0^K.$$

Using (2.10) and the above,

$$\begin{aligned}|\Lambda_N|^{-1}\sum_{i\in\Lambda_N} E|\Delta_i^N(t)| &\le 2Mt|\Lambda_N|^{-1}\sum_{i\in\Lambda_N}\epsilon_i^N\\ &\le \delta + 2Mt|\Lambda_N|^{-1}\sum_{i\in\Lambda_N} 1\{(i+\Lambda_K)\cap\Lambda_N^c\neq\emptyset\}\\ &\le \delta + 2MtdK/N\,.\end{aligned}$$

Let $N \to \infty$ and $\delta \to 0$ to obtain (b). □

L_2-Theory

For $x \in \mathbb{E}$ and $n = 1, 2, \ldots$, define

$$D_n(x) = |\Lambda_n|^{-1}\sum_{i\in\Lambda_n} x_i\,.$$

The L_2 ergodic theorem asserts that if $\mu \in \mathcal{T}_2(\mathbb{E})$, then the spatial density $D(x) = \lim_{n\to\infty} D_n(x)$ exists as an $L_2(\mu)$ limit. Furthermore, for each $\mu \in \mathcal{T}_2(\mathbb{E})$ there is a unique finite measure λ on $\Pi = (-\pi, \pi]^d$ such that

$$\langle \mu, (x_j - \theta)(x_k - \theta)\rangle = \int_\Pi \exp(iu\cdot(j-k))\,\lambda(du),$$

where $\theta = \langle \mu, x_0\rangle$. λ is called the *spectral measure* of μ. If $\mu \in \mathcal{T}_2(\mathbb{E})$ and $\langle \mu, x_0\rangle = \theta$, and λ is the corresponding spectral measure, then $\mu \in \mathcal{T}_2^\theta(\mathbb{E})$ if and only if $\lambda(\{0\}) = 0$. This is easily seen as follows. Let ϕ be the Fourier transform $\phi(u) = \sum_{k\in\mathbb{Z}^d} a(0,k)\exp(i(u\cdot k))$, $u \in \mathbb{R}^d$, so that

$$\sum_{k\in\mathbb{Z}^d} a_t(0,k)e^{i(u\cdot k)} = \exp(-t(1-\phi(u))).$$

Then we have

$$\begin{aligned}\langle \mu, (\sum_{k\in\mathbb{Z}^d} a_t(0,k)x_k - \theta)^2\rangle &= \langle \lambda, \sum_{k,l\in\mathbb{Z}^d} a_t(0,k)a_t(0,l)\exp(i\,u\cdot(k-l))\rangle\\ &= \langle \lambda, \exp(-2t\Re(1-\phi(u)))\rangle\\ &\to \lambda(\{0\})\end{aligned}$$

as $t \to \infty$ by the dominated convergence theorem, since the irreducibility of $a(i,j)$ implies that $\phi(u) \neq 1$ on $\Pi \setminus \{0\}$.

Now consider the trigonometric polynomials $p_n(u) = |\Lambda_n|^{-1}\sum_{j\in\Lambda_n}\exp(i(u\cdot j))$. These polynomials satisfy

(i) $\lim_{n\to\infty} p_n(u) = 1_{\{0\}}(u)$, and
(ii) for $\delta > 0$ there exists $\epsilon(m,\delta)$ such that if $J_\delta = (-\delta,\delta)^d \setminus \{0\}$, then

$$|p_m(u) - 1_{\{0\}}(u)| \le 1_{J_\delta}(u) + \epsilon(m,\delta),$$

and $\epsilon(m,\delta) \to 0$ as $m \to \infty$.

If $\mu \in \mathcal{T}_2(\mathbb{E})$ is ergodic, with spectral measure λ and $\langle \mu, x_0 \rangle = \theta$, then using (i) it follows from the bounded convergence theorem that

$$\langle \mu, (D_n(x) - \theta)^2 \rangle = \langle \lambda, |p_n(u)|^2 \rangle \to \lambda(\{0\})$$

as $n \to \infty$. Since this limit must be zero, $\mu \in \mathcal{T}_2^\theta(\mathbb{E})$.

The next result gives us a condition under which the weak convergence $\mu_n \Rightarrow \mu$ implies the spatial densities $D(x)$ under μ_n converge in distribution to the spatial density under μ.

Lemma 2.5. *Let $\mu, \mu_1, \mu_2, \ldots \in \mathcal{T}_2(\mathbb{E})$, with respective spectral measures $\lambda, \lambda_1, \lambda_2, \ldots$. If $\mu_n \Rightarrow \mu$, $\langle \mu_n, x_i x_j \rangle \to \langle \mu, x_i x_j \rangle$ for all i, j, and $\lambda_n(\{0\}) \to \lambda(\{0\})$, then*

$$\lim_{k \to \infty} \sup_{m,n \geq k} E^{\mu_n} |D_n(x) - D_m(x)| = 0.$$

Furthermore, if $H_n(\cdot) = \mu_n(D(x) \in \cdot)$ and $H(\cdot) = \mu(D(x) \in \cdot)$, then $H_n \Rightarrow H$.

Remark. The condition $\lambda_n(\{0\}) \to \lambda(\{0\})$ is certainly implied by the condition $\lambda(\{0\}) = 0$.

Proof. We first note that $\lambda_n \Rightarrow \lambda$. So for $\epsilon > 0$ we may choose $\delta > 0$ such that $\lambda(J_\delta) < \epsilon$ and $\lambda_n((-\delta, \delta)^d) \to \lambda((-\delta, \delta)^d)$. This and the assumptions imply $\lambda_n(J_\delta) \to \lambda(J_\delta) < \epsilon$. Thus, letting $C = \sup_n \lambda_n(\Pi)$,

$$\begin{aligned}
\sup_{m,n \geq k} \|p_n(u) - p_m(u)\|_{L_2(\lambda_n)} &\leq \sup_{m,n \geq k} (\|p_n(u) - 1_{\{0\}}(u)\|_{L_2(\lambda_n)} \\
&\quad + \|p_m(u) - 1_{\{0\}}(u)\|_{L_2(\lambda_n)}) \\
&\leq \sup_{m,n \geq k} (2\lambda_n(J_\delta)^{1/2} + C\epsilon(n, \delta) + C\epsilon(m, \delta)) \\
&\to 2\lambda(J_\delta)^{1/2} < 2\epsilon^{1/2}
\end{aligned}$$

as $k \to \infty$. This proves the first assertion, which we may write in the form: for any $\epsilon > 0$ there exists $k < \infty$ such that

$$\|D_n(x) - D_m(x)\|_{L_2(\mu_n)} < \epsilon \quad \text{for } m, n \geq k.$$

Letting $m \to \infty$ above gives $\|D_n(x) - D(x)\|_{L_2(\mu_n)} \leq \epsilon$ for all $n \geq k$, while setting $m = k$ gives $\|D_n(x) - D_k(x)\|_{L_2(\mu_n)} \leq \epsilon$ for all $n \geq k$. Putting these facts together gives

$$\|D_k(x) - D(x)\|_{L_2(\mu_n)} \leq 2\epsilon \quad \text{for } n \geq k.$$

By the L_2–ergodic theorem, by taking k larger, we may also assume that

$$\|D(x) - D_m(x)\|_{L_2(\mu)} < \epsilon \quad \text{for } m \geq k.$$

Since $\mu_n \Rightarrow \mu$ as $n \to \infty$, for any fixed m, $\mathcal{L}_{\mu_n}(D_m(x)) \Rightarrow \mathcal{L}_\mu(D_m(x))$ as $n \to \infty$. Combining these statements we see that $\mathcal{L}_{\mu_n}(D(x)) \Rightarrow \mathcal{L}_\mu(D(x))$, or $H_n \Rightarrow H$.

□

Convergence properties of the coupled process

We prove now a convergence result for the bivariate process $(x(t), y(t))$ as $t \to \infty$ which will have several important applications.

Proposition 2.6. *Assume $\widehat{a}(i,j)$ is transient. Let $\{\bar{\mu}_n\} \subset \mathcal{T}_2(\mathbb{E} \times \mathbb{E})$ satisfy $\sup_n \langle \bar{\mu}_n, x_0^2 + y_0^2 \rangle < \infty$, and every weak limit point $\bar{\mu}$ of $\{\bar{\mu}_n\}$ satisfies $\bar{\mu}\{D(x) = D(y)\} = 1$. Let $(x(t), y(t))$ be the bivariate process defined in (2.7). Then for any $t_n \to \infty$, $E^{\bar{\mu}_n}|x_i(t_n) - y_i(t_n)| \to 0$ as $n \to \infty$.*

Proof. The proof is by contradiction. If the conclusion of the proposition is false, there is a $\delta > 0$ and a subsequence $\{n'\}$ such that

$$E^{\bar{\mu}_{n'}}|x_i(t_{n'}) - y_i(t_{n'})| \to \delta \quad \text{as } n' \to \infty.$$

The uniform second moment condition implies tightness of $\mathcal{L}(x(t_n), y(t_n))$, so we may assume as well that $\bar{\mu}_{n'}$ converges weakly to some $\bar{\mu}$. Since $\bar{\mu} \in \mathcal{T}_2(\mathbb{E} \times \mathbb{E})$, and $\bar{\mu}\{D(x) = D(y)\} = 1$, there is a probability measure λ on I such that $\bar{\mu} = \int_I \bar{\mu}_\rho \, d\lambda(\rho)$, where each $\bar{\mu}_\rho \in \mathcal{T}_2(\mathbb{E} \times \mathbb{E})$ and $\bar{\mu}_\rho\{D(x) = D(y) = \rho\} = 1$ (except for ρ in a λ–null set). This is possible since $\mathbb{E} \times \mathbb{E}$ is Polish, and we can let $\bar{\mu}_\rho$ be a version of $\bar{\mu}(\cdot \mid D(x) = D(y) = \rho)$. By Proposition 2.3(b), for such ρ, $E^{\bar{\mu}_\rho}|x_i(t) - y_i(t)| \to 0$ as $t \to \infty$. Consequently, $\lim_{t\to\infty} E^{\bar{\mu}}|x_i(t) - y_i(t)| = 0$. Proposition 2.3(a) implies that $E^{\bar{\mu}_n}|x_i(t) - y_i(t)|$ is nonincreasing in t for fixed n, while for fixed t, the Feller property, together with the uniform bound on second moments, imply that $E^{\bar{\mu}_{n'}}|x_i(t) - y_i(t)| \to E^{\bar{\mu}}|x_i(t) - y_i(t)|$ as $n' \to \infty$. It therefore follows that for all $t < \infty$,

$$\limsup_{n'\to\infty} E^{\bar{\mu}_{n'}}|x_i(t_{n'}) - y_i(t_{n'})| \le \limsup_{n'\to\infty} E^{\bar{\mu}_{n'}}|x_i(t) - y_i(t)| = E^{\bar{\mu}}|x_i(t) - y_i(t)|.$$

Since the right-hand side above tends to 0 as $t \to \infty$, this is a contradiction, and the proposition is proved. □

Corollary 2.7. *Assume $\widehat{a}(i,j)$ is transient, and μ, $\{\mu_n\}$ satisfy the assumptions of Lemma 2.5. If $t_n \to \infty$ and $\mu S(t_n) \Rightarrow \nu$, then $\mu_n S(t_n) \Rightarrow \nu$.*

Proof. We will define measures $\bar{\mu}_n$ which satisfy the assumptions of Proposition 2.6, such that $\bar{\mu}_n$ has first marginal μ_n and second marginal μ. Given this, Proposition 2.6 implies $E^{\bar{\mu}_n}|x_i(t_n) - y_i(t_n)| \to 0$. Since $\mu S(t_n) \Rightarrow \nu$, this implies $\mu_n S(t_n) \Rightarrow \nu$.

To define $\bar{\mu}_n$ we introduce

$$\begin{aligned}
\mu_n^\theta(\cdot) &= \mu_n(\cdot \mid D(x) = \theta), \\
\mu^\theta(\cdot) &= \mu(\cdot \mid D(x) = \theta), \\
H_n(\cdot) &= \mu_n(D(x) \in \cdot), \\
H(\cdot) &= \mu(D(x) \in \cdot).
\end{aligned}$$

By Lemma 2.5, $H_n \Rightarrow H$ as $n \to \infty$. By using Skorohod's theorem we can construct measures $\bar{H}_n$ and $\bar{H}$ on $\mathbb{R}^2$ such that $\bar{H}_n$ has first marginal H_n, second marginal H, and $\bar{H}_n \Rightarrow \bar{H}$ as $n \to \infty$, where $\bar{H}$ is concentrated on the diagonal, $\bar{H}(\{(a,b) : a = b\}) = 1$. Finally, define

$$\bar{\mu}_n = \int_{\mathbb{E}^2} (\mu_n^\theta \times \mu^{\theta'}) \bar{H}_n(d\theta, d\theta').$$

The $\bar{\mu}_n$ have the correct marginals, so it remains to prove that if $\bar{\mu}$ is a weak limit point of $\{\bar{\mu}_n\}$, then $\bar{\mu}(D(x) = D(y)) = 1$. First, by Lemma 2.5, one can easily see that

$$\lim_{m\to\infty} \lim_{n\to\infty} E^{\mu_n} |D_m(x) - D(x)| = 0.$$

Therefore

$$\begin{aligned} E^{\bar{\mu}}|D(x) - D(y)| &= \lim_{m\to\infty} \lim_{n\to\infty} E^{\bar{\mu}_n} |D_m(x) - D_m(y)| \\ &\le \lim_{m\to\infty} \lim_{n\to\infty} \left[E^{\bar{\mu}_n} |D_m(x) - D(x)| + E^{\bar{\mu}_n} |D_m(y) - D(y)|\right] \\ &\quad + E^{\bar{\mu}_n} |D(x) - D(y)| \\ &\le \lim_{m\to\infty} \lim_{n\to\infty} \left[E^{\mu_n} |D_m(x) - D(x)| + E^{\mu} |D_m(y) - D(y)|\right] \\ &\quad + \lim_{n\to\infty} \iint |\theta - \theta'| \bar{H}_n(d\theta, d\theta') \end{aligned}$$

which is zero, showing $\bar{\mu}(D(x) = D(y)) = 1$. □

Corollary 2.8. *Assume $\widehat{a}(i,j)$ is transient. Let $\{\bar{\mu}^N\} \in \mathcal{T}_2(\mathbb{E}_N \times \mathbb{E}_N)$, $\sup_N \langle \bar{\mu}^N, x_0^2 + y_0^2 \rangle < \infty$, and suppose that every weak limit point $\bar{\mu}$ of the $\bar{\mu}^N$ has the property $\bar{\mu}\{D(x) = D(y)\} = 1$. Let $(x^N(t), y^N(t))$ be the bivariate process defined in (2.8) with initial distribution $\bar{\mu}^N$. Then $E|x_i^N(t_N) - y_i^N(t_N)| \to 0$ for any $t_N \to \infty$.*

Proof. Let $t_N \to \infty$, let $(x^N(0), y^N(0))$ have law $\bar{\mu}^N$, and let $(\tilde{x}^N(0), \tilde{y}^N(0))$ be the periodic extension of $(x^N(0), y^N(0))$ to $\mathbb{Z}^d$. We can now construct a coupled process $(\tilde{x}^N(t), \tilde{y}^N(t), x^N(t), y^N(t))$ with initial state $(\tilde{x}^N(0), \tilde{y}^N(0), x^N(0), y^N(0),) \in \mathbb{E} \times \mathbb{E} \times \mathbb{E}_N \times \mathbb{E}_N$ in the obvious way. By Proposition 2.4 there exists $l_N \to \infty$, $l_N < t_N$ such that for fixed i,

$$E|x_i^N(l_N) - \tilde{x}_i^N(l_N)| \to 0 \quad \text{and} \quad E|y_i^N(l_N) - \tilde{y}_i^N(l_N)| \to 0.$$

By the monotonicity remark after (2.8),

$$E|x_i^N(t_N) - y_i^N(t_N)| \le E|x_i^N(l_N) - y_i^N(l_N)|.$$

The family $\tilde{\mu}^N = \mathcal{L}(\tilde{x}^N, \tilde{y}^N)$ satisfies the assumptions of Proposition 2.6, so $E|\tilde{x}_i^N(l_N) - \tilde{y}_i^N(l_N)| \to 0$. By combining these results we obtain the desired conclusion. □

The next lemma is a crucial ingredient in the proofs of Theorems 1 and 2.

Lemma 2.9. *Assume $d \ge 3$. Suppose that $\mathcal{L}(x^N(0)) \in \mathcal{T}_2(\mathbb{E}_N)$, $\sup_N E|x_0^N(0)|^p < \infty$ for some $p > 2$, and for some random variable Z_0, $\mathcal{L}(\Theta^N(0)) \Rightarrow \mathcal{L}(Z_0)$ as $N \to \infty$. Let $t_N \to \infty$, $t_N \le T\beta_N$, and $\mathcal{L}(x^N(t_N)) \Rightarrow \mu$ as $N \to \infty$. In the case $t_N = O(N^2)$ assume also that for some $\theta \in I$, $\mathcal{L}(x^N(0)) \Rightarrow$ some $\nu \in \mathcal{T}_2^\theta(\mathbb{E})$. Let $\mu_N = \pi_N \mathcal{L}(x^N(t_N))$. Then $\mu, \mu_1, \mu_2, \ldots$ satisfy the assumptions of Lemma 2.5*

Proof. Since $\mu_n \Rightarrow \mu$, and $\sup\langle\mu_n, |x_0|^p\rangle < \infty$, it follows that $\langle\mu_n, x_i x_j\rangle \to \langle\mu, x_i x_j\rangle$ for all i,j. Now let $\mu, \mu^1, \mu^2, \ldots$ have spectral measures $\lambda, \lambda_1, \lambda_2, \ldots$. It is enough to prove that

$$E|\Theta^N(t_N)|^2 \to E^\mu|D(x)|^2, \tag{2.11}$$

since this implies $\lambda_N(\{0\}) \to \lambda(\{0\})$, and hence the assumptions of Lemma 2.5 are fulfilled. Since $\mathcal{L}(x^N(t_N)) \in \mathcal{T}_2(\mathbb{E}_N)$, (2.11) is equivalent to

$$Ex_0^N(t_N)\Theta^N(t_N) \to E^\mu x_0 D(x). \tag{2.12}$$

Before proceeding further we note that by Lemma 2.2(d), $E|\Theta^N(0)|^p$ is bounded. Since $\Theta^N(0) \Rightarrow Z_0$, it follows that $E|\Theta^N(0)|^2 \to E|Z_0|^2$. Similarly,

$$Ex_0^N(t_N)x_j^N(t_N) \to E^\mu x_0 x_j \tag{2.13}$$

for all j. Furthermore,

$$\begin{aligned} Ex_0^N(t_N)x_j^N(t_N) &= \sum_{k,l\in\Lambda_N} a_{t_N}^N(0,k)a_{t_N}^N(j,l)Ex_k^N(0)x_l^N(0) \\ &\quad + \int_0^{t_N} \widehat{a}_{2u}^N(0,j)Eg(x_0^N(t_N-u))\,du. \end{aligned} \tag{2.14}$$

The case $t_N/N^2 \to \infty$: By Lemma 2.1(a), uniformly in j as $N \to \infty$,

$$\begin{aligned} \sum_{k,l\in\Lambda_N} a_{t_N}^N(0,k)\; a_{t_N}^N(j,l)Ex_k^N(0)x_l^N(0) &= \sum_{k,l\in\Lambda_N} \frac{1+o(1)}{(2N)^{2d}}Ex_k^N(0)x_l^N(0) \\ &= E|\Theta^N(0)|^2 + o(1) \to EZ_0^2. \end{aligned} \tag{2.15}$$

Let s_N satisfy $s_N \le t_N$, $s_N/N^2 \to \infty$ and $s_N/N^d \to 0$. By (1.2) and Lemma 2.2(c), $C = \sup\limits_{s\le T\beta_N} Eg(x_0^N(s))$ is finite. So again by Lemma 2.1

$$\int_0^{s_N} \widehat{a}_{2u}^N(0,j)Eg(x_0^N(t_N-u))\,du \le C\int_0^{s_N} \widehat{a}_{2u}^N(0,j)\,du \to C\widehat{A}(0,j)$$

and

$$\int_{s_N}^{t_N} \widehat{a}_{2u}^N(0,j)Eg(x_0^N(t_N-u))\,du = \frac{1+o(1)}{|\Lambda_N|}\int_0^{t_N} Eg(x_0^N(u))\,du \le CT.$$

Now let $\{N_k\}$ be any sequence such that the limit

$$|\Lambda_{N_k}|^{-1}\int_0^{t_{N_k}} Eg(x_0^{N_k}(u))\,du \to \tilde{G} \tag{2.16}$$

exists. In view of (2.13) – (2.15), this implies that for every j, for some $H(0,j) \le C\widehat{A}(0,j)$,

$$\int_0^{s_{N_k}} \widehat{a}_{2u}^{N_k}(0,j)Eg(x_0^{N_k}(t_{N_k}-u))\,du \to H(0,j).$$

Using (2.13) again, we have

$$E^\mu x_0 x_j = \lim_{k\to\infty} Ex_0^{N_k}(t_{N_k})x_j^{N_k}(t_{N_k}) = EZ_0^2 + H(0,j) + \tilde{G}.$$

By (2.14) and (2.15),

$$\begin{aligned} Ex_0^{N_k}(t_{N_k})\Theta^{N_k}(t_{N_k}) &= |\Lambda_{N_k}|^{-1}\sum_{j\in\Lambda_{N_k}}(EZ_0^2+o(1)) + |\Lambda_{N_k}|^{-1}\int_0^{t_{N_k}} Eg(x_0^{N_k}(u))\,du \\ &\to EZ_0^2 + \tilde{G}. \end{aligned}$$

On the other hand,

$$E^\mu x_0 D(x) = \lim_{n\to\infty} E^\mu x_0 D_n(x) = EZ_0^2 + \tilde{G} + \lim_{n\to\infty}|\Lambda_n|^{-1}\sum_{j\in\Lambda_n} H(0,j) = EZ_0^2 + \tilde{G},$$

since $H(0,j) \le C\widehat{A}(0,j) \to 0$ as $j \to \infty$. This is enough to prove (2.12).

The case $t_N = O(N^2)$: We use here the additional assumption that for some $\theta \in I$ and $\nu \in \mathcal{T}_2^\theta(\mathbb{E})$, $\mathcal{L}(x^N(0)) \Rightarrow \nu$. Let γ_N be the spectral measure of $\pi_N\mathcal{L}(x^N(0))$, and let γ be the spectral measure of ν, so that $\gamma_N \Rightarrow \gamma$ as $N \to \infty$, and $\gamma(\{0\}) = 0$. Let $\phi(u) = \sum_{k\in\mathbb{Z}^d} a(0,k)\exp(i(u\cdot k))$ and $\phi^N(u) = \sum_{k\in\Lambda_N} a^N(0,k)\exp(i(u\cdot k))$, $u \in \mathbb{R}^d$. It is easy to see that if $\kappa_N = \sum_{k\notin\Lambda_N} a(0,k)$, then $|\phi^N(u) - \phi(u)| \le \kappa_N$ for all u. If $\phi_t(u) = \sum_{k\in\mathbb{Z}^d} a_t(0,k)\exp(i(u\cdot k))$ and $\phi_t^N(u) = \sum_{k\in\Lambda_N} a_t^N(0,k)\exp(i(u\cdot k))$, then

$$\phi_t(u) = \exp(-t(1-\phi(u))), \quad \phi_t^N(u) = \exp(-t(1-\phi^N(u))).$$

To prove (2.12) we must compute the two terms in the right-hand side of (2.14). Let $\theta_N = Ex_0^N(0)$, $\theta_N \to \theta$ as $N \to \infty$. Then

$$\begin{aligned} \sum_{k,l\in\Lambda_N} a_{t_N}^N(0,k)a_{t_N}^N(j,l) \quad & E(x_k^N(0)-\theta_N)(x_l^N(0)-\theta_N) \\ &= \sum_{k,l\in\Lambda_N} a_{t_N}^N(0,k)a_{t_N}^N(j,l)\langle\gamma_N, \exp(i\,u\cdot(k-l))\rangle \\ &= \langle\gamma_N, |\phi_{t_N}^N(u)|^2 e^{-i(u\cdot j)}\rangle \\ &= \left\{\int_{(-\delta,\delta)^d} + \int_{\Pi\setminus(-\delta,\delta)^d}\right\}|\phi_{t_N}^N(u)|^2 e^{-i(u\cdot j)}\gamma_N(du). \end{aligned}$$

Since $\gamma_N \Rightarrow \gamma$, and $\gamma(\{0\}) = 0$, we have that $\gamma_N(\{0\}) \to 0$, and for $\epsilon > 0$ we can choose $\delta > 0$ such that $\limsup \gamma_N((-\delta,\delta)^d) < \epsilon$. Thus

$$\limsup_{N\to\infty}\int_{(-\delta,\delta)^d} |\phi_{t_N}^N(u)|^2\gamma_N(du) < \epsilon.$$

On the other hand,

$$\begin{aligned}\int_{\Pi\setminus(-\delta,\delta)^d} |\phi^N_{t_N}(u)|^2\gamma_N(du) &= \int_{\Pi\setminus(-\delta,\delta)^d} \exp(-2t_N\Re(1-\phi^N(u)))\gamma_N(du)\\ &\le \int_{\Pi\setminus(-\delta,\delta)^d} \exp(-2t_N\Re(1-\phi(u)-\kappa_N))\gamma_N(du)\\ &\le \sup_{u\in\Pi\setminus(-\delta,\delta)^d} \exp(-2t_N\Re(1-\phi(u)-\kappa_N))\gamma_N(\Pi)\end{aligned}$$

which tends to 0 as $N\to\infty$, since $\kappa_N\to 0$, and the irreducibility of $a(i,j)$ implies $\phi(u)$ is bounded away from 1 on $\Pi\setminus(-\delta,\delta)$.

This shows that the first term in (2.14) converges to θ^2, and hence by (2.13), the second term must also be convergent, say $\int_0^{t_N}\widehat{a}^N_{2s}(0,j)Eg(x^N_0(t_N-s))\,ds\to H(0,j)$. We have therefore established

$$E^\mu x_0x_j = \lim_{N\to\infty} Ex^N_0(t_N)x^N_j(t_N) = EZ_0^2 + H(0,j).$$

As in the first part of the proof, it is now easy to see that (2.12) must hold. □

Regularity properties of ν_θ and $g^*(\theta)$

We will use without proof the following elementary lemma.

Lemma 2.10. *Let X be a Polish space, and suppose that $\{\mu_n\}$ is a weakly convergent sequence of probability measures on X, $\mu_n\Rightarrow\mu$ as $n\to\infty$.*

(a) *If there is a continuous function $\phi\ge 0$ on X such that $\langle\mu_n,\phi\rangle\to\langle\mu,\phi\rangle$, then $\langle\mu_n,\tilde\phi\rangle\to\langle\mu,\tilde\phi\rangle$ for every continous function $\tilde\phi$ on X such that $|\tilde\phi|\le\phi$.*

(b) *If there is a continuous function $\phi\ge 0$ on X such that $\langle\mu_n,\phi\rangle$ is bounded in n, then $\langle\mu_n,\tilde\phi\rangle\to\langle\mu,\tilde\phi\rangle$ for every continous function $\tilde\phi$ on X such that $\tilde\phi(x)/\phi(x)\to 0$ as $x\to\infty$.*

By combining coupling and moment results we can obtain information about the mapping $\theta\to\nu_\theta$.

Lemma 2.11. *Assume $\widehat{a}(i,j)$ is transient.*

(a) $\langle\nu_\theta,x_0\rangle=\theta$, $\langle\nu_\theta,x_ix_j\rangle=\theta^2+\widehat{A}(i,j)\langle\nu_\theta,g(x_0)\rangle$.

(b) *If $g(\theta_0)=0$, and $\theta\ge\theta_0$, then $\nu_\theta(x_i\ge\theta_0,\ i\in\mathbb{Z}^d)=1$. If $g(\theta_0)=g(\theta_1)=0$ for some $\theta_0<\theta_1$, then $\nu_\theta(\theta_0\le x_i\le\theta_1,\ i\in\mathbb{Z}^d)=1$ for $\theta\in[\theta_0,\theta_1]$.*

(c) *For $\theta,\theta'\in I$, $\theta<\theta'$, $\mathbb{E}$–valued random variables x,y can be defined on a common probability space such that $\mathcal{L}(x)=\nu_\theta$, $\mathcal{L}(y)=\nu_{\theta'}$, and $x_i\le y_i$ for all $i\in\mathbb{Z}^d$ with probability one. Consequently, $E|y_i-x_i|=\theta'-\theta$.*

Remark. A consequence of (b) is that $\nu_{\theta_0}=\delta_{\theta_0}$ if $g(\theta_0)=0$. Since the Lipschitz functions form a determining class for weak convergence, a consequence of (c) is that $\nu_{\theta'}\Rightarrow\nu_\theta$ whenever $\theta'\to\theta$

Proof. By (2.5), with $\mathcal{L}(x(0))=\nu_\theta$,

$$Ex_i(t)x_j(t)=E[\sum_{k,l\in\mathbb{Z}^d}a_t(i,k)a_t(j,l)x_k(0)x_l(0)]+\int_0^t\widehat{a}_{2(t-s)}(i,j)Eg(x_0(s))\,ds.$$

Since ν_θ is invariant for $x(t)$,

$$Ex_i(0)x_j(0) = \langle \nu_\theta, \sum_{k\in\mathbb{Z}^d} a_t(i,k)x_k \sum_{l\in\mathbb{Z}^d} a_t(j,l)x_l\rangle + Eg(x_0(0)) \int_0^t \widehat{a}_{2s}(i,j)\, ds$$

for every t. By Lemma 2 of [CG3] and Lemma 2.2 of [S4], $\nu_\theta \in \mathcal{T}^2_\theta(\mathbb{E})$. Thus, as $t \to \infty$, the right-hand side above tends to $\theta^2 + \widehat{A}(i,j)\langle \nu_\theta, g(x_0)\rangle$, proving (a).

For (b), let $\theta \geq \theta_0$ and let $\mathcal{L}(x(0)) = \delta_\theta$. It can be seen from (1.1) that $P(x_i(t) \geq \theta_0,\ t \geq 0, i \in \mathbb{Z}^d) = 1$. But $\mathcal{L}(x(t)) \Rightarrow \nu_\theta$, so $\nu_\theta(x_i \geq \theta_0,\ i \in \mathbb{Z}^d) = 1$. The second part of (b) now follows.

For (c), use (2.7) to construct the bivariate process $(x(t), y(t))$ with $x_i(0) \equiv \theta$ and $y_i(0) \equiv \theta'$. Then Proposition 2.3(c) implies that $\mathcal{L}(x(t)) \Rightarrow \nu_\theta$ and $\mathcal{L}(y(t)) \Rightarrow \nu_{\theta'}$ as $t \to \infty$. Furthermore, a standard approximation argument and Itô's lemma with the function $h(\rho) = \max\{0, \rho\}$ in (2.7) leads to

$$Eh(x_i(t) - y_i(t)) \leq \sum_{j\in\mathbb{Z}^d} a_t(i,j)h(x_j(0) - y_j(0)),$$

which is zero here, consequently $P(x_i(t) \leq y_i(t),\ i \in \mathbb{Z}^d) = 1$ for all $t \geq 0$. The family $\mathcal{L}(x(t), y(t))$ is tight, and letting $\bar{\mu}$ be any weak limit point as $t \to \infty$, it follows that $\bar{\mu}$ has first marginal ν_θ, second marginal $\nu_{\theta'}$, and $\bar{\mu}(x_i \leq y_i,\ i \in \mathbb{Z}^d) = 1$. The proof is completed by choosing a realization (x, y) of $\bar{\mu}$. □

Lemma 2.12. *Assume $\widehat{a}(i,j)$ is transient.*

(a) *g^* is continuous on I.*

(b) *For $\theta \in I$, $g^*(\theta) = 0$ if and only if $g(\theta) = 0$.*

(c) *If $\theta_0 < \theta_1$, $\theta_0, \theta_1 \in I$, and $g(\theta_0) = g(\theta_1) = 0$, then g^* is Lipschitz on $[\theta_0, \theta_1]$.*

(d) *If $b = \limsup_{|\theta|\to\infty} g(\theta)/\theta^2$, then $\limsup_{|\theta|\to\infty} g^*(\theta)/\theta^2 \leq b/(1 - b\widehat{A}(0,0))$.*

(e) *If $g(\theta_0) = 0$, then there is a finite constant C such that $g^*(\theta) \leq C|\theta - \theta_0|$ for θ sufficiently close to θ_0.*

Proof. Recall $g^*(\theta) = \langle \nu_\theta, g(x_0)\rangle = E^{\nu_\theta} g(x_0)$. For (a), in view of Lemma 2.11(a), it suffices to show that $\langle \nu_\theta, x_0^2\rangle$ is continuous in θ. By (1.2) there are constants $b < 1/\widehat{A}(0,0)$ and c finite such that $g(x_0) \leq c + bx_0^2$. Some rearrangement of this inequality yields

$$x_0^2 \leq \frac{c\widehat{A}(0,0) + x_0^2 - \widehat{A}(0,0)g(x_0)}{1 - b\widehat{A}(0,0)}.$$

Letting $\phi(x)$ denote the right-hand side above, it follows from Lemma 2.11(a) that $\langle \nu_\theta, \phi(x)\rangle = (c\widehat{A}(0,0) + \theta^2)/(1 - b\widehat{A}(0,0))$, i.e. $\langle \nu_\theta, \phi(x)\rangle$ is a continuous function of θ. Now Lemma 2.10(a) implies $\langle \nu_\theta, x_0^2\rangle$ is continuous in θ.

For (b), we use Lemma 2.11(a) again to write $\langle \nu_\theta, (x_i - \theta)^2\rangle = \widehat{A}(0,0)g^*(\theta)$. If $g^*(\theta_0) = 0$, then clearly ν_{θ_0} is concentrated on $\{x_i \equiv \theta_0\}$. On the other hand, if $g(\theta_0) = 0$, $\nu_{\theta_0} = \delta_{\theta_0}$, and hence $g^*(\theta_0) = 0$.

For (c) it suffices by Lemma 2.11(a) to show that $\langle \nu_\theta, x_0^2\rangle$ is Lipschitz in $\theta \in [\theta_0, \theta_1]$. To see this we compute

$$\begin{aligned} |\langle \nu_\theta, x_0^2\rangle - \langle \nu_{\theta'}, x_0^2\rangle| &= E|x_0^2(t,\theta) - x_0^2(t,\theta')| \\ &\le 2\max\{|\theta_0|, |\theta_1|\}E|x_0(t,\theta) - x_0(t,\theta')| \\ &\le 2\max\{|\theta_0|, |\theta_1|\}|\theta - \theta'|. \end{aligned}$$

For (d), let $b < b_0 < 1/\widehat{A}(0,0)$, and choose c large enough so that $g(x_0) \le c + b_0 x_0^2$. Using this in Lemma 2.11(a) it is easy to derive

$$\langle \nu_\theta, x_0^2\rangle \le \frac{c\widehat{A}(0,0) + \theta^2}{1 - b_0\widehat{A}(0,0)}.$$

Using Lemma 2.10(a) again we obtain

$$g^*(\theta) \le \frac{c + b_0\theta^2}{1 - b_0\widehat{A}(0,0)}.$$

Thus $\limsup_{|\theta|\to\infty} g^*(\theta)/\theta^2 \le b_0/(1 - b_0\widehat{A}(0,0))$. Let $b_0 \to b$ to complete the proof.

For (e), suppose $g(\theta_0) = 0$, and hence $g^*(\theta_0) = 0$. It suffices to prove that there are constants C_1, C_2 such that

$$g^*(\theta) \le C_1|\theta - \theta_0|^2 + C_2|\theta - \theta_0|. \tag{2.17}$$

First, by (1.2) there constants $b < 1/\widehat{A}(0,0)$ and c finite such that

$$g(x_0) \le b|x_0 - \theta_0|^2 + c|x_0 - \theta_0|.$$

Next, for $\theta \in I$, a little rearrangement gives

$$g(x_0) \le b[(x_0 - \theta)^2 + 2(x_0 - \theta)(\theta - \theta_0) + (\theta - \theta_0)^2] + c|x_0 - \theta_0|.$$

By integrating with respect to ν_θ we obtain

$$\begin{aligned} g^*(\theta) &\le bE^{\nu_\theta}|x_0 - \theta|^2 + b|\theta - \theta_0|^2 + cE^{\nu_\theta}|x_0 - \theta_0| \\ &= b(E^{\nu_\theta}x_0^2 - \theta^2) + b|\theta - \theta_0|^2 + cE^{\nu_\theta}|x_0 - \theta_0| \\ &= b\widehat{A}(0,0)g^*(\theta) + b|\theta - \theta_0|^2 + cE^{\nu_\theta}|x_0 - \theta_0|. \end{aligned}$$

By Lemma 2.11(c), $E^{\nu_\theta}|x_0 - \theta_0| \le |\theta - \theta_0|$, and thus we have

$$g^*(\theta) \le \frac{b|\theta - \theta_0|^2 + c|\theta - \theta_0|}{1 - b\widehat{A}(0,0)},$$

proving (2.17). □

For the final result of this section we define $C_0^1(\mathbb{E})$ to be the collection of all bounded continuous functions ϕ on $\mathbb{E}$ which have the property that for some $C_\phi < \infty$ and finite set $A_\phi \subset \mathbb{Z}^d$,

$$|\phi(x) - \phi(y)| \leq C_\phi \sum_{i \in A_\phi} |x_i - y_i|, \quad x, y \in \mathbb{E}.$$

We write $\langle \mu^N, \phi \rangle$ for $\langle \pi_N \mu^N, \phi \rangle$ for probability measures μ^N on $\mathbb{E}_N$. For $\phi \in C_0^1(\mathbb{E})$ let

$$D_n(\phi, x) = |\Lambda_n|^{-1} \sum_{i \in \Lambda_n} \phi(\sigma_i x)$$

and let $D(\phi, x)$ be the L_2 limit, $D(\phi, x) = \lim_{n \to \infty} D_n(\phi, x)$ whenever it exists. Since each ν_θ is mixing, the L_2-ergodic theorem implies that for $\phi \in C_0^1(\mathbb{E})$, $\lim_{n \to \infty} E^{\nu_\theta} |D_n(\phi, x) - \langle \nu_\theta, \phi \rangle| = 0$. By Proposition 2.3(d), $\lim_{t \to \infty} E^{\delta_\theta} D_n(\phi, x(t)) = E^{\nu_\theta} D_n(\phi, x)$. The next result provides some uniformity we need in interchanging these two limits.

Lemma 2.13. *Fix $M < \infty$ and $\phi \in C_0^1(\mathbb{E})$. If $t_n \to \infty$ as $n \to \infty$, then*

$$\sup_{|\theta| \leq M} E^{\delta_\theta} |D_n(\phi, x(t_n)) - \langle \nu_\theta, \phi \rangle| \to 0.$$

Proof. We prove the lemma by applying the dominated convergence theorem. The first step is to show pointwise convergence by coupling with the stationary process. Fix θ, and let $(x(t), y(t))$ be the bivariate process defined by (2.7) with initial distribution $\delta_\theta \times \nu_\theta$. Since ν_θ is invariant,

$$\begin{aligned} E^{\delta_\theta} |D_n(\phi(x(t_n))) - \langle \nu_\theta, \phi \rangle| \; &\leq \; E |D_n(\phi, x(t_n)) - D_n(\phi, y(t_n))| \\ &\quad + E |D_n(\phi, y(0)) - \langle \nu_\theta, \phi \rangle| \, . \end{aligned}$$

By translation invariance, the right-hand side above is no larger than

$$\begin{aligned} & E |\phi(x(t_n)) - \phi(y(t_n))| + \langle \nu_\theta, |D_n(\phi, y) - \langle \nu_\theta, \phi \rangle| \rangle \\ & \leq C_\phi \sum_{i \in A_\phi} E |x_i(t_n) - y_i(t_n)| + \langle \nu_\theta, |D_n(\phi, y) - \langle \nu_\theta, \phi \rangle| \rangle \end{aligned}$$

which tends to zero as $n \to \infty$ by Proposition 2.3(b) and the L_2 ergodic theorem. This establishes

$$E |D_n(\phi, x(t_n)) - \langle \nu_\theta, \phi \rangle| \to 0 \text{ as } n \to \infty. \tag{2.18}$$

The next step is to show that there is a constant C depending only on ϕ such that if $\theta, \theta' \in I$, and $(x(t), y(t))$ has initial distribution $\delta_\theta \times \delta_{\theta'}$, then for for all $t \geq 0$,

$$E |D_n(\phi, x(t)) - D_n(\phi, y(t))| \leq C |\theta - \theta'|. \tag{2.19}$$

This follows easily from translation invariance and the monotonicity property of Proposition 2.3(a), since

$$E\left|D_n(\phi,x(t))-D_n(\phi,y(t))\right| \le E\left|\phi(x(t))-\phi(y(t))\right| \le C_\phi \sum_{i\in A_\phi} E\left|x_i(t)-y_i(t)\right|$$
$$\le C_\phi \sum_{i\in A_\phi} E\left|x_i(0)-y_i(0)\right| = C\left|\theta-\theta'\right|.$$

Finally, there is a constant $C < \infty$ depending only on ϕ such that if $\theta, \theta' \in I$, then

(2.20)
$$\left|\langle \nu_\theta, \phi\rangle - \langle \nu_{\theta'}, \phi\rangle\right| \le C\left|\theta-\theta'\right|.$$

This is an easy consequence of Lemma 2.10(c). The lemma follows now by combining (2.18), (2.19), and (2.20). □

3 Proof of Theorem 1

We first note that we may assume without loss of generality that $\mathcal{L}(x^N(0)) \in \mathcal{T}_p(\mathbb{E}_N)$. Otherwise, we may replace $\mathcal{L}(x^N(0))$ with its symmetrization, and let $\bar{U}_N(t)$ denote the counterpart of $U_N(t)$. It is routine to see that $U_N(\cdot)$ and $\bar{U}_N(\cdot)$ are equivalent processes.

Our strategy is to establish the following key facts (where $[\,\cdot\,]$ denotes quadratic variation).

(3.1) $\{Z_N(\cdot), N \ge 1\}$ is a tight sequence of continuous L_2 martingales,

(3.2)
$$E \mid \langle U_N(t\beta_N), \phi\rangle - \langle \nu_{Z_N(t)}, \phi\rangle \mid \to 0 \text{ as } N \to \infty \text{ for } t > 0 \text{ and } \phi \in C_0^1(\mathbb{E}),$$

(3.3)
$$E \mid [Z_N](t) - \int_0^t g^*(Z_N(s))\,ds \mid \to 0 \text{ as } N \to \infty \text{ for } t > 0.$$

For given these points, we see from (3.1) and (3.3) that if $\tilde{Z}(\cdot)$ is any limit point of the $Z_N(\cdot)$, then $\tilde{Z}(\cdot)$ is a continuous L_2 martingale with increasing process $[\tilde{Z}](t) = \int_0^t g^*(\tilde{Z}(s))\,ds$. Hence $\tilde{Z}(\cdot)$ must solve

$$\tilde{Z}(t) - \tilde{Z}(0) = \int_0^t \sqrt{g^*(\tilde{Z}(s))}\,dw(s),$$

where $w(s)$ is a Brownian motion. By Lemma 2.12(e), this equation has a unique solution in the probability law sense. Therefore $Z_N \Rightarrow Z$, and thus (3.2) implies $U_N(t\beta_N) \Rightarrow \nu_{Z(t)}$, since $C_0^1(\mathbb{E})$ is a determining class for weak convergence.

Proof of (3.1). It is easy to see from (1.3), since $a(i,j)$ is doubly stochastic, that

$$d\Theta^N(t) = |\Lambda_N|^{-1} \sum_{i\in\Lambda_N} \sqrt{g(x_i^N(t))}\,dw_i(t),$$

and hence that $\Theta^N(t)$ is a martingale, with $[\Theta^N](t) = |\Lambda_N|^{-2} \int_0^t \sum_{i\in\Lambda_N} g(x_i^N(s))\,ds$. By (1.2) and Lemma 2.2(c), for $T < \infty$ there exists $C_T < \infty$ such that for all N and all $s, u \in [0, T\beta_N]$,

$$E|\Theta^N(s) - \Theta^N(u)|^2 \leq C_T |\Lambda_N|^{-1} |s-u|. \tag{3.4}$$

Turning to $Z_N(t) = \Theta^N(t\beta_N)$, it is now clear that Z_N is a martingale with increasing process

$$[Z_N](t) = \int_0^t |\Lambda_N|^{-1} \sum_{i\in\Lambda_N} g(x_i^N(s\beta_N))\,ds,$$

and for $N \geq 1$ and $s, u \in [0, T]$, $E|Z_N(s) - Z_N(u)|^2 \leq C_T|s-u|$. Furthermore, by Lemma 2.2(d) there is a $p' > 2$ such that

$$C_{T,p'} = \sup_{t\in[0,T\beta_N]} E(g(x_0^N(t)))^{p'/2} < \infty.$$

With this p', and $1/q + 2/p' = 1$, proceeding as in the proof of Lemma 2.2(b), for $0 \leq s \leq t \leq T$, we have

$$\begin{aligned} E|Z_N(t) - Z_N(s)|^{p'} &\leq (p'-1)^{p'} E|[Z_N](t) - [Z_N](s)|^{p'/2} \\ &= (p'-1)^{p'} E\left[|\Lambda_N|^{-1} \sum_{i\in\Lambda_N} \int_s^t g(x_i^N(u\beta_N))\,du\right]^{p'/2} \\ &\leq (p'-1)^{p'} |t-s|^{\frac{p'}{2}-1} |\Lambda_N|^{-1} \sum_{i\in\Lambda_N} \int_s^t E(g(x_i^N(u\beta_N)))^{p'/2} du \\ &\leq (p'-1)^{p'} C_{T,p'} |t-s|^{p'/2}. \end{aligned}$$

Since $p'/2 > 1$, this establishes tightness in continuous path space by Kolomogorov's theorem (see [EK], Prop. 3.6.3). □

Proof of (3.2). Fix $t > 0$ and $\phi \in C_0^1(\mathbb{E})$. By Lemma 2.2(c), $C = \sup_{0\leq s\leq\beta_N} E(x_0^N(s))^2 < \infty$. Let ℓ_N be a sequence chosen as in Proposition 2.4(b) for this C, with $\ell_N = o(\beta_N)$ as $N \to \infty$. In view of (2.20), with $t_N = t\beta_N - \ell_N$,

$$E|\langle \nu_{\Theta^N(t\beta_N)}, \phi\rangle - \langle \nu_{\Theta^N(t_N)}, \phi\rangle| \leq C\, E|\Theta^N(t\beta_N) - \Theta^N(t_N)|. \tag{3.5}$$

Since the right-hand side above tends to 0 as $N \to \infty$ by (3.4), in order to prove (3.2) it suffices to prove

$$E|\langle U_N(t\beta_N), \phi\rangle - \langle \nu_{\Theta^N(t_N)}, \phi\rangle| \to 0. \tag{3.6}$$

To do this, let $(\bar{x}^N(t), \bar{y}^N(t))$ be the bivariate process defined by (2.8) with initial state

$$\bar{x}_i^N(0) = x_i^N(t_N), \quad \bar{y}_i^N(0) = |\Lambda_N|^{-1} \sum_{j\in\Lambda_N} x_j^N(t_N), \quad i \in \Lambda_N.$$

Note that $D_N(\bar{x}^N(0)) = D_N(\bar{y}^N(0)) = \Theta^N(t_N)$. For the bivariate process $(\bar{x}^N(t), \bar{y}^N(t))$ define the empirical measures

$$\bar{U}_N^1(t) = |\Lambda_N|^{-1} \sum_{i \in \Lambda_N} \delta_{\sigma_i^N \bar{x}^N(t)}, \quad \bar{U}_N^2(t) = |\Lambda_N|^{-1} \sum_{i \in \Lambda_N} \delta_{\sigma_i^N \bar{y}^N(t)}.$$

By the Markov property,

$$E|\langle U_N(t\beta_N), \phi\rangle - \langle \nu_{\Theta^N(t_N)}, \phi\rangle| = E|\langle \bar{U}_N^1(\ell_N), \phi\rangle - \langle \nu_{D_N(\bar{y}^N(0))}, \phi\rangle|. \tag{3.7}$$

To prove (3.6) it thus suffices to show

$$E|\langle \bar{U}_N^1(\ell_N), \phi\rangle - \langle \bar{U}_N^2(\ell_N), \phi\rangle| \to 0, \tag{3.8}$$

and

$$E|\langle \bar{U}_N^2(\ell_N), \phi\rangle - \langle \nu_{D_N(\bar{y}^N(0))}, \phi\rangle| \to 0. \tag{3.9}$$

We begin with the proof of (3.8). By translation invariance,

$$\begin{aligned} E|\langle \bar{U}_N^1(\ell_N), \phi\rangle &- \langle \bar{U}_N^2(\ell_N), \phi\rangle| \le |\Lambda_N{}^{-1}| \sum_{i \in \Lambda_N} E|\phi(\sigma_i^N \bar{x}^N(\ell_N)) - \phi(\sigma_i^N \bar{y}^N(\ell_N))| \\ &= E|\phi(\bar{x}^N(\ell_N)) - \phi(\bar{y}^N(\ell_N))| \le C_\phi \sum_{i \in A_\phi} E|\bar{x}_i^N(\ell_N) - \bar{y}_i^N(\ell_N)| \\ &= C_\phi |A_\phi| E|\bar{x}_0^N(\ell_N) - \bar{y}_0^N(\ell_N)|. \end{aligned}$$

So it suffices to show

$$E|\bar{x}_0^N(l_N) - \bar{y}_0^N(l_N)| \to 0, \tag{3.10}$$

which we will do by appealing to Lemmas 2.9 and 2.5, and Corollary 2.8.

Let us write $\bar{\mu}^N$ for $\mathcal{L}(\bar{x}^N(0), \bar{y}^N(0))$, and note that $\bar{\mu}^N \in \mathcal{T}_2(\mathbb{E}_N \times \mathbb{E}_N)$, and that $\sup_N \langle \bar{\mu}^N, x_0^2 + y_0^2\rangle < \infty$ by Lemma 2.2(c). If we can show that $\bar{\mu}\{D(x) = D(y)\} = 1$ for every subsequential limit $\bar{\mu}$ of the $\bar{\mu}^N$, then Corollary 2.8 will imply (3.10). For notational simplicity assume $\bar{\mu}^N \Rightarrow \bar{\mu}$, and let $\tilde{\mu}^N \in \mathcal{T}_2(\mathbb{E} \times \mathbb{E})$ be the periodic extension of $\bar{\mu}^N$. Then $\tilde{\mu}^N \Rightarrow \bar{\mu}$ and by Fatou's lemma $\bar{\mu} \in \mathcal{T}_2(E \times E)$. Furthermore, $D_M(x) \to D(x)$ and $D_M(y) \to D(y)$ in $L_2(\bar{\mu})$ as $M \to \infty$. Thus, given $\epsilon > 0$ there exists $M_0 < \infty$ such that for $M \ge M_0$

$$E^{\bar{\mu}}|D(x) - D(y)| \le \epsilon + E^{\bar{\mu}}|D_M(x) - D_M(y)|.$$

By weak convergence and uniform boundedness of second moments,

$$E^{\bar{\mu}}|D_M(x) - D_M(y)| = \lim_{N \to \infty} E^{\tilde{\mu}^N}|D_M(x) - D_M(y)|.$$

We observe now that with respect to the measure $\tilde{\mu}^N$, $D_M(y) = D_N(x)$ by construction. Thus

$$\lim_{N \to \infty} E^{\tilde{\mu}^N}|D_M(x) - D_M(y)| = \lim_{N \to \infty} E^{\tilde{\mu}^N}|D_M(x) - D_N(x)|.$$

Now the law of x under $\bar{\mu}^N$ is $\mathcal{L}(x^N(t_N))$, so by Lemmas 2.9 and 2.5, the the right-hand side above tends to zero, proving $E^{\bar{\mu}}|D(x) - D(y)| = 0$. This justifies the use of Corollary 2.8.

Turning to the proof of (3.9), the basic idea is to extend to the infinite system and apply Lemma 2.13. To do this, let $\tilde{y}_i^N(0) = \bar{y}_0^N(0)$, $i \in \mathbb{Z}^d$, and let $(\bar{y}(t), \bar{y}(t))$ be the bivariate process defined by (2.9). The sequence ℓ_N was previously chosen to ensure

$$(3.11) \qquad |\Lambda_N|^{-1} \sum_{i\in\Lambda_N} E|\bar{y}_i^N(\ell_N) - \tilde{y}_i^N(\ell_N)| \to 0.$$

Since A_ϕ is finite, we can choose $K < \infty$ such that $i+j \in \Lambda_N$ for all $i \in \Lambda_{N-K}$ and $j \in A_\phi$. Thus, for some finite C'

$$E|\langle \bar{U}_N^2(\ell_N), \phi\rangle - D_N(\phi, \bar{y}^N(\ell_N))| \le |\Lambda_N|^{-1} \sum_{i\in\Lambda_N} E|\phi(\sigma_i^N \bar{y}^N(\ell_N)) - \phi(\sigma_i \tilde{y}^N(\ell_N))|$$

$$\le |\Lambda_N|^{-1} \sum_{i\in\Lambda_N} \sum_{j\in A_\phi} C_\phi E|(\sigma_i^N \bar{y}^N(\ell_N))_j - (\sigma_i \tilde{y}^N(\ell_N))_j|$$

$$\le C_\phi |A_\phi| |\Lambda_N|^{-1} \sum_{i\in\Lambda_N} E|\bar{y}_i^N(\ell_N) - \tilde{y}_i^N(\ell_N)| + C_\phi |A_\phi| C'K/N.$$

In view of (3.11), this implies

$$(3.12) \qquad E|\langle \bar{U}_N^2(\ell_N), \phi\rangle - D_N(\phi, \bar{y}^N(\ell_N))| \to 0 \quad \text{as } N \to \infty,$$

so the proof of (3.9) will be completed by proving

$$(3.13) \qquad E|D_N(\phi, \bar{y}^N(\ell_N)) - \langle \nu_{D_N(\bar{y}(0))}, \phi\rangle| \to 0.$$

To do this we first observe that

$$K = \sup_N E|D_N(\phi, \bar{y}^N(\ell_N)) - \langle \nu_{D_N(\bar{y}^N(0))}, \phi\rangle|^2 < \infty$$

by Lemma 2.2(c). Next, by a simple decomposition and the Cauchy-Schwarz inequality,

$$E|D_N(\phi, \bar{y}^N(\ell_N)) - \langle \nu_{D_N(\bar{y}^N(0))}, \phi\rangle| \le$$
$$(KP(D_N(\bar{y}^N(0)) > M))^{1/2} + \sup_{|\theta|\le M} E^{\delta_\theta}|D_N(\phi, \bar{y}(\ell_N)) - \langle \nu_\theta, \phi\rangle|.$$

By Lemma 2.13, the second term in the right-hand side above tends to zero as $N \to \infty$. Since the sequence $D_N(\bar{y}^N(0)) = \Theta^N(t_N)$ is tight, the first term in the right-hand side above can be made arbitrarily small by choosing M sufficiently large, and we are done with the proof of (3.2). □

Proof of (3.3). We begin with

$$E|[Z_N](t) - \int_0^t g^*(Z_N(s))\,ds| = E|\int_0^t (|\Lambda_N|^{-1}\sum_{i\in\Lambda_N} g(x_i^N(s\beta_N))$$

$$-\langle \nu_{Z_N(s)}, g(x_0)\rangle)\,ds| \le \int_0^t E|\langle U_N(s\beta_N), g(x_0)\rangle - \langle \nu_{Z_N(s)}, g(x_0)\rangle|\,ds.$$

We will see that the right-hand side above tends to 0 as $N \to \infty$ by (3.2) and our basic L_p estimates. We first note that if X is any random variable with $E|X|^p = K^p$, it follows from Hölder's inequality and Markov's inequality that if $1/q + 2/p = 1$, then

$$EX^2 1\{|X| \ge M\} \le (E|X|^p)^{2/p}(P|X| \ge M)^{1/q} \le K^2(E|X|^p/M^p)^{1/q},$$

and thus

$$EX^2 1\{|X| \ge M\} \le K^p/M^{p-2}. \tag{3.14}$$

By (1.2), given $\epsilon > 0$ we can write $g = g_1 + g_2$ where g_1 is a bounded, Lipschitz function on I and $0 \le g_2(\theta) \le \widehat{A}(0,0)^{-1}\theta^2 1_{\{|\theta|\ge\epsilon^{-1}\}}$, $\theta \in I$. By (3.2) and the bounded convergence theorem,

$$\int_0^t E|\langle U_N(s\beta_N), g_1(x_0)\rangle - \langle \nu_{Z_N(s)}, g_1(x_0)\rangle|\,ds \to 0,$$

so it suffices to show that

$$\int_0^t E\langle U_N(s\beta_N), g_2(x_0)\rangle\,ds + \int_0^t E\langle \nu_{Z_N(s)}, g_2(x_0)\rangle\,ds$$

is small, uniformly in N, if ϵ is small. The first term above is straightforward, since by Lemma 2.2(d) and (3.14), there is a finite constant K such that for all N,

$$\begin{aligned}\int_0^t E\langle U_N(s\beta_N), g_2(x_0)\rangle\,ds &\le \widehat{A}(0,0)^{-1}\int_0^t E|x_0^N(s\beta_N)|^2 1_{\{x_0^N(s\beta_N)\ge\epsilon^{-1}\}}\,ds\\ &\le \widehat{A}(0,0)^{-1}tK^p\epsilon^{p-2}.\end{aligned}$$

Next, as observed in the proof the proof of Lemma 2.12(d), there are constants c_0, c_1 such that $\langle \nu_\theta, x_0^2\rangle \le c_0 + c_1\theta^2$ for all $\theta \in I$. So again, there is a finite K such that for $M \ge 1$ and all N,

$$\begin{aligned}E\int_0^t \langle \nu_{Z_N(s)}, g_2(x_0)\rangle 1_{\{Z_N(s)>M\}}\,ds &\le \widehat{A}(0,0)^{-1}E\int_0^t (c_0 + c_1Z_N^2(s))1_{\{Z_N(s)>M\}}\,ds\\ &\le \widehat{A}(0,0)^{-1}t(c_0+c_1)K^p/M^{p-2}.\end{aligned}$$

Finally, for fixed M,

$$E\int_0^t \langle \nu_{Z_N(s)}, g_2(x_0)\rangle 1_{\{|Z_N(s)|\le M\}}\,ds \le t\widehat{A}(0,0)^{-1}\sup_{|\theta|\le M}\langle \nu_\theta, x_0^2 1_{\{|x_0|\ge\epsilon^{-1}\}}\rangle \to 0$$

as $\epsilon \to 0$. This is because by Lemma 2.11(c) and dominated convergence,

$$\sup_{\theta\in[-M,M]} \langle \nu_\theta, x_0^2 1_{\{|x_0|\geq\epsilon^{-1}\}}\rangle \leq \langle \nu_M, x_0^2 1_{\{|x_0|\geq\epsilon^{-1}\}}\rangle + \langle \nu_{-M}, x_0^2 1_{\{|x_0|\geq\epsilon^{-1}\}}\rangle \to 0.$$

□

Remark. It is the proof of (3.3) that relies most strongly on the assumption $\sup_N E\langle U_N(0), |x_0|^p\rangle < \infty$. Given (3.3), a proof of the tightness requirement in (3.1) can be made using only second moments, (3.3), and a theorem of Aldous (see [A]). We note that (3.3) can be proved using only second moments provided (1.2) is replaced by $\lim_{|\theta|\to\infty} g(\theta)/\theta^2 = 0$.

4 Proof of Theorem 2

We first note that the assumptions of Theorem 1 hold with $Z(0) = \rho$. Next, by (3.4), since $t_N/\beta_N \to s$,

$$E|\Theta^N(t_N) - Z_N(s)|^2 = E|\Theta^N(t_N) - \Theta^N(s\beta_N)|^2 \leq C|s - t_N/\beta_N| \to 0\,.$$

Since $Z_N(\cdot) \Rightarrow Z(\cdot)$, (1.7) must hold. Next, supposing that $\phi \in C_0^1(\mathbb{E})$ and $s > 0$, the method of proof (3.2) yields

$$E|\langle U_N(t_N), \phi\rangle - \langle \nu_{\Theta^N(t_N)}, \phi\rangle| \to 0\,. \tag{4.1}$$

Therefore

$$E\langle U_N(t_N), \phi\rangle \to \int_I Q_s(\rho, d\theta)\langle \nu_\theta, \phi\rangle\,. \tag{4.2}$$

But since $\mathcal{L}(x^N(0)) \in \mathcal{T}_2(\mathbb{E})$, the left-hand side above equals $E\phi(x^N(t_N))$, and (1.8) is proved.

Theorem 1 does not apply in the case $s = 0$, so we must prove directly that for $\phi \in C_0^1(\mathbb{E})$,

$$E\phi(x^N(t_N)) \to \langle \nu_\rho, \phi\rangle \text{ as } N \to \infty. \tag{4.3}$$

Suppose first that $t_N/N^2 \not\to \infty$. This means that we are assuming that $\mathcal{L}(x^N(0)) \in \mathcal{T}_2(\mathbb{E}_N)$, $\sup_N E|x^N(0)|^p \leq C$ for some finite constant C, $\Theta^N(0) \to_p \rho$, and $\mathcal{L}(x^N(0)) \Rightarrow$ some shift-ergodic element ν of $\mathcal{T}_2(\mathbb{E})$. The assumptions imply that $\langle \nu, x_0\rangle = \rho$, and hence $\nu \in \mathcal{T}_2^\rho$.

Put $C = \sup_N \sup_{t\leq\beta_N} Ex_0^N(t)^2$, and let $\{\ell_N\}$ be as in Proposition 2.4(a) for this C, with $\ell_N \leq t_N/2$, and let $t_N' = t_N - \ell_N$. The moment condition and Lemma 2.2(d) imply the sequence $\mathcal{L}(x^N(t_N'))$ is tight, so consider any subsequence (which for ease of notation we still denote t_N') such that $\mathcal{L}(x^N(t_N')) \Rightarrow \mu$ as $N \to \infty$ for some $\mu \in \mathcal{T}_2(\mathbb{E})$.

By (1.7), $\Theta^N(t_N') \to_p \rho$. Thus by Lemmas 2.9 and 2.5, $\mu(\{D(x) = \rho\}) = 1$, which implies that the associated spectral measure of μ must assign mass 0 to $\{0\}$, and hence $\mu \in \mathcal{T}_2^\rho(\mathbb{E})$. By Proposition 2.3(d),

$$\langle \mu S(\ell_N), \phi \rangle \to \langle \nu_\rho, \phi \rangle. \tag{4.4}$$

If we let $\tilde{\mu}^N = \pi_N \mathcal{L}(x^N(t'_N))$, then Corollary 2.7 and (4.4) imply

$$\langle \tilde{\mu}^N S(\ell_N), \phi \rangle \to \langle \nu_\rho, \phi \rangle. \tag{4.5}$$

To go from (4.5) to (4.3) we define the bivariate process $(z(t), z^N(t))$ on $\mathbb{Z}^d \times \Lambda_N$ constructed as in (2.9) such that

$$\begin{aligned} &z_i^N(0) = x_i^N(t'_N), \quad i \in \Lambda_N \\ &z_i(0) = z_j^N(0), \quad j = i \bmod (2N), \quad i \in \mathbb{Z}^d, j \in \Lambda_N. \end{aligned}$$

Let $\bar{\mu}^N$ denote the law of $(z(0), z^N(0))$. By Proposition 2.4(a), for every i,

$$E^{\bar{\mu}^N} |z_i(\ell_N) - z_i^N(\ell_N)| \to 0,$$

and hence

$$E^{\bar{\mu}^N} |\phi(z(\ell_N)) - \phi(z^N(\ell_N))| \to 0. \tag{4.6}$$

But

$$E\phi(x^N(t_N)) = E^{\bar{\mu}^N} \phi(z^N(\ell_N)),$$

and

$$\langle \mu S(\ell_N), \phi \rangle = E^{\bar{\mu}^N} \phi(z(\ell_N)),$$

so (4.3) follows from (4.5) and (4.6).

To finish we must consider the case $t_N/\beta_N \to 0, t_N/N^2 \to \infty$. Here we do not assume the existence of the weak limit $\mathcal{L}(x^N(0)) \Rightarrow \nu$. But in the just completed argument, the existence of this limit was used only to justify the application of Lemma 2.9. This is not needed if $t_N/N^2 \to \infty$.

Acknowledgement. We would like to thank Tom Mountford for spotting an error in a preliminary version of Lemma 2.5.

T. Shiga would like to dedicate this paper to the memory of his father Umehisa Shiga, who passed away while this work was in progress.

References

[A] D. Aldous, Stoppings times and tightness II, Ann. Probab. **17**, (1989), 586–595

[BCGH] J.-B. Baillon, Ph. Clément, A. Greven, F. den Hollander, On the attracting orbit of a non-linear transformation arising from renormalization of hierarchically interacting diffusions, Can. J. Math. (1994)

[B] D. Burkholder, Sharp inequalities for martingales and stochastic integrals, Astérique **157–158** (1988), 75–94

[C] J.T. Cox Coalescing random walks and voter model consensus times on the torus in $\mathbb{Z}^d$, Ann. Probab. **17** (1989), 1333–1366

[CG1] J.T. Cox and A. Greven, On the long term behavior of some finite particle systems, Probab. Th. Rel. Fields **85** (1990), 195–237

[CG2] J.T. Cox and A. Greven, On the long term behavior of some finite particle systems: a critical dimension example, Random Walks, Brownian Motion, and Interacting Particle Systems, R. Durrett and H. Kesten eds. (1991) Birkhäuser, 203–213

[CG3] J.T. Cox and A. Greven, Ergodic theorems for infinite systems of interacting diffusions, Ann. Probab. **22** (1994), 833–853

[CG4] J.T. Cox and A. Greven, The finite systems scheme: an abstract theorem and a new example, CRM Proceedings & Lecture Notes **5** (1994), 55–67

[CGS] J.T. Cox, A. Greven, T. Shiga, Finite and infinite systems of interacting diffusions: Part II, Preprint

[DG1] D.A. Dawson and A. Greven, Multiple time scale analysis of interacting diffusions, Prob. Th. Rel. Fields **95** (1993), 467–508

[DG2] D.A. Dawson and A. Greven, Models of Hierarchical Interacting Diffusions: Multiple Time Scale Phenomena, Phase Transition and Patterns of Cluster Formation, Prob. Th. Rel. Fields **96** (1993), 435–474

[CGri] J.T. Cox and D. Griffeath, Diffusive clustering in the two dimensional voter model, Ann. Probab. **14** (1986), 347–370

[D1] J.-D. Deuschel, Central limit theorem for an infinite lattice system of interacting diffusion processes, Ann. Probab. **16** (1988), 700–716

[D2] J.-D. Deuschel, Algebraic L_2 decay of critical processes on the lattice, Ann. Probab. **22** (1994), 264–288

[Du] R. Durrett, An infinite particle system with additive interactions, Adv. Appl. Probab. **11** (1979), 353–383

[EK] S.N. Ethier and T.G. Kurtz, Markov Processes – Characterization and Convergence, John Wiley & Sons (1986)

[FG] K. Fleischmann and A. Greven, Diffusive clustering in an infinite system of hierarchically interacting diffusions, Prob. Th. Rel. Fields **98** (1994), 517–566

[KO] M. Kimura and T. Ohta, Theoretical Aspects of Population Genetics, Princeton University Press, 1971

[KS] N. Konno and T. Shiga, Stochastic partial differential equations for some measure-valued diffusions, Prob. Th. Rel. Fields **79** (1988), 201–225

[L] T.M. Liggett, Interacting Particle Systems, Springer Verlag, 1985

[LS] T.M. Liggett and F. Spitzer, Ergodic theorems for coupled random walks and other systems with locally interacting components, Z. Wahrsch. verw. Gebiete **56** (1981), 443–468

[MW] E.W. Montroll and G.H. Weiss, Random walks on Lattices II, J. Math. Phys. **6** (1965), 167–181

[NS] M. Notohara and T. Shiga, Convergence to genetically uniform state in stepping stone models of population genetics, J. Math. Bio. **10** (1980), 281–294

[S1] T. Shiga, An interacting system in population genetics I, J. Math. Kyoto Univ. **20** (1980), 213–243

[S2] T. Shiga, An interacting system in population genetics II, J. Math. Kyoto Univ. **20** (1980), 723–733

[S3] T. Shiga, A certain class of infinite-dimensional diffusion processes arising in population genetics, J. Math. Soc. Japan **39** (1987), 17–25

[S4] T. Shiga, Ergodic theorems and exponential decay of sample paths for certain interacting diffusion systems, Osaka J. Math. **29** (1992), 789–807

[SS] T. Shiga and A. Shimizu, Infinite-dimensional stochastic differential equations and their applications. J. Math. Kyoto Univ. **20** (1980), 395–416

[SU] T. Shiga and K. Uchiyama, Stationary states and their stability of the stepping stone model involving mutation and selection, Prob. Th. Rel. Fields **73** (1986), 87–117

[Sp] F. Spitzer, Principles of Random Walk, Springer-Verlag, 1976

[WS] S. Watanabe and T. Yamada, On the uniqueness of solutions of stochastic differential equations, J. Math. Kyoto Univ. **11** (1971), 155–167