Probab. Theory Relat. Fields 95, 467–508 (1993)

Probability Theory and Related Fields

Multiple time scale analysis of interacting diffusions

Donald Dawson[1] **and Andreas Greven**[2]

[1] Department of Mathematics & Statistics, Carleton University, Ottawa K1S 5B6, Canada
[2] Institut für Mathematische Stochastik, Universität Göttingen, Lotzestrasse 13, W-3400 Göttingen, Germany

Received April 24, 1990; in revised form September 1, 1992

Summary. We consider the questions: how can the long term behavior of large systems of interacting components be described in terms of infinite systems? On what time scale does the infinite system give a qualitatively correct description and what happens at large (resp. critical) time scales?

Let $Y^N(t)$ be a solution $(y_i^N(t))_{i\in[-N,N]}$ of the system of stochastic differential equations ($w_i(t)$ are i.i.d. brownian motions)

$$d y_i^N(t)=\left(\frac{1}{2N+1}\sum_{j=-N}^{N} y_j^N(t)-y_i^N(t)\right)dt+\sqrt{2g(y_i^N(t))}\,dw_i(t).$$

In the McKean-Vlasov limit, $N\to\infty$, we obtain the infinite independent system

$$d y_i^\infty(t)=(E(y_i^\infty(t))-y_i^\infty(t))\,dt+\sqrt{2g(y_i^\infty(t))}\,dw_i(t),\qquad i\in\mathbf{Z}.$$

This infinite system has a one parameter set of invariant measures $\nu_\Theta=\bigotimes_{x\in\mathbf{Z}}\Gamma_\Theta$ with Γ_Θ the equilibrium measure of $dx(t)=(\Theta-x(t))\;dt+\sqrt{2g(x(t))}\,dw(t)$. Let $Q_s(\cdot,\cdot)$ be the transition kernel of the diffusion with generator $u_g(x)\left(\frac{\partial}{\partial x}\right)^2$ with $u_g(x)=\int g(y)\,\Gamma_x(dy)$. Then one main result is that as $N\to\infty$

$$\mathscr{L}((Y^N(s(2N+1))))\Rightarrow\int Q_s(\Theta',d\Theta)\,\nu_\Theta,\qquad \Theta'=E(y_0).$$

This provides a new example of a phenomenon also exhibited by the voter model and branching random walk. In particular we are also able to modify our model by adding the term $cN^{-1}(A-y_i^N(t))\,dt$ to obtain the first example in which the analog of $Q_s(\cdot,\cdot)$ converges to an honest equilibrium instead of absorption in traps as in all models previously studied in the literature. Finally, we discuss a hierarchical model with two levels from the point of view discussed above but now in two fast time scales.

Mathematics Subject Classification: 60K35

0 Introduction

The aim of this paper is to describe the long-term collective behavior of certain very *large systems* of interacting diffusions which are of interest in mathematical biology. The main idea is to describe this behavior in different time scales as the system size tends to infinity. This is in contrast to the usual McKean-Vlasov limit in which the system is studied in the original time scale.

A basic question is to determine the range of times, as a function of the system size, for which the large system has the same behavior as the infinite system and conversely, at what time scale is the finiteness of the system manifested. This raises also the question of the behavior in the critical range between these two regimes. In Cox and Greven [3, 4] and Greven [6], an approach to the analysis of this type of problem was introduced and applied to various interacting particle systems (critical branching, voter model, contact process). In this paper we shall apply this approach to a different class of models, namely *diffusions with mean-field interaction* including Feller's continuous state branching diffusion, the Fisher-Wright model and Kimura's random selection model both motivated by population genetics. With these results it is also possible to derive the asymptotic behavior of the trapping times of the finite model as the system size tends to infinity.

An important feature of this work is the extension to systems in a weak external field. In this case the finite system can approach a nondegenerate equilibrium in contrast to the models previously treated in which the system ultimately ends up in a trap.

An advantage of the mean field case is that some aspects of the proofs are more transparent and the basic heuristic principles explaining the behavior of the system can be turned into rigorous proofs. The mean field interaction allows us to treat a whole class of models and we need not use the rather specific (and restrictive) structural properties of the models such as duality relations which play a central role in lattice models. Furthermore the results in the mean field case are the building block for the analysis of *infinite hierarchical systems.* This is explored in forthcoming papers, see Dawson and Greven [6], [7]. For a survey of the role of mean field models, especially in mathematical physics, see Dawson and Gärtner [5].

The techniques developed allow us to also study *two level hierarchical* situations, which yield additional insight into multiple scale phenomena in large systems and which can be exhibited in large computer simulations of such systems. We expect then to have two time scales: the shorter in which the large system provides stable boundary conditions for smaller subsystems in which case the latter will have their global characteristics relax into an equilibrium state which is related to the infinite system in equilibrium with these global characteristics. On a longer time scale the whole system will feel it is finite and change its global characteristics thus forcing the subsystem to relax to a new equilibrium and change its local properties accordingly. For this reason we expect interesting longterm behavior of the system with various regimes depending on the time scale. In particular we shall study a simple example which displays this type of behavior.

Notation. $\mathscr{L}(X)$ denotes the law of the random variable X. $\mathscr{L}(X_N) \underset{N\to\infty}{\Rightarrow} \mathscr{L}(X)$ denotes weak convergence of X_N to X and $w(t)$ denotes Brownian motion.

1 Model and statement of main results

(a) Mean field models without external field

(i) The model. Consider $Y_t^N=(y_i^N(t))_{i\in[-N,N]}$ which satisfies the system of stochastic differential equations (ingredients are explained below)

$$(1.1)\quad dy_i^N(t)=\left(\frac{1}{2N+1}\sum_{j=-N}^{N} y_j^N(t)-y_i^N(t)\right)dt+\sqrt{2g(y_i^N(t))}\,dw_i(t),\qquad i\in[-N,N],$$
$$\mathscr{L}((Y_0^N))=\mu|_{\mathbf{R}[-N,N]},$$

where μ is an i.i.d. measure on $\mathbf{R}^{\mathbf{Z}}$. (We use the index set $\mathbf{Z}$ to allow for comparison with lattice case considered in [3, 4].)

The corresponding infinite system arises as the McKean Vlasov limit, $N\to\infty$, and is given by $Y_t=(y_i(t))_{i\in\mathbf{Z}}$, where

$$(1.2)\qquad dy_i(t)=(E^\mu(y_i(t))-y_i(t))\,dt+\sqrt{2g(y_i(t))}\,dw_i(t),$$
$$\mathscr{L}(Y_0)=\mu.$$

In (1.1) and (1.2) $(w_i(t))_{i\in\mathbf{Z}}$ are i.i.d. Brownian motions and g satisfies

$$g\geqq 0,\qquad \{x: g(x)>0\}=(a,b),$$
$$g(a)=0\qquad (g(b)=0)\ \text{if } a \text{ (respectively, } b\text{) is finite.}$$

Furthermore

$$(a,b)=(-\infty,\infty),\qquad (0,\infty)\quad\text{or}\quad (0,1).$$

Some typical examples are

$(a,b)=(0,\infty)$,	$g(x)=x$	(Feller's continuous state branching)
$(a,b)=(-\infty,\infty)$,	$g(x)=\sigma>0$	(Critical Ornstein-Uhlenbeck process)
$(a,b)=(0,1)$,	$g(x)=x(1-x)$	(Fisher-Wright diffusion)
$(a,b)=(0,1)$,	$g(x)=x^2(1-x)^2$	(Kimura's random selection model).

In this paper we shall focus on the case $(a,b)=(0,1)$ and assume

$$(1.3)\qquad g(x)>0\quad\text{for } x\in(0,1)\quad\text{and}\quad g(0)=g(1)=0,$$
$$g \text{ is Lipschitz continuous on } [0,1].$$

Under (1.3) the system of stochastic differential equations (1.1) has a *unique strong solution* (cf. Yamada and Watanabe [12, Theorem 1]).

The system (1.1) represents a special subclass of exchangeably interacting diffusions and as $t\to\infty$ the finite system reaches, for every N, one of the two traps $y_i^N\equiv 0$ or $y_i^N\equiv 1$. Hence these are the only extremal equilibria. On the other hand from (1.2) it immediately follows that $E^\mu(y_i(t))=E^\mu(y_i(0))$ for all $t>0$, in other words the mean Θ is a *conserved quantity*. It then follows that for each $0<\Theta<1$, the infinite system has for each Θ a unique equilibrium measure ν_Θ given by

(1.4) $\nu_\Theta=(\Gamma_\Theta)^{\otimes \mathbf{Z}}$, where Γ_Θ is the unique equilibrium of the diffusion

$$dx_t=(\Theta-x_t)\,dt+\sqrt{2g(x_t)}\,dw_t.$$

We define $\nu_0=\delta_{\{Y\equiv 0\}}$ and $\nu_1=\delta_{\{Y\equiv 1\}}$. Recall that the finite systems have as equilibria only the traps, so the *infinite and finite systems have different ergodic behavior.* It is however expected that in the times before the finite system reaches the traps a *quasi-equilibrium* is reached. The latter should resemble the stationary state of the infinite system. The description of this phenomenon for large finite systems is the object of this study.

(ii) The multiple timescale viewpoint. In order to describe the longterm behavior of large finite systems we introduce the following objects (cf. Cox and Greven [3, 4]).

I. $\{\nu_\Theta\}_{\Theta\in I}$ is the set of extremal invariant measures of the infinite system. Θ is called the *conserved quantity* and the range I of possible values of Θ defines the *ergodic components* of the infinite system. In our case $I=[0,1]$,

$$\nu_\Theta=(\Gamma_\Theta)^{\otimes \mathbf{Z}}, \quad \text{where} \quad \Gamma_\Theta(dy)=C(g(y))^{-1}\exp\left(\int_\Theta^y \frac{\Theta-x}{g(x)}\,dx\right)dy$$

and $C=C(\Theta,g)$ is the normalization constant.

II. The *slowly varying variable* $\Theta_s^N=\Theta^N(Y_s^N)$ where Θ^N is a consistent estimator of Θ based on the $2N+1$-particle system, that is, $\Theta^N\to\Theta$ as $N\to\infty$. In the case of the system (1.1):

$$\Theta^N(Y_s^N)=\frac{1}{2N+1}\sum_{i=-N}^{N} y_i^N(s).$$

III. $\beta(N)$ is the *timescale* such that $\mathscr{L}(\Theta_{L(N)}^N-\Theta_0^N)\Rightarrow\delta_0$ as $N\to\infty$ for all $L(N)=o(\beta(N))$, but not for $L(N)=\beta(N)$. In our case $\beta(N)=2N+1$.

IV. The *macroscopic observable* associated with the large system is denoted by $(Z_s)_{s\in\mathbf{R}^+}$, where in our situation of interacting diffusions, Z_s is a diffusion on I generated by the operator $A_g=u_g(x)\left(\frac{\partial}{\partial x}\right)^2$ with $u_g(\Theta)=E^{\nu_\Theta}(g(y_0))$, that is, $u_g(x)=\int g(y)\,\Gamma_x(dy)$. The transition kernel of the resulting diffusion will be denoted by $Q_s(\cdot,\cdot)$.

(iii) Main results. Before we state the results we adopt the convention that $T(N)=s\beta(N)$ where $s=0$, $s=\infty$, simply means that $T(N)=o(\beta(N))$, $\beta(N)=o(T(N))$, respectively.

As initial distribution for the system Y_t^N we choose the restriction of a product measure μ on $[0,1]^{\mathbf{Z}}$ to $[0,1]^{[-N,N]}$ with $E^\mu(y_0)=\int y_0(\omega)\,\mu(d\omega)=\Theta'$. We shall use the following notation for weak convergence of probability measures on

R: $\mathscr{L}(X_N) \underset{N\to\infty}{\Rightarrow} \mathscr{L}(X)$. For weak convergence of probability measures on the pathspace $C([0,\infty), \mathbf{R}^{\mathbf{Z}})$ or $D([0,\infty), \mathbf{R}^{\mathbf{Z}})$ we write:

$$\mathscr{L}((X_s^N)_{s\in\mathbf{R}^+}) \underset{N\to\infty}{\Rightarrow} \mathscr{L}((X_s)_{s\in\mathbf{R}^+}).$$

The following is our main result: (with the objects β, Θ_s^N, Z_s, Q_s, ν_Θ as defined in Subsect. (ii).)

Theorem 1 *Let* $\beta(N)=2N+1$, $\Theta'=E^\mu(y_0(0))$. Then with $Z_0=\Theta'$

(1.5) $$\mathscr{L}((\Theta_{s\beta(N)}^N)_{s\in\mathbf{R}^+}) \underset{N\to\infty}{\Rightarrow} \mathscr{L}((Z_s)_{s\in\mathbf{R}^+}).$$

For $s\geqq 0$ *let* $(Y_t^{\nu(s)})_{t\in\mathbf{R}^+}$ *be the stationary solution for (1.2) with marginal*

$$\nu(s)=\int Q_s(\Theta', d\Theta)\,\nu_\Theta.$$

Then for $s\in[0,\infty]$:

(1.6) $$\mathscr{L}((Y_{s\beta(N)+t}^N)_{t\in\mathbf{R}^+}) \underset{N\to\infty}{\Rightarrow} \mathscr{L}((Y_t^{\nu(s)})_{t\in\mathbf{R}^+})$$

and for $s>t$:

(1.7) $$\mathscr{L}(Y_{s\beta(N)}^N \mid |\Theta_{t\beta(N)}^N-\tilde{\theta}|\leqq \varepsilon_N) \underset{N\to\infty}{\Rightarrow} \int Q_{s-t}(\tilde{\theta}, d\Theta)\,\nu_\Theta,$$

$$\text{with } \varepsilon_N\to 0 \quad as \quad N\to\infty.$$

We now discuss the result. Thus on a microscopic time scale, the system is locally in an equilibrium determined by the current value of Θ_t^N which evolves on a macroscopic time scale. In other words the system is *slaved* to the macroscopic observable Θ which itself is a one dimensional diffusion with the new generator $u_g(x)\left(\frac{\partial}{\partial x}\right)^2$.

Putting $s=0$ we recapture the statement that: up to times $L(N)=o(N)$ but with $L(N)\to\infty$, we see a system that looks like an equilibrium of the infinite system, the one in whose domain of attraction μ lies. For $s=+\infty$ we have:

$$Q_\infty(\Theta', \cdot)=\theta'\,\delta_{\mathbf{0}}+(1-\theta')\,\delta_{\mathbf{1}},$$

so that for times $\gg N$ the system behaves like a finite system and is absorbed by the traps **0** and **1**.

One more remark on the scale is appropriate: in our models the change in the slowly varying variable Θ_s^N is driven by *fluctuations* of the components from local equilibrium, since Θ_s^N is a martingale for fixed N. This results in the scale of the form $\beta(N)=$system size. There are cases where the change in the slowly varying variable is caused by *large deviations* of all the components from the "current equilibrium", then we get $\beta(N)\sim e^{\gamma N}$ as for example in the contact process or ferromagnetic systems (cf. [5]).

The specific feature of the systems we consider here is that the estimator Θ^N of the conserved quantity Θ of the infinite system can be chosen such that:

$$E^{\mu}(\Theta_s^N)=E^{\mu}(\Theta_0^N) \qquad \forall s\in\mathbf{R}^+,$$

which then implies that Z_s is not only a Markov process (which holds in general) but it will be a *martingale*. An additional technical advantage in our situation is that already for fixed N, Θ_s^N is a martingale.

Notice that the existence of a time scale $\beta(N)$ is connected with the existence of multiple equilibria for the infinite system. If the infinite system has a unique equilibrium ν (compare [3]), then instead of (1.6) one would have

$$\mathscr{L}(Y^N(T(N)))\Rightarrow\nu \qquad \forall T(N)\to\infty \quad \text{as} \quad N\to\infty. \tag{1.8}$$

Complementing the results of Theorem 1 is a result on the asymptotics of the *trapping times* of the system of size $2N+1$. Define

$$T^N=\inf(t\,|\,y_i^N(t)=0 \quad \forall i\in[-N,N] \quad \text{or} \quad y_i^N(t)=1 \quad \forall i\in[-N,N]), \tag{1.9}$$

$$T=\inf(t\,|\,Z_t\in\{0,1\}). \tag{1.10}$$

Corollary 1 *Assume that* $g(x)/x^{\alpha}\to c'$ *as* $x\to 0$, $g(x)/(1-x)^{\alpha}\to c''$ *as* $x\to 1$ *for some* $\alpha\in[1,\infty)$; $c', c''\in(0,\infty)$. *Then we have* (δ_{∞} *is allowed here as limit law) for* $\alpha\neq 2$:

$$\mathscr{L}(T^N/\beta(N))\underset{N\to\infty}{\Rightarrow}\mathscr{L}(T). \tag{1.11}$$

In the case $\alpha=2$ *we need* $c', c''\leqq 1$.

This is of interest for genetic models: Suppose $(y_i^N(t))$ describes the vector of frequencies of a gene in $2N+1$ colonies of a population at time t. A standard genetic model corresponds to $g(x)=x(1-x)$ (Fisher-Wright model). Then for large N we can approximate the global *fixation time*, that is the time until the gene is present in every individual or totally extinct, by $(2N+1)$ T, and the distribution of T is fairly easy to calculate using classical diffusion theory.

Our last result in this section is more of an observation and is concerned with an special property of the system of *Fisher-Wright diffusions*, which will turn out to be very important in the study of *infinite systems*, see [6]. Consider the (nonlinear) map which maps each diffusion coefficient $g(x)$ onto $u_g(x)$, the diffusion coefficient of the associated diffusion Z_s.

$$F\colon\; g\to u_g(x)=\int g(y)\,\Gamma_x(dy) \tag{1.12}$$

(Note Γ_x depends on g so that F is *nonlinear*.)

Theorem 2

$$F(dx(1-x))(\Theta)=\frac{d}{d+1}\,\Theta(1-\Theta), \qquad d>0. \tag{1.13}$$

In a future paper [6] we establish that $x(1-x)$ is the only function g with the property that for all $d\in\mathbf{R}^+$: $F(dg)=\text{const.}\,g$ and which lies in the class g is Lipschitz, $g(x)>0$ for $x\neq 0, 1$, $g(0)=g(1)=0$.

Remark. Various other models show the very same features as described in Theorems 1, Corollary 1 and Theorem 2. For example choose

(i) $g(x)>0$ for $x>0$ and $g(0)=0$, g Lipschitz, $g(x)\leqq x^2$.
(ii) $g(x)>0$ for $x\in\mathbf{R}$, $C\geqq g(x)\geqq c>0$, g Lipschitz.

Then again Theorem 1 and Corollary 1 hold (Corollary 1 only for the first case) and for Theorem 2 we get $g(x)=dx$ (continuous state branching) as the "fixed points" of F in case (i) and $g(x)\equiv\sigma^2$ (critical Ornstein-Uhlenbeck) in case (ii). The only changes in the proofs come from the fact that now we have to make sure that $E(x_i^N(s\beta(N))^2$ exists for every s and remains bounded as $N\to\infty$. Also the calculations proving Theorem 2 have to be modified. In order to keep the proofs transparent we have focussed on the case $g(x)>0$ in $(0,1)$. Similarly, we could have included below the models (i) and (ii) in a weak external field and in hierarchical situations, but again for reasons of simplicity we restrict ourselves to the case $g(x)>0$ in $(0,1)$.

(b) Meanfield model with weak external field

The models we discussed so far have the property that the finite system ends up in a trap and accordingly the macroscopic observable Z_s also exhibits this property. The same is true for the models studied in Cox and Greven [3, 4] (voter model, critical branching, contact process).

We shall now discuss a modification of our basic model which will have the property that the system of size N approaches an honest equilibrium as $t\to\infty$ and so does Z_t. The initial laws are i.i.d. again.

Given $A\in(0,1)$ and i.i.d. Brownian motions $\{w_i(t)\}_{i\in\mathbf{Z}}$, consider X_t^N

$$(1.14)\qquad \begin{aligned} &X_t^N=\{x_i^N(t)\}_{i\in[-N,N]}\\ &dx_i^N(t)=\left(\frac{1}{2N+1}\right)\sum_{j=-N}^{N}(x_j^N(t)-x_i^N(t))\,dt+c(N)(A-x_i^N(t))\,dt\\ &\qquad+\sqrt{2g(x_i^N(t))}\,dw_i(t).\end{aligned}$$

Note that if $c(N)$ is $o(1)$, then the McKean Vlasov limit dynamics is still given by the infinite system: $X_t=(x_i(t))_{i\in N}$ with

$$(1.15)\qquad dx_i(t)=(\Theta-x_i(t))\,dt+\sqrt{2g(x_i(t))}\,dw_i(t),\qquad \Theta=E^{\mu}(x_0).$$

This system has a one parameter set of extremal equilibria $\{\nu_\Theta\}_{\Theta\in[0,1]}$:

$$(1.16)\qquad \nu_\Theta=(\Gamma_\Theta)^{\otimes\mathbf{Z}}.\quad \Gamma_\Theta\text{: equilibrium of the diffusion generated by}$$
$$A_\Theta=(\Theta-x)\frac{\partial}{\partial x}+g(x)\left(\frac{\partial}{\partial x}\right)^2.$$

However the additional term with coefficient $c(N)=cN^{-1}$ plays a important role in the critical $(O(N))$ and large $(\gg N)$ time scales.

Theorem 3 *Consider the case* $c(N)=cN^{-1}$.

(a) *The assertions* (1.5)–(1.7) *of Theorem* 1 *hold for* X_t^N, *and* X_t *with* $\beta(N)=2N+1$ *and* Z_s, $Q_s(\cdot,\cdot)$ *being the diffusion semigroup with generated by:*

$$c(A-x)\frac{\partial}{\partial x}+u_g(x)\left(\frac{\partial}{\partial x}\right)^2 \qquad u_g(x)=\int g(y)\,\Gamma_x(dy). \tag{1.17}$$

(b) *For all* $\Theta'\in[0,1]$, $Q_s(\Theta',\cdot)\Rightarrow\Lambda_A(\cdot)$ *as* $s\to\infty$ *and* λ_A *is the unique invariant measure of the semigroup* $Q_s(\cdot,\cdot)$.

For $L(N)/N\to\infty$ *as* $N\to\infty$:

$$\mathscr{L}((X^N_{L(N)+s})_{s\in\mathbb{R}^+})\underset{N\to\infty}{\Rightarrow}\mathscr{L}((X_s^{\nu^A})_{s\in\mathbb{R}^+}) \quad \textit{with} \quad \nu^A=\int\Lambda_A(d\Theta)\,\nu_\Theta. \tag{1.18}$$

(c) *Let* $\nu^{A,N}$ *be the equilibrium of* X_t^N *for a given value of* A *(and fixed* N*), then*

$$\nu^{A,N}\underset{N\to\infty}{\Rightarrow}\int\Lambda_A(d\Theta)\,\nu_\Theta. \tag{1.19}$$

A nice interpretation comes from genetics. Let $x_i^N(t)$ denote the frequency of a gene of type I in the colony i at time t. The $2N+1$ colonies are imbedded in a bigger environment to which they can emigrate, respectively from which *immigration* occurs. Let us suppose that the gene I has frequency A in the larger population. The total exchange with the environment and the $2N+1$ colonies is kept of order 1 as N gets larger. Furthermore migration occurs between the N colonies. A good model for this is given by (1.14).

Then Theorem 3 says that for times $\gg N$ the situation will stabilize in a state which is a mixture in Θ of states with independent components and mean Θ and the mean of the mixing measure is A. Making $c(N)N^{-1}\to\infty$ would cause the mixing measure shrink to δ_A, while $c(N)/N\to 0$ results in the limit measure $\Theta'\delta_0+(1-\Theta')\delta_1$. Therefore we see that the quantity $c(N)N^{-1}$ regulates the correlation between components and the formation of clusters for very large times.

(c) Two level hierarchical mean-field model

So far we have considered *finite* systems, where the system size N tends to infinity and the migration term is suitably scaled. The same multiple scale analysis should be applicable for *infinite* systems, where we also scale the migration term appropriately. An important step in this (difficult) direction is to study a two level hierarchical model, but which is interesting in its own right. Suppose we simulate a very large system on the computer but on the screen we display a large window but forming only a small fraction of the entire system. On the screen we focus now on local patterns. What local patterns do we observe here for different periods of observation? We expect that the large system changes its macroscopic properties very slowly compared to the process visible on the screen. Therefore on *small time scales* the big system provides stable boundary conditions for the system visible on the screen. The global properties of this subsystem will therefore relax into some equilibrium distribution dictated by these boundary conditions. The actual value of them will determine an equilibrium of the infinite system, which describes what we see locally in the window of observation. On a second *large timescale* of course the big system will fluctuate and so will these boundary conditions and therefore the system on the screen will change accordingly to another *quasi equilibrium.*

One would expect this kind of behavior from a large variety of systems with local interaction. In particular interesting behavior is expected from systems exhibiting multiple phases including voter models ($d \geqq 3$), stochastic Ising models below critical temperature, interacting diffusions with external drift field etc. These models are difficult to analyse since the interaction geometry is fairly complicated. In order to obtain rigorous results we therefore start by studying a simplified caricature of the situation described above, namely, a *two level hierarchical mean-field model.* These results turn out to be the cornerstone for a theory of *infinite hierarchical systems* (see [6, 7]).

Consider $M(N)$ blocks of size $2N$ that is, a system

$$X_t^N = (x_i^N(t))_{i \in [-(2M(N)+1)N, (2M(N)+1)N-1]}, \qquad M(N) \in \mathbf{N},$$

with components satisfying the system of stochastic differential equations (1.20) below. Let

$$I_N(i) = [(2j-1)N, (2j+1)N-1], \quad \text{if } (2j-1)N \leqq i < (2j+1)N$$

$$\bar{x}_j(t) = \frac{1}{2N} \sum_{(2j-1)N \leqq k < (2j+1)N} x_k^N(t).$$

$$(1.20) \qquad \begin{aligned} dx_i^N(t) &= \frac{c}{2N} \sum_{j \in I_N(i)} (x_j^N(t) - x_i^N(t))\, dt \\ &+ \frac{a}{2M(N)+1} \sum_{j=-M(N)}^{M(N)} (\bar{x}_j^N(t) - \bar{x}_k^N(t))\, dt \\ &+ \sqrt{2g(x_i^N(t))}\, dw_i(t) \quad \text{if } i \in [2(k-1)N, (2k+1)N). \end{aligned}$$

The initial law μ is i.i.d. In the sequel we will write everything for the case $a=1$ to simplify the notation.

In order to describe the situation we introduce some objects which generalize the picture from Sect. 1a (ii). In particular due to the *two levels,* Θ is replaced by the pair Θ, A; similarly Z_s is replaced by $Z_s^{1,A}$, $Z_s^{2,\Theta}$ and Θ_s^N is replaced by $\Theta_s^{1,N}$, $\Theta_s^{2,N}$ and the time scale $\beta(N)$ by two time scales $\beta_1(N)$, $\beta_2(N)$:

I. ν_Θ is the unique invariant measure of the system $X_t = (x_i(t))_{i \in \mathbf{Z}}$ solving

$$(1.21) \qquad dx_i(t) = (\Theta - x_i(t))\, dt + \sqrt{2g(x_i(t))}\, dw_i(t), \qquad i \in \mathbf{Z}.$$

II. The time scales are: $\beta_1(N) = 2N$, $\beta_2(N) = 2N(2M(N)+1)$.

III. $Z_s^{1,A}$, $Q_s^{1,A}$ is the diffusion respectively transition kernel generated by the differential operator:

$$(1.22) \qquad c(A-x)\frac{\partial}{\partial x} + v_g(x)\left(\frac{\partial}{\partial x}\right)^2, \qquad v_g(x) = E^{\nu_x}(g(x_0)).$$

We start this diffusion in its equilibrium Γ_A. (Note A in $Z^{1,A}$ does *not* refer to an initial state!)

IV. $Z_s^{2,\Theta}$, $Q_s^2(\cdot,\cdot)$ is the diffusion respectively its transition kernel with initial value $Z_0^{2,\Theta'}=\Theta'$ generated by the differential operator:

$$u_g(x)\left(\frac{\partial}{\partial x}\right)^2, \quad \text{where} \quad u_g(x)=\int v_g(y)\,\Gamma_x(dy). \tag{1.23}$$

V. The slowly varying functions on the two levels are $\Theta_s^{1,N}$, $\Theta_s^{2,N}$ given by

$$\Theta_s^{1,N}=\frac{1}{2N}\sum_{i=-N}^{N-1} x_i^N(s) \tag{1.24}$$

$$\Theta_s^{2,N}=\frac{1}{2M(N)+1}\sum_{j=-M(N)}^{M(N)} \bar{x}_j^N(s). \tag{1.25}$$

With these objects we are able to formulate our results as follows:

Theorem 4 *Let $M(N)\to\infty$ as $N\to\infty$.*

a) Then $\Theta^{2,N}$ fluctuates in the time scale $\beta_2(N)$ and is described in the limit by the diffusion $Z_s^{2,\Theta'}$; and $\Theta^{1,N}$ fluctuates on the smaller time scale $\beta_1(N)$ and is described in the limit by the diffusion $Z^{1,Z_s^{2,\Theta'}}$. That is,

$$\mathscr{L}((\Theta^{2,N}_{s\beta_2(N)})_{s\in\mathbf{R}^+}) \underset{N\to\infty}{\Rightarrow} \mathscr{L}((Z_s^{2,\Theta'})_{s\in\mathbf{R}^+}), \tag{1.26}$$

$$\mathscr{L}((\Theta^{1,N}_{s\beta_2(N)+u\beta_1(N)})_{u\in\mathbf{R}^+}) \underset{N\to\infty}{\Rightarrow} \mathscr{L}((Z_u^{1,Z_s^{2,\Theta'}})_{u\in\mathbf{R}^+}). \tag{1.27}$$

b) The components of the interacting system relax into the equilibrium determined by these processes, that is, after a very long time the process looks as follows:

$$\mathscr{L}((X^N_{s\beta_2(N)+u})_{u\in\mathbf{R}^+}) \underset{N\to\infty}{\Rightarrow} ((X_u^{v(s)})_{u\in\mathbf{R}^+}) \qquad s>0, \tag{1.28}$$

with $v(s)=\int Q_s^2(\Theta',dA)\int\Gamma_A(d\Theta)\,v_\Theta$, while after a shorter time $L(N)$ we have

$$\mathscr{L}((X^N_{L(N)+u})_{u\in\mathbf{R}^+}) \underset{N\to\infty}{\Rightarrow} \mathscr{L}((X_u^{v(\Theta')})_{u\in\mathbf{R}^+}), \tag{1.28'}$$

here $L(N)\to\infty$, $L(N)=o(\beta_2(N))$ and $v(\Theta')=\int\Gamma_{\Theta'}(d\Theta)\,v_\Theta$.

Remark. The essential point here is that in (1.28) instead of $\int Q_s(\Theta',d\Theta)\,v_\Theta$ as in Theorem 1 or $\int\Gamma_A(d\Theta)\,v_\Theta$ as in Theorem 3 we have now a combination of these operations and obtain

$$\int Q_s^2(\Theta',dA)\,\mu_A \quad \text{with} \quad \mu_A=\int\Gamma_A(d\Theta)\,v_\Theta,$$

so that we find more correlation in this system imbedded in a larger one than in the case of a free system. We obtain a *mixture of states of a system in a weak external field*, the mixing measure being a kernel of a diffusion. Furthermore the density $\Theta_{sN}^{1,N}$ visible on the screen of observation is now subordinated to the background diffusion $Z_s^{1,\Theta'}$, which is fluctuating only on a much larger timescale.

(d) Organization of proofs

The proof of our results are given in Sects. 2–4. The approach we take is guided by the goal to give proofs which can be modified to treat infinite hierarchical models (see [6, 7]) and lattice models and furthermore are in line with the structure of proof needed to get general statements for systems with interacting components (such as interacting Fisher-Wright diffusions, voter model etc.).

In Sects. 2 and 3 we prove the main ingredients needed to deduce Theorem 1. In Sect. 2 we prove, that if the density process $(\Theta^N_{s\beta(N)})$ is tight, then $\mathscr{L}(Y^N_{s\beta(N)})$ has as weak limit points as $N\to\infty$, mixtures of equilibria of the infinite system. In Sect. 3 we then use these results to prove that $\Theta^N_{s\beta(N)}$ in fact converges and with the results of Sect. 2 we can then also identify the limit.

Finally, Sect. 4(a) uses the results obtained so far to prove Theorem 1. In the same subsection we prove Corollary 1 and Theorem 2. Section 4(b) outlines how to modify the proofs to get Theorem 3 and so does Sect. 4(c) for Theorem 4.

2 Convergence to mixtures of equilibria

(i) Formulation and proof of Proposition 1

In this paragraph we shall establish that the finite systems in the time scale sN look like mixtures of equilibrium states of the infinite system. The actual identification of the mixing measure and the proof of convergence of $\mathscr{L}(Y^N_{sN})$ as $N\to\infty$ will be done in the next section.

Proposition 1 *Let $(N_k)_{k\in N}$ be a subsequence and $L(N)\uparrow\infty$, $L(N)=o(N)$ such that:*

(2.1)
$$\mathscr{L}(\Theta^{N_k}_{s\beta(N_k)}) \underset{k\to\infty}{\Rightarrow} P_s(\cdot), \qquad \mathscr{L}\big(\sup_{t\leq L(N)} |\Theta^N_{s\beta(N)}-\Theta^N_{s\beta(N)+t}|\big) \underset{N\to\infty}{\Rightarrow} \delta_0$$
$$\mathscr{L}(Y^{N_k}_{s\beta(N_k)}) \underset{k\to\infty}{\Rightarrow} \nu(s).$$

Then $\nu(s)$ necessarily has the form

(2.2)
$$\nu(s)=\int P_s(d\Theta)\,\nu_\Theta.$$

The proof will be based first of all on some facts from the ergodic theory of the infinite system which is collected in two Lemmata 2.1, 2.2, a comparison of infinite and finite systems (Lemma 2.3), and a comparison of two infinite systems starting with initial distributions close to each other (Lemma 2.4).

We shall state these results first, then derive the proposition from them and finally we shall prove the lemmata step by step. For a Polish space E let $M_1(E)$ denote the space of probability measures on E. A measure $\mu\in M_1([0,1]^{\mathbf{Z}})$ is called $L_2(\mu)$ ergodic, if the subspace of $L_2(\mu)$ which is invariant under the shift on $\mathbf{Z}$ contains only the constant functions.

Lemma 2.1 (Ergodic theorem for the infinite system) *Let μ be a translation invariant probability measure on $[0,1]^{\mathbf{Z}}$ which has the property*

(1) μ *is* $L_2(\mu)$ *ergodic*

(2) $\int y_0\,d\mu=\rho<\infty.$

Then we have for the system Y_t given by (1.2):

$$\mathscr{L}(Y_t) \underset{N\to\infty}{\Rightarrow} \nu_\rho, \tag{2.3}$$

with $\nu_\rho = (\Gamma_\Theta)^{\otimes \mathbf{Z}}$, $\Gamma_x(dz) = C \frac{1}{g(y)} \exp\left(\int_x^y \frac{x-z}{g(z)} dz\right) dy.$

Lemma 2.2 *The mapping* $[0,1] \to M_1([0,1]^{\mathbf{Z}})$ *given by*

$$\Theta \to \nu_\Theta$$

is continuous. Furthermore if f is a Lipschitz function of a single component k, define

$$\hat{f}\colon \hat{f}(\Theta) = E^{\nu_\Theta} f(y_k).$$

The function $\hat{f}$ is also Lipschitz.

Lemma 2.3 (Comparison of finite and infinite systems) *Fix some $s \in \mathbf{R}^+$ and a sequence $L(N)$. Denote by $Y_t^{\mu_N}$ the infinite system starting at $t=0$ in the distribution defined by continuing the configuration of $Y^N_{s\beta(N)-L(N)}$ periodically to $[0,1]^{\mathbf{Z}}$. Similarly, consider Y_t^N as element of $[0,1]^{\mathbf{Z}}$. Assume that $L(N)$ has the property $L(N) \to \infty$ as $N \to \infty$ but $L(N) = o(N)$. Assume that*

$$\sup_{t \leq L(N)} |\Theta^N_{s\beta(N)} - \Theta^N_{s\beta(N)-t}| \underset{N\to\infty}{\Rightarrow} 0 \quad \textit{in probability.}$$

Then the following holds:

$$|E f(Y^{\mu_N}_{L(N)}) - E f(Y^N_{s\beta(N)})| \underset{N\to\infty}{\Rightarrow} 0, \qquad \forall f \in \mathscr{C}([0,1]^{\mathbf{Z}}, \mathbf{R}). \tag{2.4}$$

Lemma 2.4 (Uniformity of the ergodic theorem for the infinite system) *Let $Y_t^{\mu_N}$ be defined as above in Lemma 2.3. Since $\{\mu_N\}$ is relatively weakly compact, we can find convergent subsequences. Let $\mu = \lim_{k\to\infty} \mu_{N_k}$ for some such sequence N_k. Then for any sequence $L(N)$ with $L(N) \to \infty$:*

$$E f(Y^{\mu_{N_k}}_{L(N_k)}) - E f(Y^{\mu}_{L(N_k)}) \underset{k\to\infty}{\longrightarrow} 0. \tag{2.5}$$

Lemma 2.5 (Stability of the estimator of the conserved quantity) *Let $(y_i^N)_{i\in[-N,N]}$ be distributed as the restriction of a translation invariant measure μ_N on $[0,1]^{\mathbf{Z}}$ to $[0,1]^{[-N,N]}$. Suppose $\mu_N \Rightarrow \mu$. Define a random variable φ on $(\mu, [0,1]^{\mathbf{Z}})$:*

$$\varphi = \lim_{n\to\infty} \frac{1}{2n+1} \sum_{i=-n}^{+n} y_i \tag{2.6}$$

and random variables φ_N *on* $(\mu_N, [0,1]^{\mathbf{Z}})$:

(2.7) $$\varphi_N = \frac{1}{2N+1} \sum_{i=-N}^{N} y_i^N.$$

Then

(2.8) $$\mathscr{L}(\varphi_N) \underset{N\to\infty}{\Rightarrow} \mathscr{L}(\varphi).$$

Proof of Proposition 1

Since $M_1([0,1]^{\mathbf{Z}}$ is compact, we can choose a subsequence $\tilde{N}_k$ of N_k such that for a given sequence $L(N)$:

(2.9) $$\mathscr{L}(Y^{\tilde{N}_k}_{s\beta(\tilde{N}_k)-L(\tilde{N}_k)}) \text{ converges weakly.}$$

The limit, denoted by μ, is translation invariant. We shall choose a sequence $L(N)$ with $L(N)\uparrow +\infty$ as $N\to\infty$, $L(N)=o(N)$. In the sequel we suppress the $\sim$ on N_k. Condition on φ (see (2.6)) and write

(2.10) $\mu = \int_{[0,1]} \mu_\rho \, d\Lambda(\rho)$. Here μ_ρ is a measure on $[0,1]^{\mathbf{Z}}$ which is translation invariant, $L_2(\mu_\rho)$ ergodic and $\int y_0 \, d\mu_\rho = \rho$. $\Lambda(\cdot)$ is a probability measure on $[0,1]$.

By assumption (2.1) we know

(2.11) $$\mathscr{L}(\Theta^{N_k}_{s\beta(N_k)}) \underset{k\to\infty}{\Rightarrow} P_s(\cdot)$$

and

(2.12) $$\mathscr{L}(\Theta^{N_k}_{s\beta(N_k)-L(N_k)}) \underset{k\to\infty}{\Rightarrow} P_s(\cdot).$$

Since Λ in (2.10) is given by

(2.13) $$\Lambda = \mathscr{L}\left(\lim_{N\to\infty}\left(\frac{1}{2N+1}\sum_{i=-N}^{N} y_i^\infty\right)\right) \quad \text{with } \mathscr{L}((y_i^\infty)_{i\in\mathbf{Z}}) = \mu,$$

Lemma 2.5 tells us that

(2.14) $$\Lambda(\cdot) = P_s(\cdot)$$

and therefore

(2.15) $$\mu = \int_{[0,1]} \mu_\rho \, P_s(d\rho).$$

Now apply Lemma 2.1 to the $L_2(\mu)$ ergodic components μ_ρ in the decomposition (2.10) of the measure μ and conclude

$$(2.16)\qquad \mathscr{L}(Y^{\mu}_{L(N)}) \underset{N\to\infty}{\Rightarrow} \int \nu_\rho\, P_s(d\rho).$$

Let μ_N be the distribution of $Y^N_{s\beta(N)-L(N)}$ continued periodically to $[0,1]^{\mathbf{Z}}$. Write now

$$(2.17)\qquad \begin{aligned}\mathscr{L}(Y^{N_k}_{s\beta(N_k)}) &= (\mathscr{L}(Y^{N_k}_{s\beta(N_k)}) - (\mathscr{L}(Y^{\mu_{N_k}}_{L(N_k)}))\\ &\quad + (\mathscr{L}(Y^{\mu_{N_k}}_{L(N_k)}) - \mathscr{L}(Y^{\mu}_{L(N_k)})) + \mathscr{L}(Y^{\mu}_{L(N_k)}).\end{aligned}$$

By Lemma 2.3, respectively 2.4, the first two signed measures converge weakly to the 0-measure and by (2.16) above the third summand converges to $\int \nu_\rho\, P_s(d\rho)$, therefore

$$(2.18)\qquad \mathscr{L}(Y^{N_k}_{s\beta(N_k)}) \underset{k\to\infty}{\Rightarrow} \int \nu_\rho\, P_s(d\rho) \qquad \text{q.e.d.}$$

(ii) Proof of Lemma 2.1 For a shift ergodic measure μ with $E^{\mu}(g_0)=\Theta$ the system consists of independent components evolving according to

$$d\,y_i(t) = (\Theta - y_i(t))\,dt + \sqrt{2g(y_i(t))}\,d\,w_i(t).$$

It suffices therefore to show that each component will tend to the unique invariant measure of the diffusion given above. These are classical results and we shall not elaborate on them further.

(iii) Proof of Lemma 2.3 The proof uses the concept of coupling and proceeds in six steps.

I. We start by rewriting the problem in a more convenient form. First view the system of size N with initial distribution $\mathscr{L}(Y^N_{s\beta(N)-L(N)})$ as a system on $[0,1]^{\mathbf{Z}}$ by periodic continuation and rewrite the drift term:

$$(2.19)\qquad \begin{aligned} d\,y_i^N(t) &= (\Theta^N - y_i^N(t))\,dt + \left(\frac{1}{2N+1}\sum_{j=-N}^{N}(y_j^N(t)-\Theta^N)\right)dt\\ &\quad + \sqrt{2g(y_i^N(t))}\,d\,\tilde w_i(t), \qquad i\in\mathbf{Z}\\ \Theta^N &= \frac{1}{2N+1}\sum_{j=-N}^{N} y_j^N(0).\end{aligned}$$

Here $\tilde w_{i+k(2N+1)}(t) = w_i(t)$ for $i\in[-N,N]$, and $\{w_i(t)\}_{i\in\mathbf{Z}}$ are i.i.d. Brownian motions.

This system has to be compared with the following system of equations the solution of which we have called $\{y_i^{\mu_N}(t)\}_{i\in\mathbf{Z}}$.

$$(2.20)\qquad \begin{aligned} d\,y_i(t) &= (\Theta^N - y_i(t))\,dt + \sqrt{2g(y_i(t))}\,d\,w_i(t), \qquad i\in\mathbf{Z}\\ \Theta^N &= \frac{1}{2N+1}\sum_{i=-N}^{+N} y_i(0), \qquad \mathscr{L}(Y_{(0)}) = \mu_N.\end{aligned}$$

That is, our problem can be formulated abstractly as follows: Compare two systems $X(t)=(x_i^N(t))_{i\in\mathbf{Z}}$ and $Z(t)=(z_i^N(t))_{i\in\mathbf{Z}}$ defined as:

$$dx_i^N(t)=f^N(t,\omega)\,dt+(\Theta^N-x_i^N(t))\,dt+\sqrt{2g(x_i^N(t))}\,d\tilde{w}_i(t) \tag{2.21}$$

with $\tilde{w}^{(t)}_{i+(2N+1)k}=w_i(t)$ for $i\in[-N,N]$ and $k\in\mathbf{Z}$, with $\{w_i(t)\}_{i\in[-N,N]}$ i.i.d. Brownian motions. The nonanticipating stochastic process $f^N(t,w)$ will be specified later.

$$dz_i^N(t)=(\Theta^N-z_i^N(t))\,dt+\sqrt{2g(z_i^N(t))}\,dw_i(t). \tag{2.22}$$

II. The problem of comparing two such systems $X(t)$, $Z(t)$ as in (2.21), (2.22) (N is fixed for the moment and suppressed in the notation) is now approached by the *coupling method.* We shall introduce a coupled dynamics, that is a dynamic of the bivariate process $(\bar{X}(t), \bar{Z}(t))$ such that:

$$\begin{aligned}\mathscr{L}(\bar{X}(t))&=\mathscr{L}(X(t))\\ \mathscr{L}(\bar{Z}(t))&=\mathscr{L}(Z(t))\\ L(\bar{X}(0),\bar{Z}(0))&=\mathscr{L}(X(0))\otimes\mathscr{L}(Z(0))\end{aligned} \tag{2.23}$$

and the property that the difference between $\bar{X}(t)$, $\bar{Z}(t)$ becomes small as $t\to\infty$. This can be achieved in our situation by using the same collection of $\{w_i(t)\}$ for both $X(t)$ and $Z(t)$ in those components with index $i\in[-N,N]$ and using the fact that the equations have a unique strong solution. The coupled dynamics is given by the following system of stochastic differential equations:

$$\begin{aligned}d\bar{x}_i^N(t)&=f^N(t,\omega)\,dt+(\theta^N-\bar{x}_i^N(t))\,dt+\sqrt{2g(\bar{x}_i^N(t))}\,d\omega_i(t)\\ d\bar{z}_i^N(t)&=(\theta^N-\bar{z}_i^N(t))\,dt+\sqrt{2g(\bar{z}_i^N(t))}\,d\omega_i(t)\qquad\text{for all } i\in[-N,N]\end{aligned} \tag{2.24}$$

$$\begin{aligned}\bar{x}^N_{i+(2N+1)k}(t)&=\bar{x}_i^N(t) &\quad i\in[-N,N],k\in\mathbf{Z}\setminus\{0\}\\ \bar{z}^N_{i+(2N+1)k}(t)&=z^N_{i+(2N+1)k} &\quad i\in[-N,N],k\in\mathbf{Z}\setminus\{0\}.\end{aligned} \tag{2.25}$$

We write the system immediately in a form suitable for studying $\bar{x}_i^N(t)-\bar{z}_i^N(t)$):

$$\begin{aligned}d\bar{x}_i^N(t)=&f^N(t,\omega)\,dt+(\Theta^N-\bar{x}_i^N(t)\wedge\bar{z}_i^N(t))\,dt+\sqrt{2g(\bar{x}_i^N(t))}\\ &\wedge\sqrt{2g(\bar{z}_i^N(t))}\,dw_i(t)-(\bar{x}_i^N(t)-\bar{z}_i^N(t))^+\,dt+(\sqrt{2g(\bar{x}_i^N(t))}\\ &-\sqrt{2g(\bar{z}_i^N(t))})^+\,dw_i(t)\qquad\text{for all } i\in[-N,N]\\ d\bar{z}_i^N(t)=&(\Theta^N-\bar{x}_i^N(t)\wedge\bar{z}_i^N(t))\,dt+(\sqrt{2g(\bar{x}_i^N(t))}\wedge\sqrt{2g(\bar{z}_i^N(t))})\,dw_i(t)\\ &-(\bar{x}_i^N(t)-\bar{z}_i^N(t))^-\,dt+(\sqrt{2g(\bar{x}_i^N(t))}-\sqrt{2g(\bar{z}_i^N(t))})^-\,dw_i(t)\\ &\text{for all } i\in[-N,N].\end{aligned} \tag{2.26}$$

In the second part of (2.25) we use that the components of $Z(t)$ evolve independent of each other.

In the following i will be always an element in $[-N, N]$. From the relations above we obtain the following dynamics for $|\bar{z}_i^N(t)-\bar{x}_i^N(t)|$:

$$(2.27)\quad d|\bar{x}_i^N(t)-\bar{z}_i^N(t)|=\operatorname{sgn}(\bar{x}_i^N(t)-\bar{z}_i^N(t))\, f_{(t,\omega)}^N\, dt-|\bar{x}_i^N(t)-\bar{z}_i^N(t)|\, dt$$
$$+\operatorname{sgn}(\bar{x}_i^N(t)-\bar{z}_i^N(t))(\sqrt{2g(\bar{x}_i^N(t))}-\sqrt{2g(\bar{z}_i^N(t))})\, dw_i(t)$$

using a result by Yamada and Watanabe [12, pp. 165, 166] to overcome the difficulty that $|x-y|$ is not a smooth function in $\mathbf{R}^2\to\mathbf{R}$. Here $\operatorname{sgn} x=-1$ for $x\leqq 0$, $+1$ for $x>0$.

III. We define now the Lyapunov function $h^N(t)$:

$$(2.28)\qquad h^N(t)=E|\bar{z}_0^N(t)-\bar{x}_0^N(t)|$$

and obtain from (2.27)

$$(2.29)\qquad \frac{d}{dt}h^N(t)=-h^N(t)+G^N(t)$$

where

$$(2.30)\qquad G^N(t)=E(\operatorname{sgn}(\bar{x}_i^N(t)-\bar{z}_i^N(t))\, f^N(t,\cdot)).$$

The linear differential equation in (2.29) can be solved explicitly as:

$$(2.31)\qquad h^N(t)=h^N(0)\, e^{-t}+\int_0^t e^{-(t-s)}G^N(s)\, ds$$

and this given us the relation

$$(2.32)\qquad h^N(T)\leqq h^N(0)\, e^{-T}+\sup_{t\leqq T}[|G^N(t)|].$$

IV. Now we shall estimate the probability of the event that $(\bar{x}_0^N(t), \bar{z}_0^N(t))$ are further than ε away from the diagonal:

$$(2.33)\quad \operatorname{Prob}(|\bar{z}_0^N(T)-\bar{x}_0^N(T)\geqq\varepsilon)\leqq\frac{1}{\varepsilon}E|\bar{z}_0^N(T)-\bar{x}_0^N(T)|$$
$$\leqq\frac{1}{\varepsilon}([E h^N(0)]\, e^{-T})+\frac{1}{\varepsilon}\sup_{t\leqq T}E|f^N(t,\cdot)|.$$

Since we know by assumption that for $L(N)=o(N)$ we have the relation:

$$(2.34)\qquad \sup_{t\leqq L(N)}E\left|\Theta^N-\frac{1}{2N+1}\sum_{i=-N}^{+N}y_i^N(t)\right|\underset{N\to\infty}{\Rightarrow}0,\qquad h^N(0)\leqq 2E^\mu(|y_0|),$$

we apply (2.33) to $T=L(N)$ and $f^N(t,\cdot)=\left(\Theta^N-\frac{1}{2N+1}\sum_{i=-N}^{+N}\bar{x}_i^N(t)\right)$ to get

$$(2.35)\qquad \overline{\lim_{N\to\infty}}\operatorname{Prob}(|\bar{z}_0^N(L(N))-\bar{x}_0^N(L(N))|\geqq\varepsilon)=0.$$

V. Suppose now that f is a Lipschitz function on $[0, 1]$ and define

(2.36) $$\operatorname{Lip}(f) = \sup_{x \neq y} \frac{|f(x) - f(y)|}{|x - y|}.$$

Now we want to see what (2.35) tells us about the difference of the systems (2.19) and (2.20). Using (2.23), the result (2.35) implies that

(2.37) $$\begin{aligned} |E f(y_0^N(L(N))) - E f(y_0(L(N)))| &= |E f(\bar{z}_0^N(L(N))) - E f(\bar{x}_0^N(L(N)))| \\ &\leq E|f(\bar{z}_0^N(L(N))) - f(\bar{x}_0^N(L(N)))| \\ &\leq \operatorname{Lip}(f)\, E|\bar{z}_0^N(L(N)) - \bar{x}_0^N(L(N))| \underset{N \to \infty}{\Rightarrow} 0. \end{aligned}$$

Since the Lipschitz functions are dense in $\mathscr{C}([0, 1])$ we have:

(2.38) $$\overline{\lim_{N \to \infty}} |E f(y_0^N(L(N))) - E f(y_0(L(N)))| = 0 \qquad \forall f \in \mathscr{C}([0, 1]).$$

This argument can be carried out not only for the component with index 0, but for every component k with $|k| \leq L(N)$ (compare 2.26). This completes the proof of Lemma 2.3.

(iv) Proof of Lemma 2.4 and 2.2

We start by proving Lemma 2.4, and the Lemma 2.2 will be a by-product. Note first that the components of the infinite systems evolve independent of each other, so we are really left with showing that for two *one*-dimensional diffusions:

(2.39) $$dA^N(t) = (\theta_N - A^N(t))\, dt + \sqrt{2g(A^N(t))}\, dw(t)$$

(2.40) $$dA(t) = (\Theta - A(t))\, dt + \sqrt{2g(A(t))}\, d\tilde{w}(t)$$

with $w(t)$, $\tilde{w}(t)$ independent Wiener processes, satisfying

(2.41) $$\mathscr{L}(A_{(0)}^N, \Theta_N) \underset{N \to \infty}{\Rightarrow} \mathscr{L}((A(0), \Theta)),$$

the following is true

(2.42) $$\mathscr{L}(A(L(N))) - \mathscr{L}(A^N(L(N))) \underset{N \to \infty}{\Rightarrow} 0.$$

It clear that $\mathscr{L}(A^N(t))$ converges as $t \to \infty$ to a unique equilibrium measure Γ^N, and $\mathscr{L}(A(t))$ converges to a unique equilibrium Γ (compare Lemma 2.1). It can be shown that $\Gamma^N \Rightarrow \Gamma$ and this will be done below. So we have an array μ_t^N of measures with $\mu_t^N \underset{t \to \infty}{\Rightarrow} \mu_\infty^N \underset{N \to \infty}{\Rightarrow} \mu_\infty^\infty$ and the question is whether the trend to equilibrium is uniform enough in N to give (2.42). This can be seen by *coupling* as follows:

Denote by $B^N(t)$, $B(t)$ the *stationary* solutions of (2.39), (2.40). We shall construct a coupling for $A^N(t)$ and $B^N(t)$, $A(t)$ and $B(t)$ and for $B^N(t)$, $B(t)$, so that we are able to estimate as follows:

$$(2.43)\quad |E f(A^N(L(N))) - E f(A(L(N)))| \leqq |E f(A^N(L(N))) - E f(B^N(L(N)))| + |E f(B^N(L(N))) - E f(B(L(N)))| + |E f(A(L(N))) - E f(B(L(N)))|$$

and show that each of these summands on the right tends to 0 as $N \to \infty$.

Recall that a coupling of two stochastic processes $X(t)$, $Y(t)$ is a bivariate process $(\bar{X}(t), \bar{Y}(t))$ such that $\mathscr{L}(\bar{X}(t)) = \mathscr{L}(X(t))$, $\mathscr{L}(\bar{Y}(t)) = \mathscr{L}(Y(t))$. We want to define the bivariate process such that $E|\bar{X}(t) - \bar{Y}(t)|$ gets as small as possible as $t \to \infty$. Again the trick is to use the *same* Brownian motion in both components.

We define the dynamics for the coupling of $A^N(t)$, $B^N(t)$ as follows

$$(2.44)\quad d\bar{A}^N(t) = (\bar{\Theta}_N - \bar{A}^N(t))\,dt + \sqrt{2g(\bar{A}^N(t))}\,dw(t)$$

$$(2.45)\quad d\bar{B}^N(t) = (\bar{\Theta}_N - \bar{B}^N(t)\,dt + \sqrt{2g(\bar{B}^N(t))}\,dw(t)$$

$$(2.46)\quad \mathscr{L}((\bar{A}^N(0), \bar{B}^N(0))) = \mathscr{L}(A^N(0)) \otimes \mathscr{L}(B^N(0)).$$

As in the proof of Lemma 2.3 we obtain this time

$$(2.47)\quad \frac{d}{dt} E|\bar{A}^N(t) - \bar{B}^N(t)| = -E|\bar{A}^N(t) - \bar{B}^N(t)|,$$

so that with the abbreviation $h^N(t) = E|\bar{A}^N(t) - \bar{B}^N(t)|$ we have

$$(2.48)\quad h^N(t) = h^N(0)\,e^{-t} \leqq 2e^{-t}.$$

Similarly, we can perform the above construction for $A(t)$, $B(t)$ and get

$$(2.49)\quad h(t) = h(0)\,e^{-t} \leqq 2e^{-t}.$$

Next we couple $B^N(t)$ and $B(t)$. Here is the dynamics of $(\bar{B}^N(t), \bar{B}(t))$:

$$(2.50)\quad \begin{aligned} d\bar{B}^N(t) &= (\bar{\Theta}_N - \bar{B}^N(t))\,dt + \sqrt{2g(\bar{B}^N(t))}\,dw(t) \\ d\bar{B}(t) &= (\bar{\Theta} - \bar{B}(t))\,dt + (\sqrt{2g(\bar{B}(t))})\,dw(t) \end{aligned}$$

$\mathscr{L}((\bar{B}^N(0), \bar{\Theta}_N), (\bar{B}(0), \bar{\Theta}))$ is chosen such that:

$$(2.51)\quad |\bar{\Theta}_N - \bar{\Theta}| \underset{N \to \infty}{\Rightarrow} 0 \quad \text{in probability.}$$

This last requirement can be met since by assumption $\mathscr{L}(\bar{\Theta}_N) \Rightarrow \mathscr{L}(\bar{\Theta})$.

Define

$$(2.52)\quad g^N(t) = E|\bar{B}^N(t) - \bar{B}(t)|.$$

We have from the dynamics the relation (recall the remark following (2.27)):

$$(2.53)\quad \frac{d}{dt} g^N(t) = -g^N(t) + E((\Theta_N - \Theta)^+ 1(\bar{B}^N_{(t)} \geqq \bar{B}(t)) - (\Theta_N - \Theta)^- 1(\bar{B}^N_{(t)} \leqq \bar{B}(t))).$$

Abbreviate the expectations on the r.h.s. of (2.53) by $L^N(t)$. The Eq. (2.53) is solved explicitly and then bounded as follows:

$$g^N(t)=g^N(0)\,e^{-t}+\int_0^t e^{-(t-s)}L^N(s)\,ds\leqq 2e^{-t}+E|\bar{\Theta}_N-\bar{\Theta}|. \tag{2.54}$$

From this we can deduce

$$\overline{\lim_{N\to\infty}}\; g^N(L(N))\leqq \overline{\lim_{N\to\infty}}\; E|\bar{\Theta}_N-\bar{\Theta}|. \tag{2.55}$$

Now insert (2.48) (2.49), (2.55) into the r.h.s. of (2.43) and obtain for all Lipschitz functions f on $[0, 1]$,

$$\overline{\lim_{N\to\infty}}\; |E\,f(A^N(L(N)))-E\,f(A(L(N)))|\leqq \mathrm{Lip}(f)\;\overline{\lim_{N\to\infty}}\; E|\bar{\Theta}_N-\bar{\Theta}|. \tag{2.56}$$

The right hand side is equal to 0 according to (2.51). Since the Lipschitz functions are dense in $C[0, 1]$, we have proved (2.42), and therefore established Lemma 2.4. □

We now prove Lemma 2.2. From (2.55) it follows especially that we have the following facts for the equilibria of the diffusion in (2.39), (2.40) denoted by Γ^{Θ_N}, Γ^{Θ} (but with Θ_N, Θ just numbers and not random variables):

$$\Gamma^{\Theta_N}\Rightarrow\Gamma^{\Theta} \quad \text{as} \quad \Theta_N\to\Theta \tag{2.57}$$

$$|E^{\Gamma^{\Theta_N}}(f)-E^{\Gamma^{\Theta}}(f)|\leqq \mathrm{Lip}(f)\,|\Theta_N-\Theta| \qquad \text{for } f \text{ Lipschitz.} \tag{2.58}$$

The first relation shows that $\Theta\to\Gamma^{\Theta}$ is a continuous map from $[0, 1]$ $\to M_1([0, 1])$, while the second relation shows that for a Lipschitz function f: $(0, 1]\to\mathbf{R}$, the function $\hat{f}$: $\theta\to E^{\Gamma^{\Theta}}(f)$ is again Lipschitz. This immediately proves Lemma 2.2 since

$$\nu_{\Theta}=(\Gamma_{\Theta})^{\otimes\mathbf{Z}} \tag{2.59}$$

and the components of the infinite system evolve independent.

(v) Some further consequences of the coupling

We shall later need two additional consequences of the couplings we constructed so far. If we combine the couplings from (2.24), (2.25) and (2.44), (2.45), (2.50) by simply using one set $\{w_i(t)_{t\in\mathbf{R}^+}\}$ of Wiener processes driving the stochastic differential equations, we obtain the following consequence of (2.35) and (2.56). We can construct $(Y^N_{s\beta(N)+u})_{u\in\mathbf{R}^+}$, $(Y^{\mu}_u)_{u\in\mathbf{R}^+}$ with $\mu=w-\lim \mathscr{L}(Y^{N_k}_{s\beta(N_k)})$ on one probability space such that:

$$E(|\bar{y}^{N_k}_j(s\beta(N_k))-\bar{y}^{\mu}_j(u)|)\underset{k\to\infty}{\Rightarrow} 0, \qquad \forall u\geqq 0, j\in\mathbf{Z}. \tag{2.60}$$

The other fact we shall need later reads as follows:
Suppose that $\mathscr{L}(Y_0^N)$ and $\mathscr{L}(Y_0)$ can be constructed on one probability space such that

$$\sup_{t \leqq K(N)} E\left|\frac{1}{2N+1}\sum_{j=-N}^{+N} y_j^N(t) - \Theta\right| \leqq \delta(N), \quad \text{with} \quad \Theta = \lim_{N\to\infty}\frac{1}{2N+1}\sum_{j=-N}^{+N} y_j(0).$$

Then there exists a sequence $\varepsilon(n)>0$, $\varepsilon(n)\to 0$ as $n\to\infty$, such that:
Let $L(N)$ be a sequence with $L(N)\uparrow+\infty$, $L(N)\leqq K(N)$. Then (Y_t^N), (Y_t) can be constructed on one probability space such that

$$E(|y_j^N(L(N)) - y_j(L(N))|) \leqq \varepsilon(L(N)) + \delta(L(N)). \tag{2.61}$$

Finally we need later on the following statement, which generalizes Proposition 1.

Proposition 1' *For every subsequence N_k along which $\mathscr{L}((\Theta^{N_k}_{s\beta(N_k)})_{s\geqq 0})$ converges the following holds*

$$\mathscr{L}((Y^{N_k}_{s_1\beta(N_k)}, \ldots, Y^{N_k}(s_\ell\,\beta(N_k))) \Rightarrow \int P(\Theta_{s_1}\in d\rho_1, \ldots \Theta_{s_\ell}\in d\rho_\ell)\, \nu_{\rho_1}\otimes\ldots\otimes\nu_{\rho_\ell}.$$

The proof is as follows:
The system $Z^N(t) = (z_i^N(t))_{i\in[N,N]}$, which evolves in the time interval $[s_1\,\beta(N_k) - L(N_k), s_1\,\beta(N_k)], \ldots [s_\ell\,\beta(N_k) - L(N_k), \quad s_\ell\,\beta(N_k)]$ with $L(N) = o(\beta(N))$, $L(N) < \min_{j=1,\ldots,\ell}(s_{j+1}\,\beta(N) - s_j\,\beta(N))$, but $L(N)\to\infty$ as $N\to\infty$; as follows

$$\begin{aligned} dz_i^N(t) &= (\Theta^{j,N} - z_i^N(t))\,dt + \sqrt{2g(z_i^N(t))}\,dw_i(t)\\ Z^N(s_j\,\beta(N) - L(N)) &= Z(s_j) \quad \text{with } j=1,\ldots,\ell \end{aligned}$$

has by construction the property that for fixed path $(Z(t))$:

$$\mathscr{L}((Z^N(s_1\,\beta(N)), \ldots, Z^N(s_\ell\,\beta(N))) = \bigotimes_{k=1}^{\ell} \mathscr{L}(Z^N(s_k\,\beta(N)).$$

Then by Lemma 2.3 the assertion follows immediately.

(vi) Proof of Lemma 2.5 Since μ has the property $\int y_0^2\,d\mu<\infty$, and is translation invariant, we know that the $L_2(\mu)$-limit of $\frac{1}{2N+1}\sum_{i=-N}^{+N} y_i$ exists (L_2-ergodic theorem). Introduce the following notation:

$$D^N(Y) = \frac{1}{2N+1}\sum_{i=-N}^{+N} y_i, \qquad D(Y) = L_2 - \lim_{N\to\infty} D^N(Y).$$

Since (as mentioned above) $\mathscr{L}_\mu(D^N(Y)) \Rightarrow \mathscr{L}_\mu(D(Y))$ as $N\to\infty$, and by assumption $\mathscr{L}_{\mu_N}(D^M(Y)) \Rightarrow \mathscr{L}_\mu(D^M(Y))$ for fixed M as $N\to\infty$, it suffices to prove that:

$$\lim_{M\to\infty}\sup_{N\geqq M} \|D^M(Y) - D^N(Y)\|_{L_2(\mu_N)} = 0. \tag{2.62}$$

This is easily done using Fourier transforms. By standard results on the spectral representation of L_2 stationary sequences (see for instance [10, Chap. 3]) there are finite measures λ, λ_N on $[-\pi, \pi]$ which represent μ, μ_N via the Hilbert space correspondence

$$y_x \leftrightarrow e^{ixu}, \quad x \in \mathbf{Z}, \quad u \in [-\pi, \pi] \quad \text{and} \quad \int y_x y_y \, d\mu = \int_{[-\pi,\pi]} e^{i(x-y)u} \, d\lambda(u).$$

In particular, $\mu_N \Rightarrow \mu$ implies $\lambda_N \Rightarrow \lambda$. If we define

$$D^N(u) = \frac{1}{2N+1} \sum_{x=-N}^{+N} e^{ixu}, \qquad u \in [-\pi, \pi],$$

then

$$\| D^M(Y) - D^N(Y) \|_{L_2(\mu_N)} = \| D^M(u) - D^N(u) \|_{L_2(\lambda_N)}.$$

The trigonometric polynomials $D^N(u)$ satisfy: (i) $D^N(u) \to 1_{\{0\}}(u)$ as $N \to \infty$, $D^N(0) = 1$, (ii) for $\delta > 0$ and $M < \infty$ there exists $\varepsilon(M, \delta)$ such that for all $N \geqq M$

$$|D^N(u) - 1_{\{0\}}(u)| \leqq 1_{(-\delta,\delta)\setminus\{0\}}(u) + \varepsilon(M, \delta) \qquad \varepsilon(M, \delta) \to 0 \quad \text{as} \quad M \to \infty.$$

Consequently with letting K_δ abbreviate $(-\delta, \delta)\setminus\{0\}$ we have:

$$\| D^M(u) - D^N(u) \|_{L_2(\lambda_N)} \leqq 2\lambda_N(K_\delta) + 2\varepsilon(M, \delta). \tag{2.63}$$

Let first M tend to infinity to get

$$\overline{\lim_{M \to \infty}} \, (\sup_{N \geqq M} \| D^M(u) - D^N(u) \|_{L_2(\lambda_N)}) \leqq 2\lambda(K_\delta) \qquad \forall \delta > 0.$$

Since $K_\delta \searrow \emptyset$ as $\delta \to 0$, we get finally letting $\delta \to 0$ that

$$\overline{\lim_{M \to \infty}} \, (\sup_{N \geqq M} \| D^N(Y) - D^M(Y) \|_{L_2(\mu_N)}) = \lim_{\delta \to \infty} 2\lambda(K_\delta) = \lambda(\emptyset) = 0.$$

This finishes the proof of Lemma 2.5.

3 Convergence of the slowly varying variable $\Theta^N_{s\beta(N)}$ to Z_s

In this section we shall use the results of the previous one, to prove the statement (1.5) of Theorem 1.

Proposition 2

$$\mathscr{L}((\Theta^N_{s\beta(N)})_{s \in \mathbf{R}^+}) \underset{N \to \infty}{\Rightarrow} \mathscr{L}((Z_s)_{s \in \mathbf{R}^+}). \tag{3.0}$$

Proof. The proof consists of three parts corresponding to Lemmas 3.1 to 3.3 proving tightness, an equation for weak limit points and finally the uniqueness of the limit point.

We start with a lemma establishing some tightness results.

Lemma 3.1 (a) $(\Theta^N_{s\beta(N)})_{s\in\mathbf{R}^+}$ *is a square integrable martingale with continuous path and increasing process*

$$\langle\Theta^N_{s\beta(N)}\rangle=\int_0^s\left(\frac{1}{2N+1}\sum_{i=-N}^{+N}g(y_i^N(u\beta(N)))\right)du.$$

(b) *The sequence* $\{\mathscr{L}((\Theta^N_{s\beta(N)})_{s\in\mathbf{R}^+})\}_{N\in\mathbf{N}}$ *is tight as a sequence of probability measures on* $\mathscr{C}([0,\infty])$.

Proof. Since the $y_i^N(t)$ are solutions of the system of the stochastic differential Eq. (1.1) it follows for N fixed, that Θ_t^N is given by:

$$\Theta_t^N=\Theta_0^N+\frac{1}{2N+1}\sum_{i=-N}^{N}\int_0^t\sqrt{2g(y_i^N(s))}\,dw_i(s). \tag{3.1}$$

Since stochastic integrals with respect to Brownian motion are local martingales, we know already that $(\Theta_t^N)_{t\in\mathbf{R}^+}$ is a local martingale so that since g is bounded $(\Theta_t^N)_{t\in\mathbf{R}^+}$ is a square integrable martingale. As a stochastic integral with respect to Brownian motion it has automatically continuous path.

By Ito's formula we derive from (3.1):

$$E(\Theta_t^N)^2=E(\Theta_0^N)^2+\frac{1}{(2N+1)^2}\sum_{i=-N}^{N}\int_0^t Eg(y_i^N(s))\,ds\leqq\frac{1}{2N+1}\|g\|_\infty\cdot t+1. \tag{3.2}$$

We see from (3.2) that for $L(N)=o(N)$ and every $\beta(N)$

$$\sup_{t\leqq L(N)}|\Theta^N_{s\beta(N)+t}-\Theta^N_{s\beta(N)}|\underset{N\to\infty}{\Rightarrow}0 \tag{3.2'}$$

in probability by the optional sampling theorem.

We have to establish the tightness of $\{(\Theta^N_{s\beta(N)})_{s\in\mathbf{R}^+}\}$. The representation (3.1) together with the fact that as processes $\mathscr{L}(\sqrt{\alpha}\,w_i(t))=\mathscr{L}(w_i(\alpha t))$ implies that $\Theta^N_{s\beta(N)}$ can be represented equivalently in law as

$$\Theta^N_{s\beta(N)}=\Theta_0^N+\frac{1}{\sqrt{2N+1}}\sum_{i=-N}^{+N}\int_0^s\sqrt{2g(y_i^N(u\beta(N)))}\,dw_i(u) \tag{3.3}$$

and then the increasing process is given by

$$\langle\Theta^N_{s\beta(N)}\rangle=\int_0^s\left(\frac{1}{2N+1}\sum_{i=-N}^{+N}g(y_i^N(u\beta(N)))\right)du \tag{3.4}$$

so that

$$0\leqq\frac{d\langle\Theta^N_{s\beta(N)}\rangle}{ds}\leqq\|g\|_\infty,\qquad\forall s\in\mathbf{R}^+. \tag{3.5}$$

This implies in particular that the path of $(\langle \Theta^N_{s\beta(N)}\rangle)_{s\geqq 0}$ are equi-continuous on compact time intervals. Therefore $(\langle \Theta^N_{s\beta(N)}\rangle)_{s\leqq T}$ form a relatively compact subset of $\mathscr{C}([0, T])$. As a consequence the sequence

$$\mathscr{L}((\langle \Theta^N_{s\beta(N)}\rangle)_{s\geqq 0})$$

is a relatively weakly compact sequence of probability measures on $\mathscr{C}([0, \infty))$. Observe furthermore that $\langle \Theta^N_{s\beta(N)}\rangle$ is strictly increasing up to the point in time were it becomes a constant (that is the underlying process has hit the traps $\mathbf{0}$ or $\mathbf{1}$). The martingale $(\Theta^N_{s\beta(N)})_{s\geqq 0}$ is a stochastic integral, and therefore a time changed Brownian motion (see Ethier and Kurtz [18, Chap. 6]):

$$\Theta^N_{s\beta(N)} = W(S^N(s)), \quad S^N(T) := \langle \Theta^N_{T\beta(N)}\rangle.$$

It is therefore clear from the path continuity of $W(\cdot)$ that the tightness of the $\mathscr{L}((\langle \Theta^N_{s\beta(N)}\rangle)_{s\in\mathbf{R}^+})$ implies the tightness of the sequence $\{(\Theta^N_{s\beta(N)})_{s\in\mathbf{R}^+}\}$ of martingales in the path space $\mathscr{C}([0, \infty])$. This finishes the proof of Lemma 3.1. □

In the next step we characterize weak limit points of $\mathscr{L}((\Theta^N_{s\beta(N)})_{s\in\mathbf{R}^+})$:

Lemma 3.2 *Let N_k be a subsequence such that:*

$$\mathscr{L}((\Theta^{N_k}_{s\beta(N_k)})_{s\in\mathbf{R}^+}) \Rightarrow \mathscr{L}((\Theta_s)_{s\in\mathbf{R}^+}). \tag{3.6}$$

Then $(\Theta_s)_{s\in\mathbf{R}^+}$ is a square integrable martingale with continuous path and

$$\left(\Theta_s^2 - \int_0^s \Phi(\Theta_u)\, du\right)_{s\geqq 0} \tag{3.7}$$

is a martingale, where Φ is given by

$$\Phi(\Theta) = E^{\nu_\Theta} g(y_0) = u_g(\Theta).$$

Proof. Since our state space is $[0,1]$ we know that Θ_s is square integrable. The problem is to show that Θ_s is a martingale and the increasing process of Θ_s is given by the integral in (3.7). The fact that Θ_s is a martingale will follow as a side result from an argument we give in the sequel of (3.10) below to prove (3.7), so it suffices to show that Θ_s satisfies (3.7).

Step 1 We start by finding the limit of $\langle \Theta^N_{s\beta(N)}\rangle$. For that purpose we apply our Proposition 1, to conclude from our assumption, namely,

$$\mathscr{L}((\Theta^{N_k}_{s\beta(N_k)})_{s\in\mathbf{R}^+}) \Rightarrow \mathscr{L}((\Theta_s)_{s\in\mathbf{R}^+})$$

and from the relative compactness of $\{\mathscr{L}(Y^{N_k}(u\beta(N_k)))\}_{k\in\mathbf{N}}$ that:

$$\mathscr{L}(Y^{N_k}_{u\beta(N_k)}) \underset{k\to\infty}{\Rightarrow} \mathscr{L}((Y^\infty(u))) \quad \forall u\in\mathbf{R}^+, \quad \text{as} \quad k\to\infty \tag{3.8}$$

$$\mathscr{L}(Y_u) = \int P_u(d\Theta)\, \nu_\Theta, \quad P_u(\cdot) = \mathscr{L}(\Theta_u).$$

This implies that in the sense of weak convergence of distributions we have for all u:

$$\left(\frac{1}{2N_k+1}\sum_{i=-N_k}^{N_k} g(y_i^{N_k}(u\beta(N_k)))\right)\underset{k\to\infty}{\Rightarrow}(\Phi(\Theta_u)).$$

To verify this apply Lemma 2.5 to $\{g(y_i^N)\}_{i=-N,\ldots,N}$ and use the fact that $(y_i^\infty(u))_{i\in\mathbf{Z}}$ is i.i.d. (see (2.2) and (1.4)) to conclude $\frac{1}{2N+1}\sum_{i=-N}^{+N} g(y_i^\infty(u))$ $\to E(g(y_0^\infty(u))$ as $N\to\infty$.

This argument will be extended to give for all $u_1,\ldots,u_m$

$$\left(\frac{1}{2N_{k+1}}\sum_{i=-N_k}^{N_k} g(y_i^{N_k}(u_\ell\,\beta(N_k)))\right)_{\ell=1,\ldots,m}\Rightarrow(\Phi(\theta_{u_\ell}))_{\ell=1,\ldots,n}$$

in the sense of weak convergence. For this purpose we need the following fact, proved in Proposition 1′.

$$\mathscr{L}((Y^{N_k}_{u_1\beta(N_k)}),\ldots,Y^{N_k}_{u_\ell\beta(N_k)})\underset{k\to\infty}{\Rightarrow}\int P(\Theta_{u_1}\in d\rho_1,\ldots,\Theta_{u_\ell}\in d\rho_\ell)\,v_{\rho_1}\otimes\ldots\otimes v_{\rho_\ell}$$

which immediately gives the desired convergence result above using the same line of argument as used above in the case $\ell=1$.

Step 2 Observe that the integral

$$\int_s^t\left(\frac{1}{2N+1}\sum_{k=1}^N g(y_k^N(s\beta(N))\right)ds$$

can also be written as

$$\frac{1}{\beta(N)}\int_{s\beta(N)}^{t\beta(N)}\left(\frac{1}{2N+1}\sum_{k=1}^N g(y_k^N(s))\right)ds$$

or

$$\sum_i\frac{1}{\beta(N)}\int_{t_i\beta(N)}^{t_{i+1}\beta(N)}\left(\frac{1}{2N+1}\sum_{k=1}^N g(y_k^N(s))\right)ds$$

for a partition $s<t_1<\ldots<t_m=t$, $|t_{i+1}-t_1|\leq\varepsilon$. Hence we have as $N\to\infty$ averaging in time and space taking place. We want to show that

$$\overline{\lim_{N\to\infty}}\,E\left(\sup_{\{t_i\}}\left|\sum_i|t_{i+1}-t_i|\,\Phi(\Theta_{t_i})-\sum_i\frac{1}{\beta(N)}\int_{t_i\beta(N)}^{t_{i+1}\beta(N)}\left(\frac{1}{2N+1}\sum_{k=1}^N g(y_k^N(s))\right)ds\right|\right)\leq\delta(\varepsilon)$$

with $\delta(\varepsilon)\to 0$ as $\varepsilon\to 0$.

For that purpose we compare for a given partition (t_i) the system in the time intervals $[t_i\,\beta(N),\ t_{i+1}\,\beta(N)]$ with a new system $(Z_i(t))$ defined with the same brownian motions used for $(y_k^N)_{k=-N,\ldots,N}$ as follows:

$$dz_i(t)=(\Theta^{k,N}-z_i(t))\,dt+\sqrt{2g(z_i(t))}\,dw_i(t),\quad i\in[-N,+N]$$
$$\theta^{k,N}=\Theta_{t_k\beta(N)}.$$

The dependence on the partition is suppressed in the notation. For such a system we would have by the independence of the components and the continuous time ergodic theorem:

$$\frac{1}{\beta(N)} \int\limits_{t_k\beta(N)}^{t_{k+1}\beta(N)} \left(\frac{1}{2N+1} \sum_{i=-N}^{N} g(z_i(s))\right) ds \underset{N\to\infty}{\longrightarrow} \langle \Gamma_{\Theta_{t_k}}, g\rangle (t_{k+1}-t_k) = \Phi(\Theta_{t_k})(t_{k+1}-t_k).$$

The sum over the righthand side forms an approximation of the integral $\int\limits_0^t \Phi(\Theta_u)\, du$. Hence as the partition becomes finer

$$\left(\int\limits_0^t \left(\frac{1}{N} \sum_{i=1}^{N} g(z_i(s\beta(N)))\right) ds\right)_{t\geq 0} \to \left(\int\limits_0^t \Phi(\Theta_s)\, ds\right)_{t\geq 0}.$$

It remains to compare the $\{y_i(t)\}_{i\in\{1,\dots,N\}}$ and the $\{z_i(t)\}_{i\in\{1,\dots,N\}}$, in the limit of finer and finer partitions.

For this purpose we resort to our coupling results in Sect. 2, (2.21)–(2.38) which immediately gives the desired estimate (3.9), since we can apply to $\{\bar{z}_i(t)\}_{i\in\mathbf{Z}}$, $\{\bar{y}_i^N(t)\}_{i\in\mathbf{Z}}$ the bound (2.38), where the – indicates the periodic continuation from a process indexed with $\{-N, N\}$ to one indexed by $\mathbf{Z}$.

We have shown now that for any collection $0\leq t_1 < t_2 < \dots < t_m$:

$$\text{(3.9)} \qquad \left(\int\limits_0^{t_k} \left[\frac{1}{2N+1} \sum_{i=-N}^{+N} g(y_i^N(u\beta(N)))\right] du\right)_{k=1,\dots,m} \underset{N\to\infty}{\Rightarrow} \left(\int\limits_0^{t_k} \Phi(\Theta_u)\, du\right)_{k=1,\dots,m}, \quad \text{in law.}$$

The paths of the processes are increasing with (in N) uniformly bounded derivative, i.e. the values lie in an equi-continuous subset of paths in $C([0,\infty])$. Hence the $\mathscr{L}((\Theta^N_{s\beta(N)})_{s\geq 0})$ are tight. This means that

$$\text{(3.10)} \qquad \mathscr{L}(\langle \Theta^{N_k}_{s\beta(N_k)}\rangle_{s\in\mathbf{R}^+}) \underset{k\to\infty}{\Rightarrow} \mathscr{L}\left(\left(\int\limits_0^s \Phi(\Theta_u)\, du\right)_{s\in\mathbf{R}^+}\right).$$

This statement is now strengthened to

$$\text{(3.10')} \qquad \mathscr{L}\left(\left(\int\limits_0^t \frac{1}{2N_k+1} \sum_{i=-N_k}^{N_k} g(y^{N_k}(u\beta(N_k))\, du, \Theta^{N_k}_{t\beta(N_k)}\right)_{t\geq 0}\right) \underset{k\to\infty}{\Rightarrow} \mathscr{L}\left(\left(\int\limits_0^t u_g(\Theta_s)\, ds, \Theta_t\right)_{t\geq 0}\right)$$

which is easy to see using the method of moments together with the previous results.

Step 3 Next we show that the r.h.s. of (3.10) gives the increasing process of Θ_s and on the way we get that Θ_s is a martingale. Observe that by the Doob-Meyer decomposition [5]:

(3.11) $$(\Theta^{N_k}_{s\beta(N_k)})^2 = M^{N_k}_s + \langle \Theta^{N_k}_{s\beta(N_k)} \rangle, \quad \text{where } M^{N_k}_s \text{ is a martingale.}$$

What we need next is the following fact:
Suppose $\{(M^N_t)_{t\in\mathbf{R}^+}\}_{N=1,\ldots}$ is a sequence of continuous martingales with respect to the σ-algebras $\{\mathscr{F}_t\}$, and all these martingales are defined on *one* probability space. If this sequence has the following additional properties:

$$\sup_N E(M^N_t)^2 \leqq Ch(t), \qquad C\in(0,\infty), \;\; h(t)<\infty \;\; \forall t\geqq 0$$
$$\mathscr{L}((M^N_t)_{t\in\mathbf{R}^+}) \underset{N\to\infty}{\Rightarrow} \mathscr{L}((M_t)_{t\in\mathbf{R}^+})$$

then

$$(M_t)_{t\in\mathbf{R}^+} \text{ is a square integrable martingale.}$$

Proof. Let $g=1_A$ with $A\in\mathscr{F}_t$. Then by assumption

$$E((M^N_{t+s}-M^N_t)\,g)=0 \qquad \forall s>0$$
$$E((M^N_{t+s}-M^N_t)\,g) \underset{N\to\infty}{\Rightarrow} E((M_{t+s}-M_t)\,g),$$

wher $g\in\mathscr{A}_t$, g is a continuous function of the path and

$$\mathscr{A}_t = \{\text{events depending on the path up to time } t\}$$

and therefore

$$E((M_{t+s}-M_t)\,g)=0 \qquad \forall g, \text{ which are } \mathscr{A}_t-\text{measurable.} \qquad \square$$

In order to apply this statement to our situation observe that

$$\mathscr{L}(\sqrt{\alpha}\, w_i(t)) = \mathscr{L}(w_i(\alpha t))$$

so that we can define $\{(\tilde{\Theta}^N_s)_{s\in\mathbf{R}^+}, N\in\mathbf{N}\}$, $\{\{(\tilde{y}^N_i(t))_{t\in\mathbf{R}^+}, i\in\mathbf{Z}\}, N\in\mathbf{N}\}$ on one probability space as strong solutions of:

$$\tilde{\Theta}^N_s = \tilde{\Theta}^N_0 + \int_0^s \frac{1}{\sqrt{2N+1}} \sum_{i=-N}^{+N} \sqrt{2g(y^N_i(u\beta(N)))}\, dw_i(u)$$
$$d\tilde{y}^N_i(t) = \left(\sum_{-N}^{+N} \tilde{y}^N_j(t) - \tilde{y}^N_i(t)\right) dt + \sqrt{2(2N+1)\, g(\tilde{y}^N_i(t))}\, dw_i(t).$$

Combining these two relations we can write

$$\tilde{\Theta}^N_t = \tilde{\Theta}^N_0 + \int_0^t \frac{1}{\sqrt{2N+1}} \sum_{-N}^{+N} \sqrt{2g(\tilde{y}^N_i(u))}\, dw_i(u)$$

and we see that

$$\mathscr{L}((\tilde{\Theta}_t^N)_{t\in\mathbf{R}^+})=\mathscr{L}((\Theta_{t\beta(N)}^N)_{t\in\mathbf{R}^+}),\qquad \forall N.$$

Therefore by defining either

$$M_t^N=\tilde{\Theta}_t^N,$$
$$\text{or}\quad M_t^N=(\tilde{\Theta}_t^N)^2-\langle\tilde{\Theta}_t^N\rangle,$$

with respect to the filtration

$$\mathscr{F}_t=\mathscr{A}\{\{(w_i(s))_{s\leqq t}\}_{i\in\mathbf{Z}}\}$$

we can use in both cases the observations made above, following (3.11).

From the assumption and (3.10) and (3.11) we conclude now from the above discussion:

(3.12) $$(\Theta_s)_{s\geqq 0}\quad\text{is a martingale}$$
$$\mathscr{L}((M_s^{N_k})_{s\in\mathbf{R}^+})\underset{k\to\infty}{\Rightarrow}\mathscr{L}((M_s)_{s\in\mathbf{R}^+}),\quad\text{and } M_s \text{ is a martingale.}$$

Step 4 Therefore by combining (3.10′) with (3.12) we obtain

(3.13) $$\Theta_s^2=M_s+\int_0^s\Phi(\Theta_u)\,du,\quad (M_s)_{s\geqq 0}\text{ is a martingale.}$$

The second term on the r.h.s. is increasing and $\mathscr{A}((\Theta_u)_{u\leqq s})$-measurable, this implies that

(3.14) $$\langle\Theta_t\rangle=\int_0^t\Phi(\Theta_u)\,du,$$

since the increasing process is uniquely determined (see Ethier and Kurtz [18, p. 74, Theorem 5.1]). According to Lemma 2.2 the map $\Phi(\cdot)$ is continuous, therefore (3.14) implies that Θ_s has continuous path. This proves (3.7) and the proof of Lemma 3.2. is finished.

The last piece of information needed, to complete the proof of Proposition 2, is that there is only one process which satisfies (3.7).

Lemma 3.3 *If we denote by $\Phi(u)=E^{\nu_\Theta}(g(y_0))$, then the following martingale problem has a unique solution:*

(3.15) $$(\Theta_s)_{s\geqq 0}\quad\textit{is a continuous martingale with values in }[0,1]$$
$$\left(\Theta_s^2-\int_0^s\Phi(\Theta_u)\,du\right)_{s\geqq 0}\quad\text{is a martingale.}$$

(3.16) *The solution of* (3.15) *is given by the diffusion generated by* $\Phi(x)\left(\frac{\partial}{\partial x}\right)^2$.

Proof. By Lemma 2.2 the map $\Phi(\cdot)$ is Lipschitz. Therefore we obtain the assertion (3.15) from Stroock and Varadhan [11, p. 152]. Finally (3.16) is a standard fact in diffusion theory.

Proof of Proposition 2 By Lemma 3.1 we can choose a subsequence N_k such that $\mathscr{L}((\Theta^{N_k}_{s\beta(N_k)})_{s\in\mathbf{R}^+})$ converges. Lemma 3.2 combined with Lemma 3.3 then implies that the weak limit point is $\mathscr{L}((Z_s)_{s\geqq 0})$ independent of the subsequence so that we have in fact convergence. □

4 Proof of Theorems 1–4

(a) Proof of Theorem 1, Corollary 1 and Theorem 2

Proof of Theorem 1 I. (1.5) was proved in Proposition 2.

II. Proof of (1.6). The combination of Propositions 1 and 2 and (3.2′) gives us ($M_1([0,1]^{\mathbf{Z}})$ is compact in the weak topology)

$$\mathscr{L}(Y^N_{s\beta(N)+u}) \underset{N\to\infty}{\Rightarrow} \int Q(\Theta', d\Theta)\, v_\Theta = v(s) \qquad \forall u\in\mathbf{R}^+. \tag{4.0}$$

From the coupling techniques used to prove Lemmas 2.3, 2.4, namely, the relation (2.60) we derive

$$\mathscr{L}((Y^N_{s\beta(N)+u_1}, Y^N_{s\beta(N)+u_1+u_2})) \underset{N\to\infty}{\Rightarrow} \mathscr{L}((Y^{v(s)}_{u_1}, Y^{v(s)}_{u_1+u_2})). \tag{4.1}$$

Since these considerations apply as well to time points $s\beta(N)+u_1, \ldots, s\beta(N)+u_k$, we have established that the finite dimensional distributions of the process $\mathscr{L}((Y^N_{s\beta(N)+t})_{t\in\mathbf{R}^+})$ converge to those of the process $\mathscr{L}((Y^{v(s)}_u)_{u\in\mathbf{R}^+})$.

It remains therefore to establish the tightness of the distributions $\mathscr{L}((Y^N_{s\beta(N)+t})_{t\in\mathbf{R}^+})$ as measures on the space $D([0,\infty), [0,1]^{\mathbf{Z}})$. To see this we use again the coupling result (2.60). Tightness on $D([0,\infty), [0,1]^{\mathbf{Z}})$ is established verifying tightness of the components on this product space. First note that a criterion for tightness of a sequence of processes in the space of the components is the following (see Ethier and Kurtz [18, p. 128, Theorem 7.2]): Let X_n be a sequence of random variables with values in $D[(0,\infty), [0,1])$. The sequence is tight if for all $\eta>0$, $T>0$ exists a $\delta>0$ such that:

$$\sup_n P(w'(X_n, \delta, T) \geqq \eta) \leqq \eta$$

$$w'(x, \delta, T) = \inf_{\{t_i\}} \max_i \sup_{s,t\in[t_{i-1},t_i]} (|X(s)-X(t)|)$$

$$t_{i+1}-t_i \geqq \delta\ \forall i, \quad 0<t_1<\ldots\leqq T\leqq t_n.$$

Next observe that by (4.0) we know that we can apply (2.60) with $N_k=k$. Therefore tightness follows from our coupling in (2.60) between $\{y^N_i(t)\}_{i\in[-N,N]}$ and $\{y_i(t)\}_{i\in\mathbf{Z}} = Y^{v(s)}(t)$ (as before – refers to be bivariate dynamics): Since $|\bar{y}^N_i(t)-\bar{y}_i(t)|$

is a semimartingale, using (2.60) and the optional sampling theorem we can obtain that:

(4.2) $$\mathrm{Prob}(|\bar{y}_i^N(t)-\bar{y}_i(t)|\leqq \varepsilon, \quad \forall t\in[s\beta(N), s\beta(N)+T]) \underset{N\to\infty}{\Rightarrow} 1.$$

We can therefore estimate for all v, $u+v\in[s\beta(N), s\beta(N)+T]$:

(4.3) $$|\bar{y}_i^N(u+v)-\bar{y}_i^N(u)|\leqq|\bar{y}_i^N(u+v)-\bar{y}_i(u+v)|+|\bar{y}_i^N(u)-\bar{y}_i(u)|+|\bar{y}_i(u+v)-\bar{y}_i(u)|.$$

(4.2) and (4.3) then implies tightness using the criterion quoted above.

III. Proof of (1.7). Denote by $S^N(t)$ the semigroup of the process defined by the system of stochastic differential Eqs. (1.1) and let ν_Θ be the invariant measure of the infinite system (1.2). Then the relation (2.5) together with (1.7) and (2.3) in connection with (2.8) applied to our situation reads as follows:

Let $\{\mu_\Theta^N\}$ $N=1,2,\dots$ be a sequence of translation invariant probability measures on $[0,1]^{\mathbf{Z}}$ satisfying

$$\mu_\Theta^N\left(\left\{\frac{1}{2N+1}\sum_{i=-N}^{N} y_i=\Theta\right\}\right)=1.$$

Then we have

(4.4) $$(\mu_\Theta^N)\, S^N(L(N))\Rightarrow \nu_\Theta.$$

Therefore we write with $\delta_N=L(N)/N$

$$\begin{aligned}(\mu_\Theta^N)\, S^N((t-s)\,\beta(N))&=(\mu_\Theta^N)\, S^N(L(N)+(t-s-\delta_N)\,\beta(N))\\ &=[(\mu_\Theta^N)\, S^N(L(N))]\, S^N((t-s-\delta_N)\,\beta(N)).\end{aligned}$$

By Lemma 2.4 and (4.4) we can replace the right side above for $N\to\infty$ by

$$\nu_\Theta\, S^N((t-s-\delta_N)\,\beta(N))$$

and by (1.6) this will be for $N\to\infty$, approximated by

$$\nu_\Theta\, S^N((t-s)\,\beta(N)) \underset{N\to\infty}{\Rightarrow} \int Q_{t-s}(\Theta, d\tilde{\Theta})\, \nu_{\tilde{\Theta}}, \qquad \text{for all } \Theta\in[0,1].$$

Now simply consider a sequence μ_Θ^N obtained by taking for each N a translation invariant version of

$$\mathscr{L}(Y_{s\beta(N)}^N\,|\,\Theta_{s\beta(N)}^N=\Theta).$$

Then we know from the argument above that with this choice of a version of the conditional probabilities:

(4.5) $$\mathscr{L}(Y_{t\beta(N)}^N\,|\,\Theta_{s\beta(N)}^N=\Theta) \underset{N\to\infty}{\Rightarrow} \int Q_{t-s}(\Theta, d\tilde{\Theta})\, \nu_{\tilde{\Theta}} \qquad \text{for all } \Theta\in[0,1].$$

Then we have, with P_s^N denoting the law of $\theta^N_{s\beta(N)}$

$$\int_{\theta-\varepsilon_N}^{\theta+\varepsilon_N} \mathscr{L}(Y^N_{t\beta(N)}|\theta^N_{s\beta(N)}=\theta)\, P_s^N(d\theta)=\mathscr{L}(Y^N_{t\beta(N)}, |\theta^N_{s\beta(N)}-\theta|\leqq\varepsilon_N).$$

Note that $\Theta\to\nu_\Theta$ is a continuous map $[0,1]\to\mathscr{M}_1([0,1]^{\mathbf{Z}})$ and $Q_{t-s}(\cdot, d\Theta)$ is a continuous map $\mathbf{R}\to M_1([0,1]^{\mathbf{Z}})$. It is then straightforward to derive from (4.5):

$$\mathscr{L}(Y^N_{t\beta(N)}|\Theta^N_{s\beta(N)}=\tilde{\Theta}\pm\varepsilon_N)\underset{N\to\infty}{\Rightarrow}\int Q_{t-s}(\tilde{\Theta}, d\Theta)\,\nu_\Theta.$$

Proof of Corollary 1

The proof of relation (1.11) consists of two steps, first a soft argument using the invariance principle established in Theorem 1 and then second a comparison argument using special properties of the process (Θ_s^N) to complete the argument. It is for this second step we need the additional regularity assumptions quoted in the assumptions of Corollary 1.

Step 1 First note that T^N, T can be characterized as

(4.6) $$\inf\{t\,|\,Y^N_{t+h}=0 \text{ or } Y^N_{t+h}=1 \text{ for all } h\geqq 0 \text{ respectively}\}.$$

This means that T^N, T are lower semi-continuous functions of the path. Therefore the result (1.6) of Theorem 1 implies ("$\geqq$" means stochastically larger here):

(4.7) $$w-\lim_{k\to\infty}\mathscr{L}(T^{N_k}/\beta(N_k))\geqq\mathscr{L}(T)$$

and here N_k is a subsequence such that $\mathscr{L}(T^{N_k}/\beta(N_k))$ converges in the weak topology of measures on $\mathbf{R}\cup\{\infty\}$. We denote such a limit point by $\mathscr{L}(T^\infty)$.

By Theorem 1 and Skorohods' Theorem $(Y_t^N)_{t\in\mathbf{R}^+}$, $(Y_t)_{t\in\mathbf{R}^+}$, T^∞ can be defined on one probability space (cf. [2]) such that

(4.8) $$(\Theta^N_{s\beta(N)})\underset{\|\cdot\|_\infty}{\Rightarrow}\Theta_s,\ T^\infty\geqq T \quad \text{a.s.} \quad \text{as} \quad N\to\infty.$$

It is therefore our task to prove that:

$$\text{Prob}(T-T^\infty<0)=0.$$

Step 2 Here we have to distinguish the cases of accessible respectively inaccessible boundaries.

Case 1 $\{0,1\}$ are accessible boundary points for $g(x)\left(\dfrac{\partial}{\partial x}\right)^2$.

Consider an enlarged probability space in which (4.8) holds. To get an estimate on $\text{Prob}(T-T^\infty\leqq 2\delta)$ for $\delta<0$ we observe that on the event $\{T-T^\infty\leqq\delta\}$ the following relations hold (compare (4.8) and the fact that $\{0,1\}$ is accessible for Θ_s^N):

For all $\varepsilon>0$ there is a $N_0(\varepsilon,\omega)$ such that on this event

$$(4.9)\qquad \begin{aligned} &0<\Theta_s^N\leqq\varepsilon \quad \forall s\in[T^N-\delta\beta(N),T^N] \quad \text{for } N\geqq N_0(\varepsilon,\omega)\\ \text{or}\quad &1-\varepsilon\leqq\Theta_s^N<1 \quad \forall s\in[T^N-\delta\beta(N),T^N] \quad \text{for } N\geqq N_0(\varepsilon,\omega). \end{aligned}$$

But the fact that Θ_s^N is a solution to (3.1), implies that Θ_s^N is a time-transformed Brownian motion: $w(\tilde{S}_t^{x,N})$. In order to exploit these facts, we introduce now a convex function $\hat{g}$ which is an "effective" minorant of g at one boundary, that is for some $\alpha\in[1,2)$, $c\in\mathbf{R}^+$,

$$\hat{g}(x)=cx^\alpha,$$

so that, using the regularity assumptions of g (g behaves like some power close to 0 and 1), we have for some prescribed $\beta>0$

$$\hat{g}(x)\leqq g(x) \qquad \forall x\in[0,1-\beta].$$

We shall see that this implies that on the event $\{\Theta_{s\beta(N)}\in(0,\varepsilon] \text{ for } s\in[T,T+\delta]\}$ we have for some $b>0$

$$(4.10)\qquad \begin{aligned} &\int_T^{T+\delta}\frac{1}{2N+1}\sum_{i=-N}^{N} g(y_i^N(s\beta(N))\,ds\\ &\quad\geqq b\int_T^{T+\delta}\frac{1}{2N+1}\sum_{i=-N}^{N}\hat{g}(y_i^N(s\beta(N)))\,ds \qquad \forall N\geqq N_0. \end{aligned}$$

In order to see that we can achieve this for some N_0, $b>0$, we first use:

$$\int_T^{T+\delta}\frac{1}{2N+1}\sum_{i=-N}^{N} g(y_i^N(s\beta(N)))\,ds=\frac{1}{\beta(N)}\int_{T\beta(N)}^{(T+\delta)\beta(N)}\frac{1}{2N+1}\sum_{i=-N}^{N} g(y_i^N(s))\,ds$$

in connection with Theorem 1 and (3.8)–(3.9) in the proof of Theorem 1. Then in order to estimate the contribution where $g(x)>\hat{g}(x)$ we use the facts that,

$$\begin{gathered}[\sup_{\Theta\leqq\varepsilon}(\Gamma_\Theta(y(t)\in[1-\beta,1]))/\sup_{\theta\leqq\varepsilon}(\Gamma_\theta(y(t)\in[0,1-\beta]))]\leqq a(\varepsilon)\underset{\varepsilon\to 0}{\Rightarrow}0,\\ \Gamma_\theta([1-\beta,1-\beta/2]\mid y\in[1-\beta,1])\geqq 1/2.\end{gathered}$$

The latter relation can be read off from the explicit form of Γ_Θ (see Sect. 1, (ii)).

Since $\hat{g}(x)$ is convex we can conclude by Jensen's inequality:

$$\frac{1}{2N+1}\sum_{i=-N}^{N}\hat{g}(x_i^N(s))\geqq\hat{g}(\Theta_s^N).$$

If we denote now by $\tilde{S}^x$ the random time transformation such that

$$\hat{\Theta}_s^x=w(\tilde{S}^x(t)) \qquad \text{is a diffusion generated by } \hat{g}(y)\left(\frac{\partial}{\partial y}\right)^2$$

starting at x, we conclude from (4.10) via the well-known relation between the time transformation $S(t)$ and the increasing process of x_t, namely $S_t = \langle x_t \rangle$, that we can construct Θ_s^N and $\hat{\Theta}_s^x$ on one probability space such that (simply use the same Brownian path)

$$\tilde{S}^{x,N}(\cdot) \geqq \tilde{S}^x(\cdot), \tag{4.11}$$

holds on the event $\{\Theta_{s\beta(N)} \in (0, \varepsilon], \forall s \in [T, T+\delta]\}$.

Since $(\hat{\Theta}_s^x)_{s \geqq 0}$ is strong Markov, has 0 and 1 as accessible boundaries and since $(\hat{\Theta}_s^x)_{s \geqq 0}$ is a martingale with continuous path we know that for $\delta > 0$:

$$\operatorname{Prob}(\inf\{s \mid \hat{\Theta}_s^x = 0\}, \delta) \uparrow 1 \quad \text{as} \quad x \downarrow 0.$$

Consequently for $0 < x < y \leqq \varepsilon$

$$\operatorname{Prob}(\hat{\Theta}_s^x \in [0, y) \quad \forall s \leqq \delta) \underset{\varepsilon \to 0}{\Rightarrow} 0$$

for every $\delta > 0$. Due to the representation (4.11) we then know that also

$$\operatorname{Prob}(\Theta_s^N \in [0, y) \quad \forall s \leqq \delta \mid \Theta_0 = x) \underset{\varepsilon \to 0}{\Rightarrow} 0.$$

From the last fact we conclude with (4.9) that for all $\delta < 0$

$$\operatorname{Prob}(T - T^\infty \leqq \delta) = 0,$$

which finishes the proof.

Case 2 $\{0, 1\}$ are inaccessible boundaries for $g(x)\left(\frac{\partial}{\partial x}\right)^2$.

Since Θ_s is a martingale we know that $\Theta_s \underset{s \to \infty}{\longrightarrow} \Theta_\infty$, where Θ_∞ takes the value 0 or 1. Since in (0, 1) the diffusion coefficient of Θ_s is strictly positive, we shall prove below that if 0, 1 are inaccessible for $g(x)\left(\frac{\partial}{\partial x}\right)^2$ then they are inaccesible for $u_g(x)\left(\frac{\partial}{\partial x}\right)^2$. Therefore we have the following

$$\left\{\begin{array}{ccc} \Theta_s & \underset{s \to \infty}{\longrightarrow} & 0 \text{ a.s.} \\ & \text{or} & \\ \Theta_s & \underset{s \to \infty}{\longrightarrow} & 1 \text{ a.s.} \end{array}\right\} \quad \text{and} \quad 0 < \Theta_s < 1 \quad \forall s \in \mathbf{R}^+. \tag{4.12}$$

Then it is clear that the result of Theorem 1, that

$$\Theta_{s\beta(N)}^N \underset{\varepsilon \to 0}{\longrightarrow} \Theta_s$$

implies now

$$T^N/N \to \infty \quad \text{a.s.} \tag{4.13}$$

and therefore $T_\infty = \infty$ a.s., that is $T_\infty = T$ a.s.

It remains to show that under our assumptions

$$(4.14)\qquad 0 \text{ is inaccesible for } g(x)\left(\frac{\partial}{\partial x}\right)^2$$
$$\Rightarrow 0 \text{ is inaccessible for } u_g(x)\left(\frac{\partial}{\partial x}\right)^2.$$

Using the criterion for inaccessibility $\int_0^\varepsilon \frac{x}{g(x)}\,dx=\infty$ (cf. Breiman Prob. 16.43), (4.14) follows from the following useful relation.

Lemma 4.1

$$(4.15)\quad g(x) \underset{x\to 0}{\sim} c x^\beta \quad \textit{for}\ \beta\geqq 2 \Rightarrow u_g(x) \begin{cases} \underset{x\to 0}{\lesssim} c' x^\beta & \beta>2 \\ \underset{x\to 0}{\lesssim} c' x^2|\log x|, c'>0 & \textit{for}\ \beta=2\, c\leqq 1. \end{cases}$$

Proof. The result is proved by using the formula $u_g(x)=\int g(y)\,\Gamma_x(dy)$ with $\Gamma_\Theta(dy) = c_\Theta^{-1}\frac{1}{g(y)}\exp\left(\int_\theta^y \frac{\theta-z}{g(z)}\,dz\right)$. We start with the case $\beta>2$.

1. The quantity C_Θ for $\Theta\to 0$ $\beta>2$.
We start by considering the following special case for g:

$$g(x)=x^k \qquad \text{for } k>2.$$

Assuming the special form of g as above, the quantity C_Θ can be calculated to be (for $k>2$)

$$\int_0^1 \frac{1}{y^k}\exp\left(-\left(\frac{\Theta}{k-1}-\frac{y}{k-2}\right)\Big/ y^{k-1}\right)dy=\int_0^\Theta+\int_\Theta^1.$$

We treat each integral separately.

The second integral is estimated using the substitution $z=(k-1)^{-1}y^{-(k-2)}$ as follows

$$\int_\Theta^1 \frac{1}{y^k}\exp\left(-\left(\frac{\Theta}{k-1}-\frac{y}{k-2}\right)\Big/ y^{k-1}\right)dy$$
$$\geqq e^{-1/\tilde{\Theta}^{k-2}}\int_\Theta^1 \frac{1}{y^k}\exp\left(\frac{1}{(k-1)\,y^{k-2}}\right)dy$$
$$=e^{-1/\tilde{\Theta}^{k-2}}\int_{1/(k-1)}^{\tilde{\Theta}} z^{k/(k-2)}e^z\,dz\sim\Theta^{-k} \quad \text{as}\quad \Theta\to 0 \quad \text{with}\quad \tilde{\Theta}=\left(\frac{\Theta}{k-1}\right)^{-(k-2)}.$$

Therefore the denominator is asymptotically at least of the order Θ^{-k}. The first integral does not contribute, it is estimated as follows using the substitution $z=(\Theta(k-1)(k-2))^{-1}y^{-(k-1)}$:

$$\begin{aligned}
&\int_0^\Theta \frac{1}{y^k}\exp\left(-\left(\frac{\Theta}{k-1}-\frac{y}{k-2}\right)\Big/ y^{k-1}\right)dy \\
&\leqq \int_0^\Theta \frac{1}{y^k}\exp\left(\frac{-1}{\Theta(k-1)(k-2)}\,\frac{1}{y^{k-1}}\right)=-\int_{c/\Theta^{k+1}}^{\infty}\Theta\frac{1}{k-2}e^{-z}dz \\
&\leqq \tilde{c}\Theta e^{-1/\Theta^{k+1}}=o(\Theta^j) \qquad \text{for all } j>0.
\end{aligned}$$

2. The behaviour of $\exp\left(\int_\Theta^y \frac{\Theta-z}{g(z)}dz\right)$ for $\Theta\to 0$.

This quantity is easily seen to be bounded above and below by numbers $c_1>0$, $c_2<\infty$, respectively.

Therefore by putting things together we have for the special choice of g that $u_g(x)\geqq x^k$ for $x\to 0$. Using monotonicity it is clear that the statement remains valid for all g with $g(x)\sim x^k$ as $x\to 0$. The details of this exercise in analysis are left to the reader. The same reasoning applies to $x\to 1$. This proves the assertion for $\beta>2$.

3. $\beta=2$

Again it is enough to consider the case $g(x)=x^2$.

The constant C_Θ is now given by

$$\int_0^1 \exp\left(-\left(\frac{\Theta}{y}+\ln y\right)\right)dy=\int_0^1\frac{1}{y}e^{-\Theta/y}dy=\int_{-\Theta}^{-\infty}\frac{1}{z}e^z dz$$
$$\sim -\ln\Theta \quad \text{as} \quad \Theta\to 0$$

using the substitution $z=-\Theta/y$.

The density is now treated as follows:

$$\int_0^1\frac{1}{y^2}\exp\left(-\left(\frac{\Theta}{y}+\ln y\right)\right)dy=\int_0^1\frac{1}{y^3}\exp(-\Theta/y)\,dy$$
$$\sim\frac{1}{\Theta^2}\int_{-\Theta}^{-\infty} z e^z dz\sim \text{const}\,\Theta^{-2} \quad \text{as} \quad \Theta\to 0.$$

Putting both things together yields

$$u_g(\Theta)\sim\text{const}\,|\ln\Theta|\,\Theta^2 \quad \text{as} \quad \Theta\to 0$$

which proves the assertion in the case $\beta=2$. □

Proof of Theorem 2 Since $g(x)=cx(1-x)$ the invariant measure is

$$\Gamma_\Theta(dx)=\frac{\text{const}}{x(1-x)}\exp\left(\int_\Theta^x\frac{\Theta-y}{cy(1-y)}dy\right)=\text{const}\,x^{\frac{\Theta}{c}-1}(1-x)^{\frac{1-\Theta}{c}-1},$$

and by (1.12)

$$F(cx(1-x))(\Theta)=c\frac{\int_0^1 x^{\Theta/c}(1-x)^{(1-\Theta)/c}dx}{\int_0^1 x^{(\Theta/c)-1}(1-x)^{((1-\Theta)/c)-1}dx}$$

$$=\frac{cB\left(\frac{\Theta}{c}+1,\frac{1-\Theta}{c}+1\right)}{B\left(\frac{\Theta}{c},\frac{1-\Theta}{c}\right)} \quad \text{(see Abramowitz and Stegun [1, p. 256])}$$

$$=c\frac{\Gamma\left(\frac{\Theta}{c}+1\right)\Gamma\left(1+\frac{1-\Theta}{c}\right)\Gamma\left(\frac{1}{c}\right)}{\Gamma\left(\frac{\Theta}{c}\right)\Gamma\left(\frac{1-\Theta}{c}\right)\Gamma\left(2+\frac{1}{c}\right)}=\frac{c}{c+1}\Theta(1-\Theta). \quad \square$$

(b) Proof of Theorem 3

Proof of Theorem 3 (a). The proof is based on the techniques from Sects. 2, 3. Here we shall explain how to adapt our previous arguments to the new situation. The first observation is that we can transform the state space of the components of the interacting system to $[-\frac{1}{2},\frac{1}{2}]$ and $A=0$, g to $g\circ h$ where h is piecewise linear. Assume now that we have done this transformation, then we can rewrite the system (1.14) in the form

$$y_i^N(t)=y_i^N(0)\,e^{-c(N)t}+\int_0^t e^{-c(N)(t-s)}d\tilde{y}_i^N(s) \tag{4.16}$$

where

$$d\tilde{y}_i^N(t)=\left(\frac{1}{2N+1}\sum_{j=-N}^{N}(y_j^N(t)-y_i^N(t))\right)dt+\sqrt{2g(y_i^N(t))}\,dw_i(t).$$

This implies that

$$\Theta_t^N=\Theta_0^N\,e^{-c(N)t}+\int_0^t e^{-c(N)(t-s)}d\tilde{\Theta}_s^N \tag{4.17}$$

$$d\tilde{\Theta}_s^N=\frac{1}{2N+1}\sum_{i=-N}^{N}\sqrt{2g(y_i^N(s))}\,dw_i(s). \tag{4.18}$$

From (4.18) it is clear that $\tilde{\Theta}_s^N$ is a martingale whose increasing process is a functional of $(y_i^N(u), i\in[-N,N])$ and with the very same arguments as in Sect. 3, we can now prove (we do not repeat the argument here) that $\mathscr{L}((\tilde{\Theta}_{sN}^N)_{s\in\mathbf{R}^+})$ is tight and that once we have proved the analog of Proposition 1, (2.1) (with Θ_s^N replaced by $\tilde{\Theta}_s^N$) we will obtain:

$$\mathscr{L}(((\Theta_{s\beta(N)}^N)_{s\in\mathbf{R}^+},(\tilde{\Theta}_{s\beta(N)}^N)_{s\in\mathbf{R}^+})\underset{N\to\infty}{\Rightarrow}\mathscr{L}((\Theta_s)_{s\in\mathbf{R}^+},(\tilde{\Theta}_s)_{s\in\mathbf{R}^+})$$

$$d\tilde{\Theta}_u=\Phi(\Theta_u)\,dw(u)$$

with

$$\Phi(\Theta) = E^{\nu_\Theta}(g(y_0)).$$

We can then realize these processes on a common probability space such that

(4.19) $$\tilde{\Theta}^N_{s\beta(N)} \underset{N\to\infty}{\Rightarrow} \tilde{\Theta}_s \quad \text{a.s.} \quad \text{in} \quad \|\cdot\|_\infty - \text{norm}.$$

This implies with (4.17) (everything takes place on $[-\frac{1}{2}, \frac{1}{2}]$) that $((\Theta_s)_{s\geqq 0}, (\tilde{\Theta}_s)_{s\geqq 0})$ satisfy:

(4.20) $$\Theta_s = \Theta_0 e^{-cs} + \int_0^s e^{-c(s-u)} d\tilde{\Theta}_u.$$

Together with Lemma 2.2 this implies that: (cf. Ethier and Kurtz [8, p. 290–307]),

(4.21) $(\Theta_s)_{s\in\mathbb{R}^+}$ (is a non-Gaussian Ornstein-Uhlenbeck type process) generated by

$$-cx\frac{\partial}{\partial x} + \Phi(x)\left(\frac{\partial}{\partial x}\right)^2.$$

It remains therefore to prove the analog of Proposition 1. First of all to get (4.19) above and second to be able to repeat the argument of Sect. 4(a) to also verify (1.6), (1.7) for the system in a weak external field. However the latter is straightforward and proceeds exactly the same way as in 4(a), since the McKean-Vlasov limit of our new system is the same as before.

The proof of Proposition 1 uses *only* the statements made in Lemmas 2.1–2.5. It therefore suffices to establish these lemmas in our new situation. Since our new system and the system in part 1(a) both have as McKean-Vlasov limit independent diffusions with linear drift term, the only new thing we have to do here is to show that the assertion of Lemma 2.3 also holds in the present situation.

That is, it remains to show that the finite system and the infinite system started in the periodic continuation of the initial state of the finite system, remain close for time $t \ll T(N)$ for some $T(N)$ such that $T(N) \to \infty$ but $T(N) = o(N)$.

Again since the McKean-Vlasov limit of our new system with the extra term $-c(N)\, y_i^N(t)\, dt$ is of the same type as for the model with $c(N)=0$ treated earlier it suffices to show, that for suitable $L(N)$, the following systems are close for all $t \leqq L(N)$ as $N\to\infty$ (recall $A=0$, transformation of the state space to $[-1/2, 1/2]$):

(4.22) $$dy_i^N(t) = \left(\frac{1}{2N+1}\sum_{j=-N}^{N}(y_j^N(t) - y_i^N(t))\right)dt - \frac{c}{N}\, y_i^N(t)\, dt + \sqrt{2g(y_i^N(t))}\, dw_i(t)$$

(4.23) $$d\tilde{y}_i^N(t) = \left(\frac{1}{2N+1}\sum_{j=-N}^{N}(\tilde{y}_j^N(t) - \tilde{y}_i^N(t))\, dt + \sqrt{2g(\tilde{y}_i^N(t))}\, dw_i(t)\right)$$

which start in one and the same initial configuration. The comparison of these systems is easy since we have the representation (cf. (4.16)):

$$(4.24)\qquad y_i^N(t)=y_i^N(0)\,e^{-ct/N}+\int_0^t e^{-c(t-s)/N}\,d\bar{y}_i^N(s)$$

$$d\bar{y}_i^N(t)=\left(\frac{1}{2N+1}\sum_{j=-N}^{N}(y_j^N(t)-y_i^N(t))\,dt+\sqrt{2g(y_i^N(t))}\,dw_i(t)\right)$$

and therefore for $t\leqq L(N)$:

$$(4.25)\qquad \begin{aligned} y_i^N(t)&=\bar{y}_i^N(0)+(\bar{y}_i^N(t)-\bar{y}_i^N(0))-(1-e^{-ct/N})\,\bar{y}_i^N(0)\\ &\quad-\int_0^t(1-e^{-c(t-s)/N})\,d\bar{y}_i^N(s)\\ &=\bar{y}_i^N(t)+O(L(N)^2/N)\quad \text{uniform in } t\leqq L(N) \text{ as } N\to\infty.\end{aligned}$$

Using this in the right side of the second equation in (4.24) (compare (2.24)–(2.32) for details) gives then that $\mathscr{L}(\tilde{y}_i(t))-\mathscr{L}(\bar{y}_i(t))\to 0$ as $N\to\infty$ for all $t\leqq L(N)$ if $L(N)$ is such that $L(N)=o(N^{\frac{1}{3}})$. This completes the proof of Theorem 3(a).

Proof of Theorem 3(b) In order to prove Theorem 3(b), it suffices to have established the following two facts:
(i) Suppose that $N_k\to\infty$ as $k\to\infty$ and $M(N)\uparrow+\infty$ as $N\to\infty$, $M(N)\gg N$. If

$$(4.26)\qquad \mathscr{L}(\Theta^{N_k}_{M(N_k)})\Rightarrow P(\cdot),\qquad \mathscr{L}(Y^{N_k}_{M(N_k)})\Rightarrow \nu,$$

then

$$\nu=\int \nu_\Theta\,dP(\Theta).$$

(ii) $$(4.27)\qquad \mathscr{L}(\Theta^{N_k}_{M(N_k)+t})_{t\in\mathbf{R}^+}\Rightarrow\mathscr{L}((\Theta_t)_{t\geqq 0})\quad\text{with}\quad \mathscr{L}(\Theta_0)=\nu^A.$$

The proof of the first assertion was given before in the proof of Theorem 3(a). (See in particular the sequel of (4.21).)

A little bit more subtle is the second fact. It is clear that

$$\mathscr{L}(\Theta_s)\underset{s\to\infty}{\Rightarrow}\Lambda^A,$$

so the question is whether the uniformity (in s) in $\mathscr{L}(\Theta^N_{s\beta(N)})\Rightarrow\mathscr{L}(\Theta_s)$ as $N\to\infty$ is sufficiently strong to give for a sequence $s(N)\to\infty$ as $N\to\infty$ that:

$$\mathscr{L}(\Theta^N_{s(N)\beta(N)})\Rightarrow\Lambda^A\quad\text{as}\quad N\to\infty.$$

In order to see this required uniformity we shall need some preparation. Define

$$(4.28)\qquad \bar{\Theta}^N_s=\Theta^N_{s\beta(N)}.$$

By definition of $\Theta^N_{s\beta(N)}$ and by the scaling property of Brownian motion, we can rewrite the stochastic differential equation for $\bar{\Theta}^N_s$ as follows:

$$(4.29)\qquad d\bar{\Theta}^N_s = c(A-\bar{\Theta}^N_s)\,ds + \frac{1}{\sqrt{2N+1}}\sum_{i=-N}^{N}[\sqrt{2g(y_i^N(s\beta(N)))}\,dw_i(s)].$$

From this equation we derive below (in order to obtain the required uniformity in s in the convergence of $\Theta^N_{s\beta(N)}$ to Θ_s) that it suffices to show

$$(4.30)\qquad E\left|\frac{1}{2N+1}\sum_{i=-N}^{N} g(y_i^N(s\beta(N))) - E^{\nu(\Theta_s)}g(y_0)\right|^2 \leqq \varepsilon(N), \quad \forall s\in[0,\infty)$$

$$\text{with}\quad \varepsilon(N)\to 0 \quad \text{as} \quad N\to\infty,$$

To see this fact note that with $w(s)$ denoting a version of standard Brownian motion and $\Phi(\Theta) = E^{\nu_\Theta}(g(y_0))$:

$$\sqrt{2\Phi(\Theta_s)}\,dw(s) = \frac{1}{\sqrt{2N+1}}\sum_{i=-N}^{N}\sqrt{2\Phi(\Theta_s)}\,dw_i(s),$$

so that we can represent Θ_s in the form

$$d\Theta_s = \frac{1}{\sqrt{2N+1}}\sum_{i=-N}^{N}\sqrt{2\Phi(\Theta_s)}\,dw_i(s).$$

Now use this coupling between Θ_s and $\bar{\Theta}^N_s$. The difference $\Theta_s - \bar{\Theta}^N_s$ is then represented as in (4.16). We see that $\bar{\Theta}^N_s$ and Θ_s on this probability space satisfy $E(\bar{\Theta}^N_s - \Theta_s)^2 \leqq \varepsilon(N)$, $\forall s \geqq 0$, if (4.30) holds.

It remains to verify (4.30). Due to Theorem 1(a) for every *fixed* $s\in[0,\infty)$ relation (4.30) holds trivially. The point is to show that the convergence as $N\to\infty$ is uniform in s. The key here is again an L_2-argument as in the proof of Lemma 2.5, which exploits the fact that $\nu(s) \Rightarrow \nu^A$ as $s\to\infty$. For that purpose we write:

$$z_i(N,s) = g(y_i^N(s\beta(N))), \qquad z_i(\infty,s) = g(y_i(s))$$
$$z_i(\infty,\infty) = g(y_i), \qquad \mathscr{L}(Y) = \nu^A.$$

We realize all these random variables on one big probability space, we call this measure P, such that:

$$z_i(N,s) \underset{N\to\infty}{\Rightarrow} z_i(\infty,s) \quad \text{a.s.}$$
$$z_i(\infty,s) \underset{s\to\infty}{\Rightarrow} z_i(\infty,\infty) \quad \text{a.s.}$$

and the marginal distributions $\mathscr{L}(\{z_i(N,s)\}_{i\in N})$ are the prescribed ones. Denote by

$$D^{N,s}=\frac{1}{2N+1}\sum_{i=-N}^{N} z_i(N,s), \quad D^{\infty,s}=\lim_{N\to\infty}\frac{1}{2N+1}\sum_{i=-N}^{N} y_i(s),$$

$$D=\lim_{N\to\infty}\frac{1}{2N+1}\sum_{i=-N}^{N} y_i.$$

Then we have to prove the following, in order to get (4.30):

$$\lim_{\substack{s\to\infty\\ N\to\infty}}\ \sup_{\substack{M\geq N\\ t\geq s}}(\|D^{N,s}-D^{M,t}\|_{L_2(P)})=0. \tag{4.31}$$

By writing

$$\begin{aligned}\|D^{N,s}&-D^{M,t}\|_{L_2(P)}\\ &\leq \|D^{N,s}-D^{\infty,s}\|_{L_2(P)}+\|D^{\infty,s}-D^{\infty,t}\|_{L_2(P)}+\|D^{\infty,t}-D^{M,t}\|_{L_2(P)}\\ &\leq \|D^{N,s}-D^{\infty,s}\|_{L_2(P)}+\|D^{\infty,s}-D\|_{L_2(P)}+\|D^{\infty,t}-D\|_{L_2(P)}+\|D^{\infty,t}\\ &\quad -D^{M,t}\|_{L_2(P)}\end{aligned} \tag{4.32}$$

we prove (4.31) following the scheme in the proof of Lemma 2.5, i.e. (2.62)–(2.63) to treat the first and fourth summands (the second and third are trivial). We omit the details and leave it to the reader to adapt the steps to the new situation. □

Proof of Theorem 3 (c) This can be derived from part (b) as follows. Suppose that $d[\cdot,\cdot]$ is a metric generating the weak topology of probability measures on $[0,1]^{\mathbb{Z}}$. For N fixed we choose a function $s(N)$ with $Ns(N)\geq N^2$ and $s(N)$ large enough so that

$$d[\mathscr{L}(Y^N_{(s(N)N)}), v^{A,N}]\to 0 \quad \text{as} \quad N\to\infty.$$

But since $s(N)\to\infty$ as $N\to\infty$ we know by Theorem 3(b) that

$$d[\mathscr{L}(Y^N_{(s(N)N)}), \textstyle\int v_\Theta\, d\Lambda_A(d\Theta)]\to 0 \quad \text{as} \quad N\to\infty.$$

Combining the last two relations gives, using the triangle inequality for the metric d, the assertion

$$d[\textstyle\int v_\Theta\, d\Lambda_A(\Theta), v^{A,N}]\to 0 \quad \text{as} \quad N\to\infty. \quad \square$$

(c) Proof of Theorem 4

Since the arguments here resemble the ones in Sects. 2 and 3 very closely we are very short here and indicate only briefly how to bring the techniques used to prove Theorem 1 and the result of Theorem 3 into play in this situation.

In particular it is easy to derive the analogue of Proposition 1 for the time scale $\beta_1(N)$, that is, given that the means converge in the time scale $\beta_1(N)$ along a subsequence, then the components converge in that time scale to a mixture of equilibria of the infinite system, the mixing measure being the limit law of the density process. This is since the term involving the means of $\bar{x}_i(t)$ values plays no role in the time scale $\beta_1(N)$.

On the other hand dealing with the averages in the time scale $\beta_2(N)$ it suffices to deal in (1.20) with the $\bar{x}_i(t)$ terms since the terms involving the components x_i themselves are zero. It is straightforward to show that it suffices to study the same system as before in Sect. 2 only replacing the diffusion coefficient g with v_g. We do not repeat these arguments here in detail, see Sect. 2 for that purpose, the modifications needed are straightforward.

In order to prove (1.26) we proceed as in Sect. 3, where the main point was to calculate the limit for $N \to \infty$ of the increasing processes $\langle \Theta^{2,N}_{s\beta_2(N)} \rangle$. We present below the necessary modifications to the argument of Sect. 3. Observe first that

$$d\Theta_t^{2,N} = \left(\frac{1}{2M(N)+1} \sum_{j=-M(N)}^{M(N)} d\bar{x}_j^N(s) \right),$$

from which we obtain:

$$d\Theta^{2,N}_{t,\beta_2(N)} = \frac{1}{\sqrt{2M(N)+1}} \sum_{i=-M(N)}^{M(N)} \left(\frac{1}{\sqrt{2N}} \sum_{j \in I_N(i)} g(x_{j,0}(t\beta_2(N)))\, dw_j(t) \right) \tag{4.33}$$

$$d\langle \Theta^{2,N}_{t\beta_2(N)} \rangle = \frac{1}{2M(N)+1} \sum_{i=-M(N)}^{M(N)} \left(\frac{1}{2N} \sum_{j \in I_N(i)} g(x_{j,0}(t\beta_2(N))) \right) dt. \tag{4.34}$$

From the representation of $\Theta_t^{2,N}$ above we see (compare Sect. 3 for details) that

$$\sup_{s \leqq T(N)} |\Theta^{2,N}_{t+s} - \Theta^{2,N}_t| \to 0 \quad \text{in probability as} \quad N \to \infty \tag{4.35}$$

if $T(N) = o(\beta_2(N))$ where t may depend on N and furthermore

$$\mathscr{L}((\Theta^{2,N}_{s\beta_2(N)})_{s \in \mathbf{R}^+}) \text{ is tight.} \tag{4.36}$$

We are then left to prove that the weak limit point for $N \to \infty$ of the sequence (4.36) is unique also in this model and to identify it via the increasing process which proves (1.26) and then we have to show that this implies (1.27), (1.28). This is done in two steps, starting with the second part.

First observe that (4.35) means that our original system (1.20) behaves, for $N \to \infty$ and times $s\beta_2(N) + t$ with times $t = o(\beta_2(N))$, like $2M(N)+1$ independent systems with weak external field. In particular in the time scales 1 and $\beta_1(N)$, $\Theta_t^{2,N}$ is essentially constant and the original system can be approximated by considering $2M(N)+1$ independent blocks $(\tilde{x}_j^i(t))_{j \in [-N, N-1]}$, $i \in [-M(N), M(N)]$ of

$$d\tilde{x}_j^i(t) = \left(\frac{1}{2N} \sum_{k=-N}^{N} \tilde{x}_k^i(t)) - \tilde{x}_j^i(t) \right) dt + \frac{c}{N}(A - \tilde{x}_j^i(t))\, dt + \sqrt{2g(\tilde{x}_j^i(t))}\, dw_j^i(t) \tag{4.37}$$

where we substitute for A the value of $\Theta_s^{2,N}$. Applying Theorem 3 to each of the $2M(N)+1$ systems we see that for $i\in[-M(N), M(N)]$ the process $\frac{1}{2N}\sum_{k=-N}^{N-1} \tilde{x}_k^i(s\beta_2(N)+\beta_1(N)\,t)$ can be approximated by $\hat{x}^i(\beta_1(N)\,t)$ where

$$(4.38)\qquad d\hat{x}^i(\beta_1(N)\,t)=c(A-\hat{x}^i(\beta_1(N)\,t))\,dt+\sqrt{2v_g(\hat{x}^i(t\beta_1(N))}\,d\hat{w}_i(t),$$
$$\hat{x}^i(0)=(2N)^{-1}\sum_{k=-N}^{N} \tilde{x}_i^i(s\beta_2(N)).$$

This means that, assuming for the moment that (1.26) holds, the proofs of (1.27) and (1.28) now follow along the same lines as the proofs of (1.6) and (1.7).

In order to verify (1.26) we have to consider the time scale $\beta_2(N)$. We must replace (4.38) with a system for $\hat{x}=\frac{1}{2N}\sum \hat{x}^i$:

$$(4.39)\qquad d\hat{x}(\beta_2(N)\,t)=\frac{1}{\sqrt{2M(N)+1}}\sum_{i=-M(N)}^{M(N)} \sqrt{2v_g(\hat{x}^i(t\beta_2(N))}\,d\hat{w}_i(t).$$

Using the same line of reasoning as in the proof of Proposition 2 in Sect. 3 we can continue (4.34) as follows

$$(4.40)\qquad d\langle\Theta_{s\beta_2(N)}^{2,N}\rangle \simeq \frac{1}{2M(N)+1}\sum_{i=-M(N)}^{M(N)} \left(\frac{1}{2N}\sum_{j=-N}^{N-1} g(\tilde{x}_j^i(s\beta_2(N)))\right)$$
$$\simeq \frac{1}{2M(N)+1}\sum_{i=-M(N)}^{M(N)} v_g(\hat{x}^i((s\beta_2(N)))$$
$$\simeq u_g(\Theta_{s\beta_2(N)}^{2,N})\,ds.$$

As in Lemma 3.3 this establishes the uniqueness of the weak limit point of $\mathscr{L}((\Theta_{t\beta_2(N)}^{2,N})_{t\in\mathbf{R}^+})$ as $N\to\infty$, as the unique solution of:

$$(\Theta_s^2)^2-\int_0^s u_g(\Theta_u^2)\,du \quad \text{is a martingale.}$$

This yields (1.26) thus completing the proof of Theorem 4.

Acknowledgements. The research of the first author was supported by an NSERC operating grant. The second author benefited greatly from the earlier joint work [3] with Ted Cox.

References

1. Abramowitz, M., Stegun, I.A.: Handbook of mathematical functions. London: Dover 1965
2. Breiman, L.: Probability. Reading, Mass: Addison-Wesley 1968
3. Cox, T., Greven, A.: On the long term behaviour of some finite particle systems. Probab. Theory Relat. Fields **85**, 195–237 (1990)
4. Cox, T., Greven, A.: On the longterm behaviour of finite particle systems: a critical dimension example. In: Durrett, R., Kesten, H. (eds.) (Prog. Probab., vol. 28) Random walks,

Brownian motion and interacting particle systems. A Festschrift in Honor of Frank Spitzer, pp. 203–213. Boston Basel Stuttgart: Birkhäuser 1991
5. Dawson, D., Gärtner, J.: Long time behavior of interacting diffusions. In: Norris, J.P. (ed.) Stochastic calculus in applications. Proceedings of the Cambridge Symposium, 1987. (Pitman Res. Notes Math. Ser., vol. 197) London: Longman 1988
6. Dawson, D., Greven, A.: Hierarchical models of interacting diffusions: multiple time scale phenomena, phase transition and cluster-formation. Probab. Theory Relat. Fields (to appear)
7. Dawson, D., Greven, A.: Multiple time scale analysis of hierarchically interacting systems. In: Cambanis, S., Ghosh, J.K., Karandikar, R.L., Sen, P.K. (eds.) Festschrift in Honor of G. Kallianpur, pp. 41–50. Berlin Heidelberg New York: Springer 1993
8. Ethier, S., Kurtz, T.: Markov processes: characterisation and convergence. New York: Wiley 1986
9. Greven, A.: Large systems with locally interacting components. In: den Hollander, F., Maassen, H. (eds.) Syllabus of the Marc Kac seminars in probability and mathematical physics. (CWI Syllabus, pp. 85–88.) Amsterdam: Math. Centrum 1988
10. Lamperti, J.: Stochastic processes. Berlin Heidelberg New York: Springer 1977
11. Stroock, D., Varadhan, S.: Multidimensional diffusion processes. Berlin Heidelberg New York; Springer 1979
12. Yamada, T., Watanabe, S.: On the uniqueness of solutions of stochastic differential equations. J. Math. Kyoto Univ. **11**, 155–167 (1971)
13. Yosida, K.: Functional analysis. Berlin Heidelberg New York: Springer 1965