Journal of Statistical Physics, Vol. 55, Nos. 3/4, 1989

Transient Bimodality in Interacting Particle Systems

P. Calderoni,[1] A. Pellegrinotti,[2] E. Presutti,[2] and M. E. Vares[3]

Received July 26, 1988

We consider a system of spins which have values ± 1 and evolve according to a jump Markov process whose generator is the sum of two generators, one describing a spin-flip *Glauber* process, the other a *Kawasaki (stirring)* evolution. It was proven elsewhere that if the Kawasaki dynamics is speeded up by a factor ε^{-2}, then, in the limit $\varepsilon \to 0$ (continuum limit), propagation of chaos holds and the local magnetization solves a reaction-diffusion equation. We choose the parameters of the Glauber interaction so that the potential of the reaction term in the reaction-diffusion equation is a double-well potential with quartic maximum at the origin. We assume further that for each ε the system is in a finite interval of Z with ε^{-1} sites and periodic boundary conditions. We specify the initial measure as the product measure with 0 spin average, thus obtaining, in the continuum limit, a constant magnetic profile equal to 0, which is a stationary unstable solution to the reaction-diffusion equation. We prove that at times of the order $\varepsilon^{-1/2}$ propagation of chaos does not hold any more and, in the limit as $\varepsilon \to 0$, the state becomes a nontrivial superposition of Bernoulli measures with parameters corresponding to the minima of the reaction potential. The coefficients of such a superposition depend on time (on the scale $\varepsilon^{-1/2}$) and at large times (on this scale) the coefficient of the term corresponding to the initial magnetization vanishes (transient bimodality). This differs from what was observed by De Masi, Presutti, and Vares, who considered a reaction potential with quadratic maximum and no bimodal effect was seen, as predicted by Broggi, Lugiato, and Colombo.

KEY WORDS: Interacting particle systems; hydrodynamic behavior; critical fluctuations; escape from unstable equilibrium; bimodality.

1. INTRODUCTION

The appearance of collective modes and self-organization phenomena is one of the most remarkable features of the dynamics of complex systems

[1] Dipartimento di Matematica dell'Università di Roma "Tor Vergata," 00167 Rome, Italy.
[2] Dipartimento di Matematica dell'Università di Roma "La Sapienza," 00187 Rome, Italy.
[3] IMPA, Rio de Janeiro, 22460 R.J., Brazil.

with many components. While the evolution in such systems is determined by *purely local* interactions among its elementary parts, yet, as time goes by, a coherent behavior establishes throughout the system as described by macroscopic fields which evolve according to closed equations, referred to as the *hydrodynamic equations for the system*, nonlinear PDEs, in the most interesting cases. Physical fluids are the archetype of such a behavior: while molecules obey the Hamiltonian laws of motion, the evolution of the thermodynamic fields describing the macroscopic properties of the system is determined by Euler and Navier–Stokes types of equations. Completely different systems behave in an analogous fashion; many examples originate from biological systems, population dynamics, economic models, computer simulations (cellular automata), and so forth. The *robustness* of this behavior is a clue for its analysis; its persistence even in oversimplified models makes a mathematical investigation possible. In stochastic interacting particle systems such a transition from microscopic to macroscopic has been observed and, in several cases, analyzed with mathematical rigor, providing a clear understanding of the mechanisms which cause the development of the collective behavior.

Such an analysis shows that the hydrodynamic equations describe the behavior of the system in the *hydrodynamic or continuum limit*, i.e., in the limit of very extended systems and very small gradients (of the extensive fields). The actual state of the system is therefore close to the predictions of the hydrodynamic equations, when approaching the hydrodynamic limit. Its small deviations are described by *fluctuation fields*, which have, in the limit, after proper normalization, *Gaussian distribution*. However, in critical situations, such fluctuations are amplified and they eventually become macroscopic. On such a longer time scale the Gaussian nature of their distribution is lost due to the appearance of nonlinear effects so that their evolution is described by nonlinear stochastic PDEs.

A very beautiful example of such a behavior can be found in the experimental work of Meyer *et al.*[18] They consider a Rayleigh–Bénard cell. By increasing the temperature difference between the two plates of the cell past a critical value, convective flows appear in the fluid. Slightly above the critical value the nonconvective state is unstable. In usual conditions the onset of convection is determined by external conditions, e.g., the geometry and nature of the walls of the container. The main problem in order to see possible intrinsic fluctuations in the system is to screen these and other external effects. This is achieved under very careful experimental conditions, as shown by the convective patterns, which no longer reflect the geometry of the container: at the early stage of the convection "*they are irregularly arranged and vary randomly between experimental runs, suggesting the importance of stochastic effects during the pattern evolution.*"

We refer to ref. 18 for more details and to their figures for evidence of the above statements.

Also from a mathematical point of view these phenomena have considerable interest. A hydrodynamic description in terms of macroscopic fields involves the validity of the law of large numbers, since the local value of a field represents the average of a corresponding local observable in a microscopically *infinite* but macroscopically *infinitesimal* region. By the law of large numbers such an average has a sharp value. In the phenomena we wish to discuss this is not the case. On a longer time scale the law of large numbers fails and the macroscopic description is given by a statistical mixture of profiles rather than by a single one. We shall observe this phenomenon in a spin system model whose hydrodynamic equation is a reaction-diffusion equation. The initial state is chosen in such a way that the limiting macroscopic profile is a stationary unstable solution to the macroscopic equation. Therefore on the time scale where the hydrodynamic equation is derived, the actual microscopic state does not change remarkably. On a longer time scale, deviations become macroscopically important and eventually the system leaves its initial unstable equilibrium. The phenomenon is very sensitive to the structure of the microscopic interaction. We consider here the case when the microscopic interaction leads to a reactive potential with a quartic maximum and compare the results with those in ref. 10 for the quadratic case. A qualitative and quantitative discussion and its physical implications are presented in the next section, proofs are reported in Sections 3 and 4, and in the Appendices some estimates of a more technical nature are established.

A few more considerations before closing this section. Critical fluctuations appear also in the analysis of the long-time behavior of shock wave profiles. As observed in refs. 8, 9, 15, and 20 for some asymmetric simple exclusion processes, traveling wave profiles remain stable even at the microscopic level. Only their location becomes random: they rigidly fluctuate in space. As for the escape from an unstable equilibrium, we have here again a macroscopic solution (for an observer moving with the speed of the wave) which may not give the actual behavior of the system. The microscopic state in this case is, however, not unstable but only marginally stable. The onset of critical fluctuations occurs then on a much longer time scale and it produces a Brownian-like motion of the shock around its average position, the phenomenon being structurally similar to that occurring when studying small random perturbations of dynamical systems with stable manifolds.[6]

2. RESULTS

The Model. For each positive ε such that ε^{-1} is an integer, we let Z_ε be the set of all integers with the identification $x = x + \varepsilon^{-1}$. We then consider a ± 1-valued spin system in Z_ε, denote by σ a spin configuration, i.e., an element of $\{-1, 1\}^{Z_\varepsilon}$, and by $\sigma(x)$, $x \in Z_\varepsilon$, the spin at x in the configuration σ. The *Glauber + Kawasaki* evolution on Z_ε is the jump Markov process on $\{-1, 1\}^{Z_\varepsilon}$ with generator

$$L^\varepsilon = \varepsilon^{-2} L_0 + L_G \tag{2.1}$$

where, for any function f on $\{-1, 1\}^{Z_\varepsilon}$,

$$L_0 f(\sigma) = \sum_{x \in Z_\varepsilon} \tfrac{1}{2}[f(\sigma^{x,x+1}) - f(\sigma)] \tag{2.2}$$

$$L_G f(\sigma) = \sum_{x \in Z_\varepsilon} c(x, \sigma)[f(\sigma^x) - f(\sigma)] \tag{2.3}$$

$\sigma^{x,x+1}$ is obtained from σ by interchanging its values at x and $x+1$ (*stirring or Kawasaki dynamics*), σ^x by flipping the spin at x (*Glauber dynamics*). The function $c(x, \sigma)$ is the flip intensity of the Glauber evolution and it is assumed to be (i) strictly positive, (ii) translationally invariant, i.e.,

$$c(x, \sigma) = c(0, \sigma + x)$$

($\sigma + x$ being the translate of σ to the left by x), and (iii) depending on finitely many spins of σ.

We shall eventually restrict ourselves to a very specific choice of the Glauber intensities, namely

$$\begin{aligned} c(x, \sigma) = {} & [1 - \tfrac{1}{2}\sigma(x)\,\sigma(x+1)][1 - \tfrac{1}{2}\sigma(x)\,\sigma(x-1)] \\ & [1 - c\sigma(x)\,\sigma(x+2)\,\sigma(x+3)\,\sigma(x+4)] \end{aligned} \tag{2.4}$$

where $1 > c > 1/4$, and to simplify notation

$$c = 3/4$$

The reason for these choices will become clear later, for the moment it is convenient to keep $c(x, \sigma)$ in its general form, namely as only specified by the conditions (i), (ii), and (iii).

To complete the definition of the model, we choose the initial measure μ^ε as the product probability on $\{-1, 1\}^{Z_\varepsilon}$ whose mean values are

$$\mu^\varepsilon(\sigma(x)) = m(\varepsilon x) \tag{2.5}$$

where $\mu^\varepsilon(f)$ denotes the expectation of f and $m(r)$ in (2.5) is a smooth function on the unit circle whose absolute value does not exceed 1. We shall denote by μ_t^ε the distribution of the spins at time t. Further assumptions on $m(r)$ will be made in due time.

The hydrodynamic limit for this class of models was studied first in ref. 7, actually in a more general setup. The hydrodynamic equations are reaction-diffusion equations; more precisely, the following holds. For all integers n

$$\lim_{\varepsilon \to 0} \sup_{x_1 \cdots x_n} \left| \mu_t^\varepsilon \left(\prod_{i=1}^{n} \sigma(x_i) \right) - \prod_{i=1}^{n} m(\varepsilon x_i, t) \right| = 0 \tag{2.6}$$

where the sup in (2.6) is over all n-tuplets of distinct sites in Z_ε and $m(r, t)$ solves the equation

$$\frac{\partial}{\partial t} m = \frac{\partial^2}{\partial r^2} m - V'(m), \qquad m(r, 0) = m(r) \tag{2.7}$$

$$-V'(m) = \nu_m(-2\sigma(0)\, c(0, \sigma)) \tag{2.8}$$

Finally, ν_m is the Bernoulli measure (product measure) with spin average equal to m.

Equation (2.7) is the hydrodynamic equation for the model, while (2.6) represents a very strong form of the propagation-of-chaos property. Equation (2.8) has a very simple interpretation as the average spin change due to the Glauber interaction, where the average is taken with respect to a Bernoulli measure, an invariant measure for the stirring evolution alone. In fact, as is intuitively clear, when $\varepsilon \to 0$ the prefactor ε^{-2} in front of the stirring generator forces the spin configurations for the full evolution to be typical for the stirring evolution alone.

These models have a very rich and interesting structure. If, for instance, V is a double-well potential, there is a soliton solution to (2.7) [namely a stationary solution which *connects* the two stable points (minima) of the potential; here we consider the whole space and not the unit circle, so the spin model has to be changed accordingly]. Traveling waves connecting the stable to the unstable point of the potential are also stationary solutions for a moving observer. Another critical solution to (2.7) is the constant profile with magnetization equal to the unstable point of the potential. In all these cases the considerations of the introduction apply.

In this paper we shall restrict our attention to the case of the space-homogeneous, unstable equilibrium profile. This has already been studied in ref. 10 for a model with flipping intensities

$$c(x, \sigma) = \; = 1 - \gamma\sigma(x)[\sigma(x+1) + \sigma(x-1)] + \gamma^2\sigma(x+1)\,\sigma(x-1) \tag{2.9a}$$

$$1/2 < \gamma \leqslant 1 \tag{2.9b}$$

so that by (2.8) the potential $V(m)$ becomes

$$V(m) = \tfrac{1}{2}\alpha m^2 + \tfrac{1}{4}\beta m^4 \tag{2.10a}$$

$$\alpha = 2(2\gamma - 1), \qquad \beta = 2\gamma^2 \tag{2.10b}$$

Hence $m(r) = 0$, for all r (in the unit circle), is a stationary solution, so that, by (2.6), μ_t^ε converges to the Bernoulli measure with parameter 0, when $\varepsilon \to 0$ if t is kept fixed. If instead we let t depend on ε in such a way that $t \to \infty$ as $\varepsilon \to 0$, then it might happen that the intrinsic fluctuations describing the difference between the actual particle model and the limiting equation will make the particle model leave the unstable equilibrium so that the magnetization *at such long times* will reach a finite nonzero value. This is indeed what happens. In fact, as proven in ref. 10,

$$\lim_{\varepsilon \to 0} \mu^\varepsilon_{\tau \log \varepsilon^{-1}} = c(\tau)\nu_0 + \frac{1 - c(\tau)}{2}(\nu_{m^*} + \nu_{-m^*}) \tag{2.11}$$

which holds for all $\tau \neq (2\alpha)^{-1}$; $c(\tau)$ in (2.11) is the step function with value 1 for $\tau < (2\alpha)^{-1}$ and 0 if the reverse inequality holds; $\pm m^*$ are the minima of the potential $V(m)$. We leave out the details of what happens when $\tau = (2\alpha)^{-1}$, they can be found in ref. 10.

The above result states that there is a new *critical time scale* $(\log \varepsilon^{-1})$ which describes the escape from the unstable equilibrium. What is *a priori* surprising is that such an escape is sharp and it happens at a deterministic time, when the phenomenon is observed in the proper time scale. There is a critical time $= (2\alpha)^{-1}$ before which the magnetization is still 0, while afterward it equals either one of the two stable values $\pm m^*$ (with equal probability). Even though the whole phenomenon originates from a stochastic fluctuation, the time when the magnetization becomes macroscopic is fixed and it is not stochastic; the event is *completely predictable.* These kinds of phenomena are studied in laser physics and we learnt from ref. 3 that the case considered above is in a sense exceptional, namely for a reactive potential with a maximum flatter than quadratic one expects a completely different phenomenology for the escape from the unstable equilibrium.

To check this on a true particle system we have chosen the intensities as in (2.4) so that the potential $V(m)$ in (2.8) becomes

$$V(m) = -\tfrac{1}{4}[2(c - \tfrac{1}{4})] m^4 + \tfrac{1}{6}(\tfrac{3}{2}c) m^6$$

which, for $c = 3/4$, equals

$$V(m) = -\tfrac{1}{4}m^4 + \tfrac{3}{16}m^6 \tag{2.12}$$

From now on we shall restrict our considerations to this case, namely in Z_ε, with intensities as in (2.4) and with $c = 3/4$. Our main result is the following.

2.1. Theorem. Let μ^ε be the product measure on $\{-1, 1\}^{Z_\varepsilon}$ with average 0. Then

$$\lim_{\varepsilon \to 0} \mu^\varepsilon_{\varepsilon^{-1/2}t} = [1 - c(t)]\, \nu_0 + \frac{c(t)}{2}(\nu_{m^*} + \nu_{-m^*}) \tag{2.13}$$

where ν_m is the Bernoulli measure with average m and $[1 - c(t)]$ denotes the probability that $x(t)$ is finite, where $x(t)$ solves

$$dx = x^3\, dt + dw, \qquad x(0) = 0 \tag{2.14}$$

$w(t)$ being the standard Wiener motion.

Remarks. The escape from the unstable equilibrium for a potential with quartic maximum is therefore a truly stochastic event. Assume we run a computer simulation. Then for the quadratic case the escape time is always the same (on the time scale $\log \varepsilon^{-1}$) no matter what run is actually observed. In the quartic case the time scale is *much longer*, i.e., $\varepsilon^{-1/2}$. On this cale the time of escape is stochastic; it changes from one run to the other. It is not, however, *completely unpredictable*: by this we mean that the time one has to wait after t to find a finite magnetization (assuming that at time t it was still infinitesimal) is *stochastically* shorter than if $t = 0$.

The distribution of an unpredictable escape time is exponential. To have such a distribution it is not enough to have a maximum of the potential which is *very flat*; one really needs to replace the unstable maximum with a stable minimum so that we have a three-well potential with the two extremal minima possibly *much deeper* than the central one at $m = 0$. In this case the system has to *fight* against a drift to reach a finite magnetization and the escape (on a much longer time scale) has an exponential distribution. We refer to ref. 3 for a very clear and interesting discussion on these points in the context of optical physics, and to ref. 4 for a mathematically rigorous analysis of the tunneling effect in a particle system, the contact process.

The proof of Theorem 2.1 is based on the following argument, as pointed out by D. A. Dawson, to whom we are indebted for this and for several other enlightning discussions; see also ref. 5, where a similar approach is used to study critical fluctuations in a mean field model of interacting Brownians. Consider the one-dimensional analogue of our problem, namely

$$dm_\varepsilon = [m_\varepsilon^3 - (9/8)m_\varepsilon^5]\, dt + \varepsilon^{1/2}\, dw, \qquad m_\varepsilon(0) = 0 \tag{2.15}$$

The above equation is based on the assumption that the relevant fluctuations of the actual particle system with respect to the limiting equation (2.7) are essentially Brownian and have strength as in (2.15). Such an assumption is supported by the analysis of the magnetic fluctuation field

$$\varepsilon^{1/2} \sum_{x \in Z_\varepsilon} \sigma(x, t)$$

which converges to a Gaussian field whose covariance diverges proportionally to t (hence producing the same effect as $\varepsilon^{1/2}\, dw$); see refs. 7 and 16, where the stability of the hydrodynamic behavior is in general related to the fluctuation–dissipation theorem and to the asymptotic behavior of the fluctuation field. Notice that such a procedure, namely to add to the hydrodynamic equation a white noise with strength determined by the covariance of the fluctuation field, gives the correct prediction for the escape from the unstable equilibrium in the quadratic case. One final remark on this point: the assumption that the system is on a finite macroscopic volume plays here a fundamental role in the reduction to one degree of freedom; in fact, on the longer time scale we need to introduce to see the escape, the space structure, still present in (2.7), is completely lost. This greatly simplifies our proofs.

Anyway, let us take (2.15) for granted; then, by a scaling argument it easily follows that

$$x(t) = \lim_{\varepsilon \to 0} x_\varepsilon(t); \qquad x_\varepsilon(t) \equiv \varepsilon^{-1/4} m_\varepsilon(\varepsilon^{-1/2} t) \tag{2.16}$$

solves (2.14). One can also prove that the time for $x(t)$ to reach ∞ has the same limiting distribution when $\varepsilon \to 0$ as the the hitting time at $\pm m^*$ for the original process m_ε (times being properly renormalized).

Our study of the escape for the particle system mimics the above approach; we list below some of the main problems we have encountered and outline the strategy used to overcome them. The analogue of m_ε is the total magnetization

$$\bar{\sigma}_t = \varepsilon \sum_x \sigma(x, t) \tag{2.17}$$

As in (2.16), we renormalize it as

$$X_t^\varepsilon = \varepsilon^{-1/4} \bar{\sigma}_{\varepsilon^{-1/2} t} \tag{2.18}$$

We think of this as a process on $\mathscr{D}(R_+, R)$ (the path space of the jump process described by the variable X_t^ε) and we want to prove that its law $\mathscr{P}^\varepsilon$ converges to the law $\mathscr{P}$ of $x(t)$, as given by (2.14). What we actually need is something more, namely that the hitting times for X_t^ε converge to those of $x(t)$. More precisely, for all $R>0$ call $\tau(R)$ the first time when the absolute value of the canonical variable in $\mathscr{D}(R_+, R)$ equals or exceeds R. Then we choose R as a suitable function of ε which diverges when $\varepsilon \to 0$, and we want to prove that the law induced on this $\tau(R)$ by $\mathscr{P}^\varepsilon$ behaves like that induced by $\mathscr{P}$ when $\varepsilon \to 0$. Of course, one would like to choose R so that this implies that the time for the magnetization to become finite converges on the time scale $\varepsilon^{-1/2}$ to the explosion time in (2.14).

The main difficulty when trying to prove something like this is that X_t^ε does not satisfy a closed stochastic differential equation. In fact, when L^ε acts on X_t^ε (or on a function of X_t^ε), then the result is no longer a function of X_t^ε. One finds an expression involving product of spins at sites close to each other [cf. (2.3) and (2.4)]. To deal with this problem we have exploited the fact that we are in a (macroscopically) finite volume (having ε^{-1} sites) and that the main term in the generator is $\varepsilon^{-2}L_0$. This is in fact in some sense close to the generator of a diffusion, so that after times of the order of ε^{-a} (here a is any positive number), the process with generator $\varepsilon^{-2}L_0$ will *homogenize any initial local disturbance*. For this argument to be effective we certainly need the influence of the Glauber interaction to be negligible up to times of the order of ε^{-a}. In (2.8) we scale the values of $\bar{\sigma}_t$ by a factor $\varepsilon^{1/4}$. Then if we trust (2.7) as capable of describing the evolution of the magnetic field to such an accuracy ($\varepsilon^{1/4}$), we would predict that when $\bar{\sigma}_t$ reaches values of the order of $\varepsilon^{1/12}$ then the magnetization varies significantly over times of the order ε^{-a} (on the magnetization scale $\varepsilon^{1/4}$). For this reason we need to choose R, the hitting value for X_t^ε, much smaller than $\varepsilon^{1/12-1/4}$. To be definite, we fix once and for all the value of a equal to 10^{-6}. Then we choose R *only as large as* ε^{-10a} [our same techniques would allow for a larger value of R as $\varepsilon^{1/12-1/4+10a}$ (the value 10 in the previous expressions is not optimal). The proof would, however, be considerably longer].

Once we establish the convergence of the law of this hitting time we are only one small step closer to the end. In fact, the absolute value of the total magnetization $\bar{\sigma}_t$ at this time has only reached the value $\varepsilon^{1/4-10a}$. Let us go back for a moment to the example (2.15) and change the initial condition as $m_\varepsilon(0) = \pm\varepsilon^\eta$, $0<\eta<1/4$ ($\eta = 1/4 - n10a$, $n = 1, 2,\ldots$). The new normalization is then

$$x_\varepsilon(t) = \varepsilon^{-\eta} m_\varepsilon(\varepsilon^{-2\eta} t) \tag{2.19}$$

which converges in distribution when $\varepsilon \to 0$ to the solution of the equation

$$dx = x^3\, dt, \qquad x(0) = \pm 1 \tag{2.20}$$

We proceed analogously in our particle model. We set $\eta = \varepsilon^{1/4-10a}$ and define (counting times starting from the previous hitting time at $\varepsilon^{1/4-10a}$, for notational simplicity)

$$X_t^{\varepsilon,\eta} = \varepsilon^{-\eta} \bar{\sigma}_{\varepsilon^{-2\eta}t} \tag{2.21}$$

Call $\mathscr{P}^{\varepsilon,\eta}$ the law induced by $X_t^{\varepsilon,\eta}$. The strategy is again to prove that $\mathscr{P}^{\varepsilon,\eta}$ induces on $\tau(\varepsilon^{-10a})$ a law which in the limit is supported by a single value, the explosion time for $x(t)$ solution to (2.20). At this time the magnetization has value $\varepsilon^{1/4-20a}$, we set then $\eta = \varepsilon^{1/4-20a}$ and start again. We repeat this procedure till $|\bar{\sigma}_t| \geqslant \varepsilon^{100a}$. When the magnetization reaches this value it is *so close* to a finite nonzero value that its further evolution can be studied almost explicitly and in this way we finally reach the value $\pm m^*$. The total escape time is then the sum of all the above hitting times; they are finitely many and only the first one contributes effectively to the sum in the limit as $\varepsilon \to 0$, because of the different normalizations.

3. THE FIRST STAGE OF THE ESCAPE

Following the strategy outlined at the end of the previous section in this one, we start establishing the main ingredients for proving that the first time when $|\bar{\sigma}_t| \geqslant \varepsilon^{100a}$ has the same law when $\varepsilon \to 0$ as the explosion time in (2.14). To complete the proof of this statement, we need some properties on the typical configurations of the process which will be established in Section 4. We also postpone to the Appendix C the proof of some technical estimates that we shall use in this section.

First some notation. We set $a = 10^{-6}$ and

$$H = \{\eta = 1/4 - n10a\colon \quad 0 \leqslant n < n^*,\ 1/4 - n^*10a = 100a\} \tag{3.1}$$

For each $\eta \in H$ we define $X_t^{\varepsilon,\eta}$ as in (2.21); $\mathscr{P}^{\varepsilon,\eta}$ and $\tau(R)$ are also defined at the end of Section 2. To state the main result in this section, we need to specify the initial spin configuration corresponding to the initial value $X_0^{\varepsilon,\eta}$. Given $\eta \in H$, we shall say that the family $\nu^{\varepsilon,\eta}$ is *allowed* if it equals μ^ε when $\eta = 1/4$, while for $\eta < 1/4$, (i) there is $\gamma^{\varepsilon,\eta} = \pm 1$ such that $|X_0^{\varepsilon,\eta} - \gamma^{\varepsilon,\eta}| \leqslant \varepsilon^{1-\eta}$, almost surely; and (ii)

$$\lim_{\varepsilon \to 0} \nu^{\varepsilon,\eta}(\|\sigma\| > \varepsilon^{\eta - 2a}) = 0 \tag{3.2}$$

where $\|\sigma\|$ is defined in (4.1) [roughly speaking, this norm is related to the maximum of the absolute value of the average magnetization over all possible intervals of Z_ε with length of the order of $\varepsilon^{-1+1/5}$ (the choice of these numbers is not optimal; we need intervals infinitesimal on the macroscopic scale ε^{-1}, yet very large microscopically)]. Therefore (3.2) ensures that *local values of the magnetization* cannot exceed by too much the total average magnetization.

Theorem 3.1. Let $\eta \in H$ and let $\nu^{\varepsilon,\eta}$ be an allowed state. For all $R>0$ denote by $\mathscr{P}_R^{\varepsilon,\eta}$ the law $\mathscr{P}^{\varepsilon,\eta}$ for the process stopped at $\tau(R)$. If $\eta=1/4$, then $\mathscr{P}_R^{\varepsilon,\eta}$ converges to $\mathscr{P}_R$, the law of the process (2.14) stopped when it reaches R. If $\eta<1/4$, then $\mathscr{P}_R^{\varepsilon,\eta}$ converges to the law supported by the deterministic trajectory $x(t)$ solution to (2.20), with initial datum $=\gamma^{\varepsilon,\eta}$, as $\varepsilon \to 0$ (cf. the definition of $\nu^{\varepsilon,\eta}$) and stopped when its absolute value reaches R.

Finally, the law of $\tau(\varepsilon^{-10a})$ under $\mathscr{P}^{\varepsilon,\eta}$ converges to the distribution of the explosion time for the process (2.14) if $\eta=1/4$, and to the explosion time for (2.20) if $\eta<1/4$.

Remarks. When $\eta=1/4$ the limiting law of $X_t^{\varepsilon,\eta}$ is described by a nonlinear stochastic differential equation. This is an example of a way to approximate stochastic differential equations by discrete particle models. A severe drawback to our results (from this point of view) is the limitation to finite volumes, an assumption that we used in an essential way. The analogous result in an infinite volume would yield a stochastic PDE. Even more interesting and intriguing is the two-dimensional case, as pointed out to us by Jona-Lasinio, to whom we are indebted for many enlightening discussions on this aspect of the problem. The model in the two-dimensional case in fact should provide a discrete approximation for stochastic quantization being somehow related to the generalized stochastic differential equation introduced in ref. 14 [here we think of a potential $V(m)=-\lambda m^4$, $\lambda>0$, obtained, for instance, by choosing $\gamma=1/2$ in (2.9a)].

The remainder of this section is devoted to the proof of Theorem 3.1 obtained via several intermediate steps.

3.1. A Change of Variables

In this first step we compactify the space, introducing the variable

$$Y_t^{\varepsilon,\eta}=\operatorname{arctg}([X_t^{\varepsilon,\eta}]^3) \tag{3.3}$$

The new process lives in a different space, $\mathscr{D}(R_+, [-\pi/2, \pi/2])$. However, for notational simplicity its law will be denoted by the same symbol $\mathscr{P}^{\varepsilon,\eta}$ as for the process $X_t^{\varepsilon,\eta}$. Furthermore, the stopping times $\tau(R)$ introduced before when employed in this new space denote the stopping times at $\operatorname{arctg}(R^3)$. Finally, $T(\phi)$ denotes the first time when the absolute value of the canonical variable in $\mathscr{D}(R_+, [-\pi/2, \pi/2])$ reaches ϕ.

3.2. The Martingale Problem

To study the convergence problem, we shall use the Stroock–Varadhan martingale theory.[19] Let F be any *smooth* function on $[-\pi/2, \pi/2]$; then

$$F(Y_t^{\varepsilon,\eta}) - F(Y_0^{\varepsilon,\eta}) - \int_0^t ds\, \varepsilon^{-2\eta} L^\varepsilon F(Y_s^{\varepsilon,\eta}) = M_t^{\varepsilon,\eta}$$

where $M_t^{\varepsilon,\eta}$ is a martingale. Since $L_0 F = 0$ (because the stirring does not change the total magnetization), only the Glauber generator contributes to the above expression. After a Taylor expansion we get

$$F(Y_t^{\varepsilon,\eta}) - F(Y_0^{\varepsilon,\eta}) - \int_0^t ds\, F'(Y_s^{\varepsilon,\eta})\, A_1^{\varepsilon,\eta}(s)$$

$$-\tfrac{1}{2}\int_0^t ds\, F''(Y_s^{\varepsilon,\eta})\, A_2^{\varepsilon,\eta}(s) + R_t^{\varepsilon,\eta} = M_t^{\varepsilon,\eta} \tag{3.4}$$

$$A_1^{\varepsilon,\eta}(s) = f'(X_s^{\varepsilon,\eta})\, \gamma_1^{\varepsilon,\eta}(s) + \tfrac{1}{2} f''(X_s^{\varepsilon,\eta})\, \gamma_2^{\varepsilon,\eta}(s) \tag{3.5a}$$

$$A_2^{\varepsilon,\eta}(s) = f'(X_s^{\varepsilon,\eta})^2\, \gamma_2^{\varepsilon,\eta}(s) \tag{3.5b}$$

where F' and F'' are the first and the second derivatives of F, while f' and f'' are the first and second derivative of the function $f(\cdot) = \text{arctg}(\cdot)^3$. Here $R_t^{\varepsilon,\eta}$ is the remainder term in the Taylor expansion; it vanishes in sup norm when $\varepsilon \to 0$. Finally,

$$\gamma_1^{\varepsilon,\eta}(s) = \varepsilon^{1-3\eta} \sum_y \Bigg[\prod_{j=-1}^{1} \sigma(y+j) - \tfrac{3}{8} \prod_{j=2}^{4} \sigma(y+j)\{4\sigma(y)[\sigma(y+1)+\sigma(y-1)] - \sigma(y+1)\,\sigma(y-1)\}\Bigg] \tag{3.6}$$

$$\gamma_2^{\varepsilon,\eta}(s) = \varepsilon^{2-4\eta} \sum_y \Bigg[1 - \sigma(y)\,\sigma(y+1) + \tfrac{1}{4}\sigma(y)\,\sigma(y+2) + \tfrac{3}{8} \prod_{i=0}^{3} \sigma(y+i) - \tfrac{3}{4}\sigma(y) \prod_{i=2}^{4} \sigma(y+i) + \tfrac{3}{8}\sigma(y) \prod_{i=3}^{5} \sigma(y+i) - \tfrac{3}{16} \prod_{i=0}^{5} \sigma(y+i)\Bigg] \tag{3.7}$$

and σ in the two equations above is evaluated at the time $\varepsilon^{-2\eta}s$.

Notice that the second integral in (3.4) is vanishingly small when $\varepsilon \to 0$ if $\eta < 1/4$; this follows from (3.7). As a consequence, the limiting process will be deterministic, in contrast to the case when $\eta = 1/4$.

As usual with martingale problems, we first need to prove tightness and then to identify the limiting points by proving that they satisfy a martingale equation which has unique solution. Tightness is the most serious problem in our case. A sufficient condition for tightness comes from a uniform L^2 estimate for both $\gamma_i^{\varepsilon,\eta}$, $i=1,2$. While this is true for $\gamma_2^{\varepsilon,\eta}$, for all $\eta \in H$, this is not the case for $\gamma_1^{\varepsilon,\eta}$. In fact, if we square it and consider the contribution of the diagonal terms, we get a factor $\varepsilon^{1-6\eta}$, which diverges for the higher values of η. We shall overcome this problem by taking suitable time averages of γ_1, as we shall see later.

3.3. The Modified Process

Instead of stopping the process at $\tau(\varepsilon^{-10a})$, it is convenient to suitably redefine it after this time. It is also convenient to introduce a new stopping time rather than the above one. We set in fact

$$\tau^{\varepsilon,\eta} = \inf\{s_n = n\varepsilon^{-a+2\eta}\colon \quad |X_{s_n}^{\varepsilon,\eta}| \geq \varepsilon^{-10a}\} \tag{3.8}$$

(namely we introduce a tie grid ε^{-a} for the original process, hence $\varepsilon^{-a+2\eta}$ for the $X^{\varepsilon,\eta}$ process, and we look at the first time on this grid when the $X^{\varepsilon,\eta}$ process reaches or exceeds the value ε^{-10a}). We then define a process $Z_t^{\varepsilon,\eta}$ which is identical to $Y_t^{\varepsilon,\eta}$ for $t \leq \tau^{\varepsilon,\eta}$ while its value afterward is determined by (2.14) [or (2.20) according to the value of η] with initial value $x(\tau^{\varepsilon,\eta}) = [\mathrm{tg}(Y^{\varepsilon,\eta})]^{1/3}$ (Y being evaluated at $\tau^{\varepsilon,\eta}$). Then, for these values of t, $Z_t^{\varepsilon,\eta} = \mathrm{arctg}[x(t)]^3$. Such a process is then stopped when it reaches $\pm\pi/2$. We call $\mathcal{Q}^{\varepsilon,\eta}$ its law and we want to prove that, like $\mathcal{P}^{\varepsilon,\eta}$, it converges to $\mathcal{P}^\eta$, the law of $\mathrm{arctg}[x(t)^3]$, where $x(t)$ solves (2.14) or (2.20) according to the value of η.

3.4. A Statement Equivalent to Theorem 3.1

Since the following discussion works for all η, we omit writing it explicitly. *If we assume that* $\mathcal{Q}^\varepsilon \to P$, then we can easily deduce that $\tau(\varepsilon^{-10a}) \to \tau$ [where τ denotes the explosion time for (2.14) or (2.20), according to cases]. *In this way we shall reduce the proof of the theorem to only showing the above assumed convergence.*

To prove this statement, we argue as follows. Call $T(\pi/2)$ the time when the absolute value of the canonical variable in $\mathcal{D}(R_+, [-\pi/2, \pi/2])$ equals $\pi/2$ (see the notation introduced in Section 3.1). Then by the above assumption the law of $T(\pi/2)$ under $\mathcal{Q}^\varepsilon$ approaches that of τ, the explosion time for (2.14) or (2.20) according to cases. Calling $T(\phi)$ the stopping time at ϕ of the absolute value of the canonical variable in $\mathcal{D}(R_+, [-\pi/2, \pi/2])$, then, still by assumption, its law under $\mathcal{Q}^\varepsilon$ approaches that under $\mathcal{P}$.

Then, since (i) for $\phi \neq \pi/2$,

$$T(\phi) \leqslant \tau(\varepsilon^{-10a}) \leqslant \tau^{\varepsilon,\eta} \leqslant T(\pi/2)$$

and (ii) the distribution under $\mathscr{P}$ of $T(\phi)$ converges to that of $T(\pi/2)$ (of course the same as the law of τ), then, by using the previous considerations on the convergence of the distribution of $T(\phi)$ and $T(\pi/2)$ when $\varepsilon \to 0$, we deduce that the distribution of $\tau(\varepsilon^{-10a})$ under $\mathscr{Q}^\varepsilon$, identical to that under $\mathscr{P}^\varepsilon$, by the second inequality above, converges to the distribution of τ.

3.5. Tightness

Since the processes defined by (2.14) and (2.20) are tight, tightness of $Z_t^{\varepsilon,\eta}$ easily follows from the tightness of $Y^{\varepsilon,\eta}(t \wedge \tau^{\varepsilon,\eta})$), i.e., the same as the process Z stopped at $\tau^{\varepsilon,\eta}$.

Note: In the equations below we omit the superscript η to simplify notation; all expressions depend on η even though η does not appear explicitly.

Writing (3.4) with $F(x) = x$, we have that

$$Y_t^\varepsilon - Y_0^\varepsilon = \int_0^{t \wedge \tau^\varepsilon} ds\, f'(X_s^\varepsilon)\, \gamma_1^\varepsilon(s) + \tfrac{1}{2} \int_0^{t \wedge \tau^\varepsilon} ds\, f''(X_s^\varepsilon)\, \gamma_2^\varepsilon(s) - R_{t \wedge \tau^\varepsilon}^\varepsilon + M_{t \wedge \tau^\varepsilon}^\varepsilon$$

where $M_{t \wedge \tau^\varepsilon}^\varepsilon$ is a martingale, γ_2^ε is a uniformly bounded function, and R^ε converges to 0 in uniform norm when $\varepsilon \to 0$, uniformly on t in the compacts.

Therefore, tightness for $Y^\varepsilon(t \wedge \tau^\varepsilon)$ is a consequence of the tightness of $M_{t \wedge \tau^\varepsilon}^\varepsilon$ and $\Gamma^\varepsilon(t \wedge \tau^\varepsilon)$, where

$$\Gamma^\varepsilon(t) = \int_0^{t \wedge \tau^\varepsilon} ds\, f'(X_s^\varepsilon)\, \gamma_1^\varepsilon(s) \tag{3.9}$$

Since γ_2^ε is uniformly bounded, the tightness of the martingales $M_{t \wedge \tau^\varepsilon}^\varepsilon$ easily follows from Lemma 2.6 of ref. 12.

To prove the tightness of $\Gamma^\varepsilon(t \wedge \tau^\varepsilon)$, we use the Chensov moment condition[2]; thus, we need to show that for all $T < \infty$ there are positive constants c, $\gamma > 0$, and $\sigma > 1$ such that for all $s < t < T$,

$$E\left[\left|\int_{s \wedge \tau^\varepsilon}^{t \wedge \tau^\varepsilon} ds\, f'(X_s^\varepsilon)\, \gamma_1^\varepsilon(s)\right|^\gamma\right] \leqslant c\,|t-s|^\sigma \tag{3.10}$$

where E here and in the following denotes the expectation with respect to the process starting from ν^ε (cf. Theorem 3.1). As already mentioned, if γ_1^ε

is L^2, then the above condition is trivially satisfied. In fact, we can set $\gamma=2$ in (3.10) and get $\sigma=2$ (after using the Cauchy–Schwarz inequality). However, if η is close enough to 1/4, γ_1^ε is not in L^2 and we have to proceed more carefully. The proof we shall give does apply to all values of η; however, for the above reason it can be simplified if η is not too close to 1/4.

We are going to prove (3.10) with $\gamma=2$ and $\sigma=9/8$. Let $s_k=k\varepsilon^{2\eta-a}$, $k\geqslant 0$. Then if $|t-s|\geqslant\varepsilon^{2\eta-a}$, there are n and n' such that $s_{n'-1}<s<s_{n'}$ and $s_{n-1}<t\leqslant s_n$. Therefore, using the Cauchy–Schwarz inequality, we obtain

$$\begin{aligned}E[(\Gamma^\varepsilon(t)-\Gamma^\varepsilon(s))^2]\leqslant(n-n'+2)\Bigg\{&E[(\Gamma^\varepsilon(t\wedge\tau^\varepsilon)-\Gamma^\varepsilon(s_{n-1}\wedge\tau^\varepsilon))^2]\\&+\sum_{k=n'+1}^{n-1}E[(\Gamma^\varepsilon(s_k\wedge\tau^\varepsilon)-\Gamma^\varepsilon(s_{k-1}\wedge\tau^\varepsilon))^2]\\&+E[(\Gamma^\varepsilon(s_{n'}\wedge\tau^\varepsilon)-\Gamma^\varepsilon(s\wedge\tau^\varepsilon))^2]\Bigg\}\end{aligned}\tag{3.11}$$

First observe that by the definition of τ^ε

$$\begin{aligned}&(\Gamma^\varepsilon(s_k\wedge\tau^\varepsilon)-\Gamma^\varepsilon(s_{k-1}\wedge\tau^\varepsilon))^2\\&\quad=(\Gamma^\varepsilon(s_k\wedge\tau^\varepsilon)-\Gamma^\varepsilon(s^{k-1}\wedge\tau^\varepsilon))^2\,1(\tau^\varepsilon>s_{k-1})\end{aligned}$$

Hence

$$\begin{aligned}&(\Gamma^\varepsilon(s_k\wedge\tau^\varepsilon)-\Gamma^\varepsilon((s_{k-1}\wedge\tau^\varepsilon))^2\\&\quad\leqslant(\Gamma^\varepsilon(s_k\wedge\tau^\varepsilon)-\Gamma^\varepsilon(s_{k-1}\wedge\tau^\varepsilon))^2\,1(\tau^\varepsilon\geqslant s_{k-2})\\&\quad\leqslant 2\int_{s_{k-1}}^{s_k}du\int_{s_{k-1}}^{u}du'\,E[f'(X_u^\varepsilon)\,\gamma_1^\varepsilon(u)\,f'(X_{u'}^\varepsilon)\,\gamma_2^\varepsilon(u')\,1(\tau^\varepsilon\geqslant s_{k-2})]\end{aligned}\tag{3.12}$$

To proceed we need an estimate for the expectation in the last integral. This cannot be uniform with respect to the integration times, since we have already seen that γ_1^ε is not in L^2 (for η close enough to 1/4). The next lemma gives the required estimate. Its proof, given in Appendix C, is based on some good *mixing properties* of the stirring process.

Lemma 3.2. Let $\eta\in H$ and let $t>s$ be such that $t-s<\varepsilon^{2\eta-a}$. Then there is a constant c independent of ε, t, and s such that

$$\left|\int_s^t ds'\,E[f'(X_t^{\varepsilon,\eta})\,\gamma_1^{\varepsilon,\eta}(t)\,f'(X_{s'}^{\varepsilon,\eta})\,\gamma_1^{\varepsilon,\eta}(s')\,1(\tau^{\varepsilon,\eta}>s-\varepsilon^{2\eta-a})]\right|\leqslant c\varepsilon^{2\eta-a}\tag{3.13}$$

Notice that rewriting the time integral using nonrescaled times, i.e., those of the original process, we get an integral over times of the order of ε^{-a}, the same for all values of η.

3.6. Tightness (Continued)

By (3.13) there is a constant c such that

$$\sum_{k=n'+1}^{n} E[(\Gamma^{\varepsilon}(s_k \wedge \tau^{\varepsilon}) - \Gamma^{\varepsilon}(s_{k-1} \wedge \tau^{\varepsilon}))^2] \leqslant c(n-n')\varepsilon^{4\eta-2a} \quad (3.14)$$

so that the contribution of the second sum in the right-hand side of (3.11) is bounded by $c|t-s|^2$; in fact, $|t-s| \geqslant (n-n'+2)\varepsilon^{2\eta-a}$. On the other hand, the first and the third terms on the right-hand side of (3.11) are estimated using Lemma 3.1, yielding a bound $c(n-n'+2)\varepsilon^{4\eta-2a} \leqslant c|t-s|^2$.

In conclusion, we have so far proven the estimate necessary for applying the Chensov criterion with $\gamma = \sigma = 2$, but only for $|t-s| \geqslant \varepsilon^{2\eta-a}$. Let us now consider the case $|t-s| < \varepsilon^{2\eta-a}$. There are two possibilities: (a) both s and t are in the same interval $[s_{n-1}, s_n]$ or (b) they belong to two next intervals, say $(s_{n-2}, s_{n-1}]$ and $(s_{n-1}, s_n]$. We can reduce case (b) to (a) because

$$\begin{aligned} &E[(\Gamma^{\varepsilon}(t \wedge \tau^{\varepsilon}) - \Gamma^{\varepsilon}(s \wedge \tau^{\varepsilon}))^2] \\ &\quad \leqslant 2\{E[(\Gamma^{\varepsilon}(t \wedge \tau^{\varepsilon}) - \Gamma^{\varepsilon}(s_{n-1} \wedge \tau^{\varepsilon}))^2] \\ &\qquad + E[(\Gamma^{\varepsilon}(s_{n-1} \wedge \tau^{\varepsilon}) - \Gamma^{\varepsilon}(s \wedge \tau^{\varepsilon}))^2]\} \end{aligned}$$

so we consider case (a). We set $\theta = 8\eta - a$ and use Lemma 3.2 as well as the bound $|\gamma_1^{\varepsilon}(s)\gamma_1^{\varepsilon}(s')| \leqslant c\varepsilon^{-6\eta}$ [easily derived from (3.6)] to get

$$\begin{aligned} &E[(\Gamma^{\varepsilon}(t \wedge \tau^{\varepsilon}) - \Gamma^{\varepsilon}(s \wedge \tau^{\varepsilon}))^2] \\ &\quad = E[(\Gamma^{\varepsilon}(t \wedge \tau^{\varepsilon}) - \Gamma^{\varepsilon}(s \wedge \tau^{\varepsilon}))^2]\, 1(\varepsilon^{\theta} < |t-s| \leqslant \varepsilon^{2\eta-a}) \\ &\qquad + E[(\Gamma^{\varepsilon}(t \wedge \tau^{\varepsilon}) - \Gamma^{\varepsilon}(s \wedge \tau^{\varepsilon}))^2]\, 1(|t-s| \leqslant \varepsilon^{\theta}) \\ &\quad \leqslant c[\varepsilon^{2\eta-a}|t-s|\, 1(\varepsilon^{\theta} < |t-s| \leqslant \varepsilon^{2\eta-a}) + \varepsilon^{-6\eta}|t-s|^2\, 1(|t-s| \leqslant \varepsilon^{\theta})] \\ &\quad \leqslant c[|t-s|^{1+(2\eta-a)\theta^{-1}} + |t-s|^{2-6\eta\theta^{-1}}] \\ &\quad \leqslant c|t-s|^{(10\eta-2a)\theta^{-1}} \leqslant c|t-s|^{9/8} \end{aligned}$$

which completes the proof of the tightness.

3.7. The Limiting Process

By tightness the law $\mathscr{Q}^{\varepsilon,\eta}$ of the process $Z^{\varepsilon,\eta}$ converges by subsequences. We shall use this fact to show that any limiting point satisfies a martingale relation. We shall then prove that this martingale relation uniquely defines a process which is the law of the solution to (2.14) or (2.20) according to the value of η; to draw this conclusion, we need to know that any limiting point has support on $C(R_+, [-\pi/2, \pi/2])$, but this follows from the fact that the jumps of $X^{\varepsilon,\eta}$ are $\pm\varepsilon^{1-\eta}$.

We shall use the martingale characterization of the diffusion processes on compact sets (ref. 13, Chapter IV, pp. 208–214) to reduce the whole problem to the proof of the following statement. For all functions $F, \Psi_1,..., \Psi_k$ in C^∞ and with compact support, and for all $0 \leqslant t_1 < \cdots < t_k < s < t$,

$$\lim_{\varepsilon \to 0} \left| E[F(Z_t^{\varepsilon,\eta}) - F(Z_s^{\varepsilon,\eta})] - \int_{s \wedge T(\pi/2)}^{t \wedge T(\pi/2)} ds' \, (LF)(Z_{s'}^{\varepsilon,\eta}) \prod_{j=1}^{k} \Psi_j(Z_{t_j}^{\varepsilon,\eta}) \right| = 0 \tag{3.15}$$

where $T(\pi/2)$ is the hitting time at $\pi/2$ and L is the following operator:

$$(LF)(y) = \frac{dF}{dy}(y)[f'(\text{tg } y) \text{ tg } y + \tfrac{1}{2} f''(\text{tg } y)] + \frac{d^2F}{dy^2}(y)[f'(\text{tg } y)]^2$$

if $|y| < \pi/2$ and 0 otherwise; recall that f' and f'' are, respectively, the first and the second derivatives of the function $\text{arctg}(\cdot)$.

Since by its definition Z_t^ε (*note*: *we keep omitting the dependence on* η *to simplify notation*) is given by the solution to (2.14) or (2.20) according to the value of η (see Section 3.3); then by Theorem 7.2 and the related construction on p. 214 of ref. 13, it is enough to prove that

$$\lim_{\varepsilon \to 0} \Bigg| E[F(Z_{t \wedge \tau^\varepsilon}^\varepsilon) - F(Z_{s \wedge \tau^\varepsilon}^\varepsilon)] - E\left[\int_{s \wedge \tau^\varepsilon}^{t \wedge \tau^\varepsilon} ds' \, (LF)(Z_{s'}^\varepsilon) \prod_{j=1}^{k} \Psi_j(Z_{t_j}^\varepsilon)\right] \Bigg| = 0 \tag{3.16}$$

We set [cf. (3.4)]

$$N_t^\varepsilon = F(Z_{t \wedge \tau^\varepsilon}^\varepsilon) - F(Z_{s \wedge \tau^\varepsilon}^\varepsilon) - \int_{s \wedge \tau^\varepsilon}^{t \wedge \tau^\varepsilon} ds' \, [F'(Y_{s'}^\varepsilon) \, A_1^\varepsilon(s') + \tfrac{1}{2} F''(Y_{s'}^\varepsilon) \, A_2^\varepsilon(s')] + R_{t \wedge \tau^\varepsilon}^\varepsilon - R_{s \wedge \tau^\varepsilon}^\varepsilon \tag{3.17}$$

Then, by (3.4), N_t^ε is a martingale.

Thus, (3.16) will easily follow from (3.17) once we show that

$$\lim_{\varepsilon \to 0} \int_s^t ds' \, E\left[F_1(X_{s'}^\varepsilon)[\gamma_1^\varepsilon(s') - (X_{s'}^\varepsilon)^3] \, 1(\tau^\varepsilon > s') \prod_{j=1}^k \Psi_j \right] = 0 \tag{3.18}$$

and

$$\lim_{\varepsilon \to 0} \int_s^t ds' \, E\left[F_2(X_{s'}^\varepsilon)[\gamma_2^\varepsilon(s') - v^\eta] \prod_{j=1}^k \Psi_j \right] = 0 \tag{3.19}$$

where v^η equals 1 if $\eta = 1/4$ and 0 otherwise. In (3.18) and (3.19) we have used the short-hand notation

$$\prod_{j=1}^k \Psi_j \equiv \prod_{j=1}^k \Psi_j(Z_{t_j}^\varepsilon)$$

and

$$F_i(x) = \frac{\partial^i}{\partial x^i} F(\operatorname{arctg} x^3)$$

for $i = 1, 2.$.

We start by proving (3.18). As in the proof of tightness, we split the time integral between s and t into a sum of integrals over intervals $(s_k, s_{k+1}]$ of length $\varepsilon^{2\eta - a}$. We apply twice the Cauchy–Schwarz inequality and we get

$$\begin{aligned}
&\left| E\left[\int_s^t ds' \, F_1(X_{s'}^\varepsilon)\{\gamma_1^\varepsilon(s') - (X_{s'}^\varepsilon)^3\} \, 1(\tau^\varepsilon > s') \prod_{j=1}^k \Psi_j \right] \right|^2 \\
&\quad \leqslant (n - n' + 2) \left(E\left[\left\{ \int_{s_{n-1}}^t ds' \, F_1(X_{s'}^\varepsilon)[\gamma_1^\varepsilon(s') - (X_{s'}^\varepsilon)^3] \right\}^2 \right.\right. \\
&\qquad \left. \times 1(\tau^\varepsilon > s_{n-2}) \prod_{j=1}^k \Psi_j \right] \\
&\qquad + E\left[\left\{ \int_s^{s_{n'}} ds' \, F_1(X_{s'}^\varepsilon)[\gamma_1^\varepsilon(s') - (X^\varepsilon)^3] \right\}^2 1(\tau^\varepsilon > s_{n'-1}) \prod_{j=1}^k \Psi_j \right] \\
&\qquad \left. + \sum_{k=n'}^{n-2} E\left[\left\{ \int_{s_k}^{s_{k+1}} ds' \, F_1(X_{s'}^\varepsilon)[\gamma_1^\varepsilon(s') - (X_{s'}^\varepsilon)^3] \right\}^2 1(\tau^\varepsilon > s_{k-1}) \prod_{j=1}^k \Psi_j \right] \right)
\end{aligned} \tag{3.20}$$

where we have used the same notation as in (3.11).

We estimate the first two terms on the right-hand side of (3.20) by

$$-c(n - n' + 2)\varepsilon^{2(2\eta - a)} + \varepsilon^{(2\eta - a)9/8}$$

We have used Cauchy–Schwarz to expand the square, noticing that $|F_1(X^\varepsilon_{s'})(X^\varepsilon_{s'})^3|$ is uniformly bounded and that those terms which contain γ^ε_1 reconstruct the function Γ^ε, so that we can apply (3.10) with $\gamma = 2$ and $\sigma = 9/8$. For the third term in (3.20) we use the following lemma, proven in Appendix C together with Lemma 3.2.

Lemma 3.3. Let $F \in C^\infty$ and have compact support, and let F_1 be as above. Let t and s be as in Lemma 3.2. Then there is $\beta > 0$ and a constant c so that

$$\left| E\left[\int_s^t ds'\, F_1(X^\varepsilon_{s'})[\gamma^\varepsilon_1(s') - (X^\varepsilon_{s'})^3]\, F_1(X^\varepsilon_t)[\gamma^\varepsilon_1(t) - (X^\varepsilon_t)^3]\, 1(\tau^\varepsilon > s - \varepsilon^{2\eta - a})\right]\right|$$
$$\leqslant c\varepsilon^\beta \varepsilon^{2\eta - a} \tag{3.21}$$

3.8. The Limiting Process (Continued)

Since $|t - s|$ is bounded by assumption, the same argument we have used to prove tightness shows that the third term on the right-hand side of (3.20) is bounded by ε^β.

We are left with the proof of (3.19), which is obvious if $\eta < 1/4$. If $\eta = 1/4$, we use the fact that F_2 and the Ψ's are bounded, so that by the Cauchy–Schwarz inequality after making explicit the difference $\gamma^\varepsilon_2 - 1$, we are reduced to considering a term of the form

$$E\left[\left(\varepsilon \sum_x h(x, t)\right)^2\right] \tag{3.22}$$

where the expectation is with respect to the process starting from an allowed measure $\nu^{\varepsilon,\eta}$ with $\eta = 1/4$ and $t = \varepsilon^{-a}$; here the time is not renormalized; it refers to the process with generator L^ε. Finally, $h(x, t)$ stands for the x-shifted function h on $\{-1, 1\}^{Z_\varepsilon}$ evaluated at time t. The function h in turn is a finite sum of products of $\sigma(x)$'s such that in each of these terms there is at least one σ. In Appendix C we prove that (3.22) vanishes when $\varepsilon \to 0$ [the proof being similar to those of Lemmas 3.2 and 3.3], so that Theorem 3.1 is proven.

4. TYPICAL CONFIGURATIONS AND TRAJECTORIES

In this section we complete the proof of Theorem 2.1, modulo a few estimates of more technical nature which will be proven in Appendix B.

There are essentially two problems for proving Theorem 2.1 after Theorem 3.1 has been established. In fact, from Theorem 3.1 we get an explicit estimate on the time it takes for the total magnetization $\bar{\sigma}_t$ to increase from the value ε^{η} to $\varepsilon^{\eta - 10a}$, $\eta \in H$ (see the beginning of Section 3 for notation). However, this result is based on the assumption that the distribution of the spins at the time when $\bar{\sigma}_t$ reaches ε^{η} is *allowed*, in the sense specified at the beginning of Section 3. To prove this, we shall study and characterize the typical trajectories of the process for time intervals of the order ε^{-a}, i.e., these are infinitely long times when $\varepsilon \to 0$, yet infinitesimal on the time scale when the escape is observed.

The second problem concerns the final step of the escape, which was not at all considered in Theorem 3.1, namely when the magnetization increases past ε^{100a} to reach one of the two equilibrium values $\pm m^*$. The answer to this question is again based on a good characterization of the typical trajectories of the process which allows us to establish in a very strong form the propagation-of-chaos property and to prove that the magnetization, after reaching the value ε^{100a}, follows closely the deterministic equation (2.7).

The key ingredient in our analysis is the estimate (4.6) below, whose proof is given in Appendix A. The estimate is valid for any initial configuration and it does not refer specifically to the problem of the escape from the unstable equilibrium. Given a configuration $\sigma \in \{-1, 1\}^{Z_\varepsilon}$, we define $m_{\varepsilon,t} \equiv \{m_\varepsilon(x, t; \sigma), t \geqslant 0, x \in Z_\varepsilon\}$ as follows:

$$m_\varepsilon(x, t; \sigma) = \sum_z P_t^\varepsilon(x \to z)\, \sigma(z) + \int_0^t ds \sum_z P_{t-s}^\varepsilon(x \to z)\, g(z, m_{\varepsilon,s}) \tag{4.1}$$

where P_t^ε is the transition probability of a simple symmetric random walk which jumps on its nearest neighbors (in Z_ε) with intensity ε^{-2}; g is given by [we write below $m_\varepsilon(x, t)$ instead of $m_\varepsilon(x, t; \sigma)$ for notational simplicity]

$$\begin{aligned} g(x, m_{\varepsilon,t}) = {} & [m_\varepsilon(x+1, t) + m_\varepsilon(x-1, t) - 2m_\varepsilon(x, t)] \\ & - \tfrac{1}{2} \prod_{j=-1}^{1} m_\varepsilon(x+j, t) + \tfrac{3}{2} \prod_{j=2}^{4} m_\varepsilon(x+j, t) \\ & - \tfrac{3}{4} \prod_{j=2}^{4} m_\varepsilon(x+j, t) \{ m_\varepsilon(x, t)[m_\varepsilon(x+1, t) - m_\varepsilon(x-1, t)] \\ & - \tfrac{1}{2} m_\varepsilon(x+1, t)\, m_\varepsilon(x-1, t) \} \end{aligned} \tag{4.2}$$

Therefore, $m_{\varepsilon,t}$ is the solution to a discretized version of (2.7). Indeed, one expects that the typical trajectories of our process are somehow related to $m_{\varepsilon,t}$. Of course, it does not make sense to compare $m_{\varepsilon,t}$ and σ_t (the random spin configuration at time t) site by site [since $\sigma_t(x) = \pm 1$]; one should rather compare suitable space averages of the two quantities. This is done in some convenient way for our purposes by introducing the following quantity. For any real-valued function f on $\{-1, 1\}^{Z_\varepsilon}$ we set

$$\|f\| = \sup_x \left| \sum_z P^\varepsilon_{\varepsilon^{2/5}}(x \to z)\, f(z) \right| \tag{4.3}$$

(the choice of the time $\varepsilon^{2/5}$ is not optimal). Roughly speaking, for any fixed x the above is an average of f over an interval centered at x and having length of the order of $\cong [\varepsilon^{-2}\varepsilon^{2/5}]^{1/2}$, an interval which becomes infinite when $\varepsilon \to 0$ but which is still infinitesimal if measured on a macroscopic scale, $\cong \varepsilon^{-1}$. Using (4.3), we can then compare the configurations of the process at time t and $m_{\varepsilon,t}$ by estimating the probability that $\{\|\sigma_t - m_{\varepsilon,t}\| \geqslant \varepsilon^\zeta\}$, where $\zeta > 0$ has to be suitably fixed. An upper bound for such a probability can be derived using the Chebyshev inequality with power $n = 2k$, $k \geqslant 1$. Denoting by E_σ the expectation of the Stirring + Glauber process starting from σ and by P_σ its law, we have that

$$\begin{aligned} &P_\sigma(\|\sigma_t - m_{\varepsilon,t}\| \geqslant \varepsilon^\zeta) \\ &\quad \leqslant \varepsilon^{-1} \sup_x P_\sigma\left(\left| \sum_z P^\varepsilon_{\varepsilon^{2/5}}(x \to z)[\sigma(z, t) - m_\varepsilon(z, t; \sigma)] \right| > \varepsilon^\zeta \right) \\ &\quad \leqslant \varepsilon^{-1} \sup_x \varepsilon^{-\zeta n} \sum_{z_1 \cdots z_n} \prod_{i=1}^n P^\varepsilon_{\varepsilon^{2/5}}(x \to z_i) \left| E_\sigma\left(\prod_{i=1}^n [\sigma(z_i, t) - m_\varepsilon(z_i, t; \sigma)] \right) \right| \end{aligned} \tag{4.4}$$

It is then natural to give the following definition. For all $n \geqslant 1$, all n-tuplets $\mathbf{x} = (x_1, \ldots, x_n)$ of n distinct sites in Z_ε, all configurations σ, and all $\varepsilon > 0$ and $t \geqslant 0$ we set

$$v^\varepsilon_n(\mathbf{x}, t; \sigma) = E_\sigma\left(\prod_{i=1}^n [\sigma(x_i, t) - m_\varepsilon(x_i, t; \sigma)] \right) \tag{4.5}$$

We shall say that n is the degree of v^ε_n and such functions will be referred to as *v-functions*. Analogous definitions have been given for the simple symmetric exclusion process,[11] the weakly asymmetric simple exclusion process,[10] and the Boghosian–Levermore cellular automaton.[15] We have in all these cases (including ours) a bound on the v-functions which *decreases* exponentially on the degree n of the v-function, just what is needed to compensate the diverging factor $\varepsilon^{-\zeta n}$ in (4.4), if ζ is not too large. In Appendix A we prove such a bound [cf. (A.1)], which we report here in the following particular form. For all n, $\mathbf{x}$, σ, and $t \leqslant 2\varepsilon^a$ there is a constant c (depending on n and a) such that

$$|v^\varepsilon_n(\mathbf{x}, t; \sigma)| \leqslant c(\varepsilon^{-2} t)^{-n/8} \tag{4.6}$$

[actually we can prove that the same inequality, with a possibly different constant c, holds for t varying on any compact; since we do not need this stronger property we have simply stated (4.6), which is easier to prove]. From (4.6), (4.4), and well-known properties of independent random walks, it easily follows that for all positive k there exists c_k so that

$$P_\sigma(\|\sigma_t - m_{\varepsilon,t}\| \geqslant \varepsilon^\alpha) \leqslant c_k \varepsilon^k \tag{4.7}$$

$$\alpha = 1/4 - a \tag{4.8}$$

uniformly on $\varepsilon^a \leqslant t \leqslant 2\varepsilon^a$.

Equations (4.6) and (4.7) are intimately related; we have seen how to derive (4.7) from (4.6). Next we shall use (4.7) to extend (4.6) in the following sense.

Theorem 4.1. For all $\sigma \in \{-1, 1\}^{Z_\varepsilon}$ such that $\|\sigma\| \leqslant \varepsilon^{98a}$, for all n, $\mathbf{x} = (x_1,..., x_n)$, $\varepsilon > 0$, and $2\varepsilon^a \leqslant t \leqslant 2\varepsilon^{-a}$

$$|v_n^\varepsilon(\mathbf{x}, t; \sigma)| \leqslant c\varepsilon^{n\delta} \tag{4.9}$$

$$\delta = 1/4 - 3a \tag{4.10}$$

Remarks. With some more effort we could prove (4.9) with $\delta = 1/4$, but (4.10) is sufficient for our needs. The number 98 of course is not optimal; it arises from the fact that $\|\sigma_t\|$ might be ε^{-2a} larger than $|\bar{\sigma}_t|$ and we shall use Theorem 4.1 only for $|\bar{\sigma}_t| \leqslant \varepsilon^{100a}$.

The proof of Theorem 4.1 is an easy consequence of (4.6) and of the following lemma, as we shall see right after stating it.

Lemma 4.2. Let $\|\sigma\| \leqslant \varepsilon^{98a}$. Then there is a constant c such that the following hold. (i) For all $x \in Z_\varepsilon$

$$|m_\varepsilon(x, t; \sigma)| \leqslant c\|\sigma\|, \qquad \varepsilon^{2/5} \leqslant t \leqslant 2\varepsilon^{-a} \tag{4.11a}$$

$$|m_\varepsilon(x, t; \sigma) - \bar{\sigma}| \leqslant c(\|\sigma\|^3 \varepsilon^{-a} + \varepsilon^{2/5}) \qquad \varepsilon^{-a} \leqslant t \leqslant 2\varepsilon^{-a} \tag{4.11b}$$

(ii) For all positive k and b such that $0 < b \leqslant a$ there is c so that for all $x \in Z_\varepsilon$

$$|m_\varepsilon(x+1, t; \sigma) - m_\varepsilon(x, t; \sigma)| \leqslant c\varepsilon^k, \qquad \varepsilon^{-b} \leqslant t \leqslant 2\varepsilon^{-a} \tag{4.12}$$

(iii) Finally, let r be any number between 1 and 2 and $\sigma^{(n)}$, $n \geqslant 0$, any sequence of configurations such that $\sigma^{(0)} = \sigma$ and

$$\|m_\varepsilon(\cdot, r\varepsilon^a; \sigma^{(n)}) - \sigma^{(n+1)}\| \leqslant \varepsilon^{\alpha - a} \tag{4.13}$$

[$\alpha = 1/4 - a$; cf. (4.7)]. Then

$$\sup_{n\varepsilon^a r \leqslant 2\varepsilon^{-a}} \sup_x |m_\varepsilon(x, \varepsilon^a r; \sigma^{(n-1)}) - m_\varepsilon(x, n\varepsilon^a r; \sigma)| \leqslant c\varepsilon^{\alpha - 3a} \tag{4.14}$$

Remarks. Roughly speaking, the above lemma states that if the magnetization is *small*, then it stays small for a time ε^{-a}, here the dynamics is that defined by (4.1). Furthermore, during the evolution the magnetization becomes very *flat*. This effect is due to the finite-volume assumption and greatly simplifies all our considerations. Finally, if at all times $n\varepsilon^a r$ the magnetization is *slightly* changed (to mimic what happens in the particle system), then the overall effect is negligible; Lemma 4.2 just states quantitatively these considerations, which sound pretty obvious. In fact, they are quite elementary properties of the deterministic evolution (4.1) and we prefer to postpone their proofs to Appendix B.

Proof of Theorem 4.2. Let σ, n, $\mathbf{x}$, and t be as in Theorem 4.1. Let $r \in [1, 2]$ be such that $t = N\varepsilon^a r$, N being some positive integer. On the right-hand side of (4.5) we condition on what has happened till time $(N-1)\varepsilon^a r$ and we get

$$v_n^\varepsilon(\mathbf{x}, t; \sigma) = E_\sigma \left[E_\sigma^{(N-1)} \left(\prod_{i=1}^n [\sigma(x_i, \varepsilon^a r) - m_\varepsilon(x, t; \sigma)] \right) \right] \tag{4.15}$$

where $\sigma^{(N-1)}$ is the configuration at time $(N-1)\varepsilon^a r$.

In each square bracket we add and subtract the term $m_\varepsilon(x, r\varepsilon^a; \sigma^{(N-1)})$ and expand the product. The generic term that we obtain in this way looks like

$$v_m^\varepsilon(\mathbf{x}', r\varepsilon^a; \sigma^{(N-1)}) \prod_{x \in I} [m_\varepsilon(x, r\varepsilon^a; \sigma^{(N-1)}) - m_\varepsilon(x, N\varepsilon^a r; \sigma)] \tag{4.16}$$

where $\mathbf{x}'$ is a subset of $\mathbf{x}$ with m sites $(0 \leqslant m \leqslant n)$ and I is its complement in $\mathbf{x}$, the second factor being a constant with respect to the internal expectation in (4.15). The v-function is estimated by (4.6). By (4.14), the difference of the m's is bounded by $\varepsilon^{\alpha - 3a}$ if the sequence $\sigma^{(i)} = \sigma_{r\varepsilon^a i}$, $i = 0,\ldots, N-1$, satisfies the assumptions of Lemma 4.2 [cf. (4.13)]. This occurs with probability larger than $1 - c\varepsilon^k$, for any given k, as easily follows by using (4.7). Since we are integrating uniformly bounded functions, it is enough to choose the above k so that $k > n/4$ to prove (4.9) with $\delta = \alpha - 3a$, i.e., as given by (4.10). In this way Theorem 4.1 is proven.

In Appendix C we use Theorem 4.1 to prove Lemmas 3.2 and 3.3 and that the expression in (3.22) vanishes as $\varepsilon \to 0$. This will not be difficult; we need some extra notation and considerations introduced in Appendix A. So we are forced to postpone their proofs after Appendix A, and since Theorem 4.1 requires Lemma 4.2 proven in Appendix B, we shift the above proofs to Appendix C.

Before stating the next result, we need to introduce the following stopping times. For $\eta: \eta + 10a \in H$, let

$$t^{\varepsilon,\eta} = \inf\{t: \quad |\bar{\sigma}_t| \geqslant \varepsilon^{\eta}\} \wedge \varepsilon^{-2} \tag{4.17}$$

Proposition 4.3. For any $\eta: \eta + 10a \in H$ and for any positive k there is c so that

$$P_{\mu^\varepsilon}[\sup_{t \leqslant t^{\varepsilon,\eta}} \|\sigma_t\| > \varepsilon^{\eta - 2a}] \leqslant c\varepsilon^k \tag{4.18}$$

where μ^ε is as in Theorem 2.1.

Corollary to Proposition 4.3. The distribution of $\varepsilon^{1/2} t^{\varepsilon, 100a}$ under P_{μ^ε} converges to the law of the explosion time in (2.14).

We obviously have

$$t^{\varepsilon, 100a} = t^{\varepsilon, 1/4 - 10a} + \sum_{1}^{n^* - 1} [t^{\varepsilon, \eta_{n+1}} - t^{\varepsilon, \eta_n}] \tag{4.19}$$

where $\eta_n = 1/4 - n10a$, while n^* is defined in (3.1). Given any $\zeta > 0$ and n as in the above sum, denote η_n by η and η_{n+1} by η', then

$$\begin{aligned} &P_{\mu^\varepsilon}[|\varepsilon^{2\eta}(t^{\varepsilon,\eta'} - t^{\varepsilon,\eta}) - \tfrac{1}{2}| > \zeta] \\ &\quad \leqslant P_{\mu^\varepsilon}[\|\sigma_{t^{\varepsilon,\eta}}\| > \varepsilon^{\eta - 2a}] + P_{\nu^\varepsilon}[|\varepsilon^{2\eta}(t^{\varepsilon,\eta'} - t^{\varepsilon,\eta}) - \tfrac{1}{2}| > \zeta] \end{aligned} \tag{4.20}$$

where ν^ε is an allowed state (cf. the beginning of Section 3). The second term on the right-hand side of (4.20) can be made, by Theorem 3.1, as small as desired by choosing ε small enough; in fact, 1/2 is the explosion time for (2.20). The first term, by Proposition 4.3, is smaller than $c\varepsilon^k$, for any given k on the set $t^{\varepsilon, 100a} \leqslant \varepsilon^{-2}$. We shall see later that the complement of this set has vanishingly small probability. Therefore the probability that the sum on the right of (4.19) is larger than

$$[\tfrac{1}{2} + \zeta] \sum_{n=1}^{n^* - 1} \varepsilon^{-2\eta_n}$$

vanishes when $\varepsilon \to 0$. Then, because of the time normalization factor $\varepsilon^{1/2}$, this shows that the only contribution to $t^{\varepsilon, 100a}$ comes from the first term on the right of (4.19). Since $\mu^\varepsilon(\|\sigma\| > \varepsilon^{1/4 - 2a})$ also vanishes when $\varepsilon \to 0$, we can apply again Theorem 3.1 to conclude that the distribution of $t^{\varepsilon, 1/4 - 10a}$

converges to that of the explosion time for (2.14). Finally, by (4.19), if $t^{\varepsilon,100a} \geqslant \varepsilon^{-2}$, then for any $R>0$ we can find η_n such that $t^{\varepsilon,\eta_n} < \varepsilon^{-2}$ while $t^{\varepsilon,\eta_n+1} - t^{\varepsilon,\eta_n} > \varepsilon^{-2\eta_n}R$, or $t^{\varepsilon,1/4-10a} \geqslant \varepsilon^{-2}$. Again by Theorem 3.1 the probability of all these events vanishes when $\varepsilon \to 0$.

Proof of Proposition 4.3. Let us fix η as in Proposition 4.3 and set

$$A_k = \{\exists t \in [k\varepsilon^{-a}, (k+1)\varepsilon^{-a}]: \quad \|\sigma_t\| > \varepsilon^{\eta-2a}\}$$

and for $h \geqslant k$

$$B_{h,k} = P_{\mu^\varepsilon}(\{t^{\varepsilon,\eta} > h\varepsilon^{-a}\} \cap A_k)$$

It is enough to show that

$$\lim_{\varepsilon\to 0}\left[\sum_{h\varepsilon^{-a}\leqslant\varepsilon^{-2}} \sum_{k=0}^{h} B_{h,k} + P_{\mu^\varepsilon}(t^{\varepsilon,\eta}=0)\right]=0$$

Since obviously

$$\lim_{\varepsilon\to 0} P_{\mu^\varepsilon}(t^{\varepsilon,\eta}=0)=0$$

and since $B_{h,k} \leqslant B_{k,k}$, we shall prove Proposition 4.3 once we show that for any positive m there is c such that

$$B_{k,k} \leqslant c\varepsilon^m$$

We denote by $\sigma^{(k-1)}$ the random configuration at time $(k-1)\varepsilon^{-a}$, then

$$\begin{aligned} B_{k,k} &\leqslant E_{\mu^\varepsilon}[1(t^{\varepsilon,\eta} > (k-1)\varepsilon^{-a})\, P_{\sigma^{(k-1)}}(A_1)] \\ &\leqslant B_{k-1,k-1} + \sup_{\|\sigma\| \leqslant \varepsilon^{\eta-2a}} P_\sigma(A_1) \end{aligned}$$

Since for all m there is c such that

$$P_\sigma(A_1) \leqslant c\varepsilon^m \tag{4.21}$$

uniformly on $\|\sigma\| \leqslant \varepsilon^{\eta-2a}$, as we shall prove below, then, by iteration we reduce the proof to that of estimating $B_{0,0}$. It is easy to see that the analysis in Appendix A applies as well to a process starting from a product measure μ^ε rather than from a single configuration σ. The function m_ε is the defined by the initial condition $\mu^\varepsilon(\sigma(x))$, $x \in Z_\varepsilon$, and $v_n^\varepsilon(\mathbf{x}, t; \mu^\varepsilon)$, defined accordingly, satisfies (4.6) and (4.9). The same argument as for $B_{k,k}$, $k>0$, applies then to $B_{0,0}$; we omit details on this point.

To prove (4.21), we introduce a time grid which partitions the time interval $[\varepsilon^{-a}, 2\varepsilon^{-a}]$ into consecutive subintervals of length ε^γ, $\gamma>0$ will be chosen later. The probability that two spins flip in the same subinterval goes like $\cong \varepsilon^{2\gamma}\varepsilon^{-2}$ (this last factor counts the number of pairs of spins), since the flip intensity for a single spin is uniformly bounded. Since there are $\varepsilon^{-\gamma-a}$ subintervals, the probability that in any of these there are two spin flis is bounded by $\cong \varepsilon^{\gamma-2-a}$. We can choose γ so large that this probability vanishes as fast as any given power of ε. The probability that there is n such that, at $t=\varepsilon^{-a}+n\varepsilon^{\gamma}$ (and less than $2\varepsilon^{-a}$), $\|\sigma_t\|>2\varepsilon^{\eta}$ vanishes by (4.9) and (4.11b) as fast as any given power of ε. Since there is only one spin flip per subinterval if $\|\sigma_i\|<2\varepsilon^{\eta}$ at all $t=\varepsilon^{-a}+n\varepsilon^{\gamma}$, then $\|\sigma_t\|\leqslant \varepsilon^{\eta-2a}$ at all t in $[\varepsilon^{-a}, 2\varepsilon^{-a}]$.

Proof of Theorem 2.1 (Continued). The last ingredient that we are still missing is the analogue of Lemma 4.2 for the evolution of the magnetization past ε^{100a}. This is done in the following proposition, also proven in Appendix B.

Proposition 4.4. Let $\|\sigma\|\leqslant\varepsilon^{98a}$ and $|\bar{\sigma}\mp\varepsilon^{100a}|\leqslant\varepsilon$. For any $T>0$ let sup′ (respectively sup″) denote the supremum over all σ as above, all $r\in[1,2]$, all n such that $n\varepsilon^a r\leqslant\varepsilon^{-300a}+T\varepsilon^{-1/2}$ (respectively, $\varepsilon^{-300a}\leqslant n\varepsilon^a r\leqslant\varepsilon^{-300a}+T\varepsilon^{-1/2}$), all $x\in Z_\varepsilon$, and all sequences $\sigma^{(n)}$, $n\leqslant N$, satisfying (4.13). Then

$$\sup{}'|m_\varepsilon(x,\varepsilon^a r;\sigma^{(n-1)})-m_\varepsilon(x,n\varepsilon^a r;\sigma)|=0 \tag{4.22}$$

and

$$\lim_{\varepsilon\to 0}\sup{}''|m_\varepsilon(x,t;\sigma)\mp m^*|=0 \tag{4.23}$$

Corollary to Proposition 4.4. Same assumptions and notation as in Proposition 4.4. Then for any n, upon setting $T_\varepsilon=\varepsilon^{-300a}+T\varepsilon^{-1/2}$, we get

$$\lim_{\varepsilon\to 0}\sup_{x}\sup_{\varepsilon^{-800a}\leqslant t\leqslant T_\varepsilon}|v_n^\varepsilon(\mathbf{x},t;\sigma)|=0 \tag{4.24}$$

uniformly on σ (provided it satisfies the conditions in Proposition 4.4) [the first sup in (4.24) is over all n-tuplets $\mathbf{x}$ of distinct sites of Z_ε].

The proof of the Corollary follows from Proposition 4.4 in quite the same way as the proof of Theorem 4.1 follows from Lemma 4.2. We therefore omit the details.

Notice that from (4.24) and (4.23)

$$\lim_{\varepsilon\to 0}\sup_{\mathbf{x}}\sup_{\varepsilon^{-800a}\leqslant t\leqslant T_\varepsilon}\left|E_\sigma\left(\prod_{i=1}^{n}\sigma(x_i,t)\right)-(\pm m^*)^n\right|=0 \tag{4.25}$$

uniformly on σ (provided it satisfies the conditions of Proposition 4.4).

We now conclude the proof of Theorem 2.1. Let $t>0$; then, by the *strong Markov property*, setting $\eta = 100a$, $\mathbf{x} = (x_1,\dots, x_n)$

$$
\begin{aligned}
E_{\mu^\varepsilon}&\left(\prod_{i=1}^{n} \sigma(x_i, \varepsilon^{-1/2}t)\right)\\
&= \int dP^{\mu^\varepsilon}\, E_{\sigma_{t^{\varepsilon,\eta}}}\left(\prod_{i=1}^{n} \sigma(x_i, \varepsilon^{-1/2}t - t^{\varepsilon,\eta})\, 1(t^{\varepsilon,\eta} < \varepsilon^{-1/2}t - \varepsilon^{-300a}\right)\\
&\quad + E_{\mu^\varepsilon}\Bigg[\prod_{i=1}^{n} \sigma(x_i, \varepsilon^{-1/2}t)\{1(t^{\varepsilon,\eta} > \varepsilon^{-1/2}t)\\
&\quad + 1(\varepsilon^{-1/2}t - \varepsilon^{-300a} \leqslant t^{\varepsilon,\eta} \leqslant \varepsilon^{-1/2}t)\}\Bigg]
\end{aligned}
\tag{4.26}
$$

By Proposition 4.3 with $\eta = 100a$, $\sigma_{t^{\varepsilon,\eta}}$ satisfies the assumptions of Proposition 4.4 with a probability which goes to 1 as $\varepsilon \to 0$. Therefore, the first term on the right-hand side of (4.26) behaves like

$$
\begin{aligned}
(m^*)^n &\int dP_{\mu^\varepsilon}[\operatorname{sign}(\bar{\sigma}_{t^{\varepsilon,\eta}})^n]\, 1(t^{\varepsilon,\eta} \leqslant \varepsilon^{-1/2}t - \varepsilon^{-300a})\\
&= [\tfrac{1}{2}(m^*)^n + \tfrac{1}{2}(-m^*)^n]\, P_{\mu^\varepsilon}(t^{\varepsilon,\eta} \leqslant \varepsilon^{-1/2} - \varepsilon^{-300a})
\end{aligned}
$$

because of the symmetry of the process under global spin reversal. Finally, by the Corollary to Proposition 4.3, the latter probability converges to $c(t)$ [cf. (2.13)], so that this reconstructs the second term on the right-hand side of (2.13). By this same argument, i.e., by using the Corollary to Proposition 4.3, we have that

$$
\lim_{\varepsilon \to 0} P_{\mu^\varepsilon}(\varepsilon^{-1/2}t - \varepsilon^{-300a} \leqslant t^{\varepsilon,\eta} \leqslant \varepsilon^{-1/2}t) = 0
$$

so that we are left with the integral over $\{t^{\varepsilon,\eta} > \varepsilon^{-1/2}t\}$. We can replace this condition by $\{t^{\varepsilon,\eta} > \varepsilon^{-1/2}t - \varepsilon^{-a}\}$ for the same reasons as above. Hence we have

$$
\begin{aligned}
E_{\mu^\varepsilon}&\left[\prod_{i=1}^{n} \sigma(x_i, \varepsilon^{-1/2}t)\, 1(t^{\varepsilon,\eta} > \varepsilon^{-1/2}t - \varepsilon^{-a})\right]\\
&= E_{\mu^\varepsilon}\left[1(t^{\varepsilon,\eta} > \varepsilon^{-1/2}t - \varepsilon^{-a}) E_{\sigma'}\left(\prod_{i=1}^{n} \sigma(x_i, \varepsilon^{-a})\right)\right]
\end{aligned}
\tag{4.27}
$$

where σ' above is short hand for the configuration at time $\varepsilon^{-1/2}t - \varepsilon^{-a}$. By using Theorem 4.1 and (4.11) (the conditions for applying Lemma 4.2 are ensured by using Proposition 4.3), the expectation on σ' vanishes if $n \geqslant 1$

as $\varepsilon \to 0$; hence we reconstruct the first term on the right-hand side of (2.13) (using again the Corollary to Proposition 4.3 for estimating the probability that $t^{\varepsilon,\eta} > \varepsilon^{-1/2} - \varepsilon^{-a}$).

APPENDIX A

In this Appendix we prove that for all $n \geqslant 1$, all n-tuplets $\mathbf{x} = (x_1, \ldots, x_n)$ of distinct sites in Z_ε, all $\beta > 0$, and $\sigma \in \{-1, 1\}^{Z_\varepsilon}$

$$|v_n^\varepsilon(\mathbf{x}, t; \sigma)| \leqslant c_n(\varepsilon^{-2}t)^{-n/8}, \qquad 0 < t \leqslant \varepsilon^\beta \tag{A.1}$$

where $v_n^\varepsilon(\mathbf{x}, t; \sigma)$ is defined in (4.5), c_n is independent of ε, σ, and t, when t is in $(0, \varepsilon^\beta]$. The same inequality holds for the symmetric simple exclusion process[11] and the weakly asymmetric simple exclusion process[9]; cf. also ref. 15, where a cellular automaton variant of this last process (of interest for computer simulations) is considered. The proof of (A.1) is similar to the one for the weakly asymmetric simple exclusion process.

To simplify notation, we shall write $v^\varepsilon(\mathbf{x}, t)$ for $v_n^\varepsilon(\mathbf{x}, t; \sigma)$. We start computing the time derivative of v^ε, which, after some simple algebra, may be proven to have the following expression:

$$\frac{d}{dt} v^\varepsilon(\mathbf{x}, t) = \varepsilon^{-2} L v_n^\varepsilon(\mathbf{x}, t) + \psi_\varepsilon(\mathbf{x}, t) \tag{A.2}$$

where L is the generator of the symmetric simple exclusion process acting on the function $\mathbf{x} \to v^\varepsilon(\mathbf{x}, t)$, thought of as a function on $\{0, 1\}^{Z_\varepsilon}$ and supported by configurations with n particles, their positions being denoted by $\mathbf{x}$ (recall that L is defined as L_0, $\{-1, 1\}^{Z_\varepsilon}$ being replaced by $\{0, 1\}^{Z_\varepsilon}$, by changing all the -1's into 0's). For this reason in the sequel we shall often speak of the degree of a v-function as a number of particles, their positions being specified by the argument of the v-function.

The term ψ_ε has the following structure:

$$\begin{aligned}\psi_\varepsilon(\mathbf{w}, t) = {} & \varepsilon^{-2} \sum_{i \neq j} 1(x_j = x_i + 1)\{a_\varepsilon(x_i, x_j, t)[v^\varepsilon(\mathbf{x}^i, t) - v^\varepsilon(\mathbf{x}^j, t)] \\ & + b_\varepsilon(x_i, x_j, t)\, v^\varepsilon(\mathbf{x}^{i,j}, t)\} \\ & + \sum_{i=1}^{n} \sum_{\Delta, \Gamma \subset Z_\varepsilon} c_\varepsilon(i, \Delta, \Gamma, t, \mathbf{x})\, v^\varepsilon(\mathbf{x}_\Gamma^\Delta, t)\end{aligned} \tag{A.3}$$

where $\mathbf{x}^i = \mathbf{x}/x_i$, $\mathbf{x}^{i,j} = \mathbf{x}/[\{x_i\} \cup \{x_j\}]$, and $\mathbf{x}_\Gamma^\Delta = \mathbf{x} \cup \{x_i : i \in \Gamma\}/\{x_i : i \in \Delta\}$. The sum over Δ and Γ in (A.3) is restricted to $\Delta \subset \mathbf{x}$ and $\Gamma : \Gamma \cap \mathbf{x} = \varnothing$; other restrictions will be specified later. In the sequel we shall refer to the above, respectively, as "a-terms," "b-terms," and "c-terms." The functions a_ε, b_ε, and c_ε have the following bounds:

$$|a_\varepsilon(x_i, x_j, t)| \leqslant a\,|m_\varepsilon(x_i, t) - m_\varepsilon(x_j, t)| \tag{A.4}$$

$$|b_\varepsilon(x_i, x_j, t)| \leqslant b\,|m_\varepsilon(x_i, t) - m_\varepsilon(x_j, t)|^2 \tag{A.5}$$

$$c_\varepsilon(i, \Delta, \Gamma, t, \mathbf{x}) \leqslant c \tag{A.6}$$

where a, b, and c are universal constants. Finally, $c_\varepsilon(i, \Delta, \Gamma, \mathbf{x}, t) = 0$ unless both Δ and Γ are contained in the interval of Z_ε with endpoints $x_i - 1$ and $x_i + 4$. Furthermore, $c_\varepsilon(i, \Delta, \Gamma, \mathbf{x}, t) = 0$ if $\Delta = \{x_i\}$, $\Gamma = \varnothing$, and $|x_i - x_j| > 4$ for all j. In particular, therefore, if $|x_i - x_j| > 4$ for all $i \neq j$, then the degrees of the v-function appearing in (A.3) are not smaller than n. On the other hand, if there is a v-function in (A.3) with degree $n-(k+1)$, $k \geqslant 1$, then there is x_i and k distinct sites (particles) in $\mathbf{x}$, all different from x_i and at distance $\leqslant 4$ from x_i. The smallest degree is therefore $n-6$ and this occurs when in $\mathbf{x}$ there are six consecutive sites. This is all that we need to know about the coefficients a_ε, b_ε, and c_ε. As already mentioned, we interpret the degree of the v-function as number of particles; in this language the various terms in ψ_ε represent a birth–death process, so that the a- and b-terms describe deaths (of 1 and 2 particles, respectively), while the c-terms have, according to cases, either births or deaths or simultaneously births and deaths.

The following inequality will be used to estimate the a and b coefficients:

$$|m_\varepsilon(x, t) - m_\varepsilon(x+1, t)| \leqslant c(T)(\varepsilon^{-2} t)^{-1/2}, \qquad 0 < t \leqslant T \tag{A.7}$$

valid for all $T > 0$ and uniformly on ε and the initial configuration σ.

To prove (A.7), we go back to the definition of m_ε [cf. (4.1)] and we get

$$\begin{aligned}
|m_\varepsilon(x, t) - m_\varepsilon(x+1, t)| &\leqslant \sum_y |P_s^\varepsilon(x \to y) - P_s^\varepsilon(x+1 \to y)| \\
&\quad + \int_0^t ds\, c \sum_y |P_t^\varepsilon(x \to y) - P_t^\varepsilon(x+1, t)| \\
&\leqslant c(\varepsilon^{-2} t)^{-1/2} + c\varepsilon \sqrt{t} \\
&\leqslant c(T)(\varepsilon^{-2} t)^{-1/2}
\end{aligned} \tag{A.8}$$

because

$$c\varepsilon \sqrt{t} = c\varepsilon \frac{1}{\sqrt{t}}\, t \leqslant cT(\varepsilon^{-2} t)^{-1/2}$$

where c above and in the sequel is a constant whose value changes from line to line. Hence (A.7) is proven.

From (A.2) we get

$$v_n^\varepsilon(\mathbf{x}, t) = \int_0^t ds \sum_{\mathbf{y}} P_{t-s}^\varepsilon(\mathbf{x} \to \mathbf{y})\, \psi_\varepsilon(\mathbf{y}, s) \tag{A.9}$$

where $P_t^\varepsilon(\mathbf{x} \to \mathbf{y})$ denotes the probability that n stirring (simple exclusion) particles initially at $\mathbf{x}$ reach $\mathbf{y}$ at time t, i.e., such a probability is determined by the semigroup $\exp(\varepsilon^{-2} Lt)$. Notice that ψ_ε consists of a sum of terms each having a v-function as a factor; therefore, we can rewrite it using again (A.9). By iterating such a procedure, we obtain a final expression for the v-function where we have a stirring motion followed by some birth–death process followed again by stirring and so forth. The problem is then to estimate the transition probabilities of the stirring process. To do this it is, however, more convenient to start the iteration in a slightly different way. First we relate the stirring motion to the independent one via couplings, then we express the expectation in (A.9) using these couplings, and this is the expression to iterate rather than (A.9) itself.

The coupled process. This is a jump Markov process which describes the evolution of n *interacting* and n *independent* particles; here n is an arbitrarily fixed positive integer. We label the particles using the same labeling for the interacting and independent particles, for instance, i taking values in $\{1,..., n\}$. A priority list π is a permutation of the labels, $\pi = (\pi_1,..., \pi_n)$. Given π, we introduce the generator $\mathscr{L}_\pi$ and correspondingly a Markov jump process on $Z_\varepsilon^n \times Z_\varepsilon^n$ as

$$\begin{aligned}(\mathscr{L}_\pi f)(\mathbf{x}, \mathbf{x}^0) = \sum_{i=1}^n \sum_{d_i = \pm 1} \{ & 1(x_{\pi_j} \neq x_{\pi_i} + d_i, \forall j) \\ & \times [f(\mathbf{x} + d_i \mathbf{e}_{\pi_i}, \mathbf{x}^0 + d_i \mathbf{e}_{\pi_i}) - f(\mathbf{x}, \mathbf{x}^0)] \\ & + 1(\exists j > i\colon \quad x_{\pi_j} = x_{\pi_i} + d_i) \\ & \times [f(\mathbf{x} + d_i \mathbf{e}_{\pi_i} - d_i \mathbf{e}_{\pi_j}, \mathbf{x}^0 + d_i \mathbf{e}_{\pi_i}) - f(\mathbf{x}, \mathbf{x}^0)] \\ & + 1(\exists j < i\colon \quad x_{\pi_j} = x_{\pi_i} + d_i)[f(\mathbf{x}, \mathbf{x}^0 + d_i \mathbf{e}_{\pi_i}) - f(\mathbf{x}, \mathbf{x}^0)] \}\end{aligned} \tag{A.10}$$

where $\mathbf{e}_i = (\delta_{j,i}, j = 1,..., n)$. [The interpretation of the above expression is the following. When an independent particle jumps, the corresponding interacting particle tries to do the same. If the jump would lead the interacting particle to an empty site or to a site occupied by an interacting particle with lower priority, then the jump is actually performed (in the latter case the low-priority particle makes the opposite jump, so that the exclusion condition is satisfied and only an exchange takes place). The jump, on the contrary, is suppressed if it would lead the interacting particle to a site occupied by a particle with higher priority].

It is easy (trivial) to see that the marginal of the above process on the interacting (respectively independent) particles is the stirring (independent) process.

We can then rewrite (A.9) as

$$v^\varepsilon(\mathbf{x}, t) = \int_0^t ds\, E(\psi_\varepsilon(\mathbf{x}(t-s), s)) \tag{A.11}$$

where E denotes the expectation with respect to the coupled process, with π being the identity permutation of $\{1,..., n\}$, and $\mathbf{x}(0) = \mathbf{x}^0(0) = \mathbf{x}$. We shall iterate (A.11) rather than (A.9); at the successive stages of the iteration the meaning of the expectation in (A.11) will, however, change.

We describe in detail the first step of the iteration; the successive ones will be defined analogously. We think of the time s in (A.11) as fixed and we denote by $\mathbf{y}$, $\mathbf{y}^0$ the values of $\mathbf{x}(t-s)$, $\mathbf{x}^0(t-s)$, namely the configuration of the n stirring and n independent particles at time $t-s$ for the process considered in (A.11). We make explicit ψ_ε in (A.11) using (A.3) and we get v-functions with different degrees. We rewrite them using (A.11), where the new expectation E refers to a coupled process starting at time $t-s$ with a particle number, a priority list, and the initial position of the particles determined by $\mathbf{y}$, $\mathbf{y}^0$, i.e., (1) the positions of the particles at the end of the previous time interval, (2) by the particular term of ψ_ε under consideration, as we are going to discuss below.

The terms arising from (A.3) are labeled by λ_1; the subscript 1 indicates that this is the first step of the iteration. λ_1 takes finitely many values; their number is determined by the particles number, in this case n. Each value of λ_1 specifies a term in (A.3); if this is a b-term, then λ_1 specifies a pair of particles j, i. To take into account the numerical factors present in (A.3), we introduce a function $d_\varepsilon(\lambda_1, \mathbf{y}, t)$. For the above value of λ_1, this function equals $\varepsilon^{-2} b_\varepsilon(y_i, y_j, t)\ 1(y_j = y_i + 1)$ [cf. (A.3)]. The two particles with labels i and j disappear at time $t-s$ and the initial configuration for the new process starting at time $t-s$ is $\mathbf{z}$, $\mathbf{z}^0$, which is obtained from $\mathbf{y}$, $\mathbf{y}^0$ by dropping the stirring and independent particles with labels i and j. The new priority list also in this case is defined in some fixed but arbitrary fashion.

If the term selected is a c-term, then λ_1 specifies an index i and two sets Δ', Γ' both contained in the interval with endpoints $\{-1\}$ and $\{4\}$. In this case the function $d_\varepsilon(\lambda_1, \mathbf{y}, t) = c_\varepsilon(i, \Delta' + y_i, \Gamma' + y_i, t, \mathbf{y})$ (i, Δ', and Γ' being specified by λ_1), and 0 if $\Delta' + y_i$ and $\Gamma' + y_i$ are not among the allowed values, according to (A.3). For a c-term the following two possibilities may arise:

1. $|\Gamma| \geqslant |\Delta|$. We then choose in some fixed, arbitrary fashion $|\Delta|$ sites in Γ and label the particles which are born there using the labels of the particles which were in Δ. The remaining particles in Γ are newly born particles, which are labeled with new labels never used before. Same labels are used for the corresponding independent particles, which are then placed on the same sites as the stirring ones with same label. Hence, summarizing, when $|\Gamma| \geqslant |\Delta|$, the stirring particles in Δ are displaced, each moving at most by four sites, while the corresponding independent particles are not moved. Furthermore, there are $|\Gamma| - |\Delta|$ new particles on $|\Gamma| - |\Delta|$ sites of Γ. They are distinguished by new labels, and at the same places and with the same labels are created independent particles.

2. $|\Delta| > |\Gamma|$. We choose in some fixed, arbitrary way $|\Gamma|$ sites in $|\Delta|$ and use the labels of the particles on these sites to name the particles in Γ (later we shall see that it is convenient to name the particles in some specific fashion). The particles in the remaining sites of Δ die and disappear at time $t-s$ together with the corresponding independent particles. Thus, in this case only $|\Delta| - |\Gamma|$ particles in Δ die; the others are displaced, moving to Γ; each one therefore moves at most by four sites.

The initial configuration $\mathbf{z}$, $\mathbf{z}^0$ for the new process starting at time $t-s$ is then obtained from $\mathbf{y}$, $\mathbf{y}^0$ following the two rules above. The priority list in any such case is again fixed in some arbitrary fashion.

We have left to the end the analysis of the a-terms, since they require some special care. In fact, in the a-terms there are discrete gradients of v-functions [cf. (A.3)], namely differences $v^\varepsilon(\mathbf{y}^i, s) - v^\varepsilon(\mathbf{y}^j, s)$, with $y_j = y_i + 1$, which we cannot afford to neglect. By using (A.9), we can write the above gradient as

$$v^\varepsilon(\mathbf{y}^i, s) - v^\varepsilon(\mathbf{y}^j, s) = \int_0^s ds' \sum_{\mathbf{z}} [P^\varepsilon_{s-s'}(\mathbf{y}^i \to \mathbf{z}) - P^\varepsilon_{s-s'}(\mathbf{y}^j \to \mathbf{z})]\, \psi_\varepsilon(\mathbf{z}, s') \tag{A.12}$$

where $\mathbf{z}$ is a configuration with $n-1$ particles in Z_ε. *Since the marginal of the process of n stirring particles on any subset of $m < n$ particles is again a stirring process*, as can be easily checked, then we have from (A.12)

$$v^\varepsilon(\mathbf{y}^i, s) - v^\varepsilon(\mathbf{y}^j, s) = \int_0^s ds' \sum_{\mathbf{w}} P^\varepsilon_{s-s'}(\mathbf{y} \to \mathbf{w})[\psi_\varepsilon(\mathbf{w}^i, s') - \psi_\varepsilon(\mathbf{w}^j, s')] \tag{A.13}$$

where $\mathbf{w}$ is a configuration with n particles. The symmetric part of $P^\varepsilon_{s-s'}(\mathbf{y} \to \mathbf{w})$ under the exchange of w_i and w_j does not contribute to

(A.13). With this in mind we introduce the priority list π in such a way that $\pi_1 = i$ and $\pi_2 = j$. The other entries of π are fixed in some arbitrary fashion. Let E denote the expectation for the coupled process with such a priority list, considering $t-s$ as initial time and $\mathbf{y}$, $\mathbf{y}^0$ as initial configuration. Let τ be the first time after $t-s$ such that

$$[x_i^0(\tau) - x_i^0(t-s)] - [x_j^0(\tau) - x_j^0(t-s)] = x_j(t-s) - x_i(t-s) = 1 \qquad \text{(A.14)}$$

We denote by $g_{s,s'}(\lambda_1)$ the characteristic function that $\tau > (t-s) + (s-s')/2$, where s' is the time specified by (A.12) (the meaning of the label λ_1 will be discussed later; it classifies the various possibilities occurring in the present analysis). We then have

$$v^\varepsilon(\mathbf{y}^i, s) - v^\varepsilon(\mathbf{y}^j, s) = \int_0^s ds' \, E(g_{s,s'}(\lambda_1)[\psi_\varepsilon(\mathbf{x}(t-s')^i, s') - \psi_\varepsilon(\mathbf{x}(t-s')^j, s')]) \qquad \text{(A.15)}$$

because, conditioned on $\{\tau < t-s'\}$, the marginal on the position of the stirring particles at time $t-s'$ is symmetric under the exchange of particles i and j. This is an easy consequence of the fact that before τ the stirring particles i and j have the same displacement as the corresponding independent ones, because of the above choice of the priority list and due to the definition of τ. From (A.14) it then follows that at time τ the positions of the particles i and j have the same probability as when exchanged. From this argument it is clear that (A.15) would hold also if g were defined as the characteristic function that $\tau > t-s'$, the reason for our choice of g will become clear later. For more details on the above argument we refer to ref. 10 and proceed with our analysis.

When referring to an a-term the label λ_1 indicates a pair i, j [cf. (A.3)] and also one of the two indices, either i or j. This last indicates whether on the next iteration we shall pick up the term $\psi_\varepsilon(\mathbf{w}^i, s')$ or $\psi_\varepsilon(\mathbf{w}^j, s')$ in (A.15). Therefore, $d_\varepsilon(\mathbf{y}, \lambda_1, s) = \pm\varepsilon^{-2} a_\varepsilon(y_i, y_j, s)\ 1(y_j = y_i + 1)$ with a plus sign if λ_1 specifies the label i and the opposite sign if it specifies j. Furthermore, in the next expectation corresponding to the time interval $t-s$, $t-s'$ there will appear the function $g_{s,s'}(\lambda_1)$.

In this way we have completed the description of the first iteration of (A.11); the successive ones are performed analogously using the same convention and notation introduced above.

We iterate (A.11) N times, choosing N so that

$$\frac{\beta}{5}(N-n) > \frac{n}{8} + 1 \qquad \text{(A.16)}$$

and we have

$$v^{\varepsilon}(\mathbf{x}, t) = \sum_{m \leqslant N} \sum_{\lambda_1,\dots,\lambda_m} \int_0^t ds_1 \cdots \int_0^{s_{m-1}} ds_m$$

$$\times E\left(\prod_{i=1}^{m} d_{\varepsilon}(\mathbf{x}(t-s_i), \lambda_i, s_i \prod_{\lambda_j \in A} g_{s_j, s_{j+1}}(\lambda_j) \right) + R \qquad \text{(A.17)}$$

where the sum over $\lambda_1,\dots, \lambda_m$ is restricted to values of λ such that for all $i<m$ there is a nonzero number of particles in the time interval $t-s_i$, $t-s_{i+1}$, while at time $t-s_m$ all particles die and none survive. R is the *remainder term*, namely it takes into account the case when some particle survives at time $t-s_N$; therefore

$$R \equiv \sum_{\lambda_1,\dots,\lambda_N} \int_0^t ds_1 \cdots \int_0^{s_{N-1}} ds_N$$

$$\times E\left[\prod_{i=1}^{N} d_{\varepsilon}(\mathbf{x}(t-s_i), \lambda_i, s_i) \prod_{j<N:\, \lambda_j \in A} g_{s_j, s_{j+1}}(\lambda_j)\, v^{\varepsilon}(\mathbf{x}(t-s_N), s_N) \right] \qquad \text{(A.18)}$$

where in the last expression $\mathbf{x}(t-s_N)$ is understood to be a nonempty configuration.

The expectation in (A.17) and (A.18) once m, $s_1,\dots, s_m$, and $\lambda_1,\dots, \lambda_m$ have been fixed refers to a process which, in each time interval $t-s_{i-1}$, $t-s_i$, $i=1,\dots, m$, $s_0 \equiv 0$, is a coupled process of stirring and independent particles, the particle number and the priority list being specified by λ_{i-1} (for $i=1$ the particle number is n, the initial particle number, and the priority list is given by the trivial permutation, as already discussed). At the times $t-s_i$ some birth–death–displacement process takes place, as specified by the value of λ_i and by the particle configuration at time $(t-s_i)^-$, in agreement with the description given before. We shall denote by P the law of this process and by E its expectation without explicitly writing the dependence on $\lambda_1,\dots, \lambda_m$ and $s_1,\dots, s_m$, which should be thought of as fixed.

A key ingredient in estimating (A.18) is the following.

Proposition A.1. Fix m, $\lambda_1,\dots, \lambda_m$ in agreement with (A.18), and let a be any positive number. Then for any k there is a constant $c(k, a)$ such that

$$E(1-h) \leqslant c(k, a)\varepsilon^k \qquad \text{(A.19)}$$

where h is the characteristic function of the event

$$|[x_i(t-s)-x_i(t-s')]-[x_i^0(t-s)-x_i^0(t-s')]| \leqslant (\varepsilon^{-2}|s-s'|)^{1/4+a} \qquad \forall i, s, s': \quad s-s' \geqslant \varepsilon^{2-a} \tag{A.20}$$

Notice that here a does not have the same meaning as in the other sections (where it denoted the number 10^{-6}). It should also be understood that in (A.19) only triples i, s, s' are allowed for which particle i is present at the times $t-s$ and $t-s'$; such triples are determined once $\lambda_1,..., \lambda_m$ and $s_1,..., s_m$ are given.

For the proof of the above proposition we refer to refs. 10 and 11 and references therein.

Let ζ be the partition on the state space of the above process (λ_i and s_i having been fixed), which specifies all the increments of the independent particles in all the time intervals $t-s_i$, $t-s_i+(s_i-s_{i+1})/2$, for all i such that $\lambda_i \in A$ (i.e., all the a-terms). Denote by I the expectation of one of the terms in (A.17) obtained by fixing the values of the λ_i and the s_i. We then have

$$I \leqslant E\left(\prod_{\lambda_j \in A} g_{s_j, s_{j+1}}(\lambda_j) I_\zeta\right) + c\varepsilon^{k-2N} \tag{A.21}$$

where

$$I_\zeta = E_\zeta\left(h \prod_{i=1}^{m} |d_\varepsilon(\mathbf{x}(t-s_i), \lambda_i, s_i)|\right) \tag{A.22}$$

E_ζ in (A.22) denotes the conditional expectation on ζ.

To derive (A.21), we have noticed that by its definition g drops out of the conditional expectation and we have used Proposition A.1 to estimate the contribution of the term containing $(1-h)$. The factor ε^{-2N} collects all the factors ε^{-2} appearing in (A.3) (they are certainly less than N). Finally, c is the product of $c(k, a)$ times another constant which bounds the product of the a-, b-, and c-functions. Since the sum over the λ's contains finitely many terms, it is enough to choose $k = 2N + n$ in (A.21) to have the contribution from the last term in (A.21) bounded by the right-hand side of (A.1).

To estimate (A.22), we shall exploit the presence of h to reduce the expectation to one involving only independent particles; we first need to introduce some more notation.

Notation. *Old particles* are those initially present, i.e., those with labels 1,..., n. It is, however, convenient to rename particles during their evolution so that it never happens that an old particle is the only particle that at some given time dies and when this happens there is no other old particle *close* to it. More precisely, notice first that only in a-terms or c-terms can one single particle die and that in both cases there is at least one more particle which is involved in such an event. For the a-term, in fact, two particles are needed, for the c-term a particle, say i, and a set of other particles, those in Δ; cf. the definition of the c-terms. If in either case, there is just one old particle among those involved in the event and this, according to our former convention, was the one which died, then we switch labels, so that the label of the old particle survives. The choice of the other particle whose label disappears (if several possibilities are present) is fixed in some arbitrary fashion.

Because of the above convention, if an old particle dies and no other old particle is close to it, then necessarily at least another particle dies at the same time. We choose in some arbitrary fashion one of them (if several are present) and call it *an auxiliary particle to the old particle which is dying*. We then call *normal* a particle neither old nor auxiliary.

Particle j is a *direct descendant* of particle i if there is k so that λ_k specifies i and the birth of new particles among which there is particle j. Particle j is a *descendant* of particle i if there is a chain of direct descendant particles connecting i and j.

We denote by $e_\varepsilon(\lambda_i, \mathbf{x}(t-s_i), s_i)$ the function obtained from $|d_\varepsilon(\lambda_i, \mathbf{x}(t-s_i), s_i)|$ by dropping all conditions on the positions of normal particles which are contained in the characteristic functions of death events. More precisely, in an a-term which does not involve two old particles the characteristic function that these two particles should be close is dropped from a_ε. For a b-term the same happens if the two particles involved are both normal (in the other cases the two particles are either both old or one old and the other auxiliary). In the characteristic function of a c-term we only keep the condition that old and auxiliary particles possibly involved in the event are suitably close; the condition on the normal particles, if present, is dropped.

Using the above notation, we have

$$I_\zeta \leqslant E\left(h \prod_{i=1}^{m} e_\varepsilon(\lambda_i, \mathbf{x}(t-s_i), s_i)\right) \tag{A.23}$$

We now relax the conditions contained in the characteristic functions referring to the deaths of the old and the auxiliary particles, the only ones left in (A.23). Let i be the label of an auxiliary particle, let $o(i)$ be the label of the old particle to which it is auxiliary, let $t(i)$ be the time when i is born and $t'(i)$ the time when it dies, let x_i [respectively $x_{o(i)}$] be the position of particle i [respectively $o(i)$] at time $t(i)$, and finally let Δx_i^0 [respectively

$\Delta x^0_{o(i)}$] be the increment in the time interval $t(i)$, $t'(i)$ of the independent particle i [respectively $o(i)$]. Then, if $h \neq 0$,

$$|[\Delta x_i^0 - \Delta x^0_{o(i)}] - [x_i - x_{o(i)}]| \leqslant c[\varepsilon^{-2}|t'(i) - t(i)|]^{1/4+a} \qquad \text{(A.24)}$$

if $|t'(i) - t(i)| \geqslant \varepsilon^{2-a}$. An analogous condition is derived for the death of a set of old particles; in this case the initial time is 0 and the initial positions are fixed by the argument of the v-function we are considering. We therefore get an upper bound on I_ζ in (A.23) by changing the characteristic functions describing the deaths of the particles using conditions as in (A.24). If $|t'(i) - t(i)| \leqslant \varepsilon^{2-a}$, we simply bound by 1 the corresponding characteristic function. What we have then is not yet an expression involving just the evolution of independent particles; in fact, we still have in the argument of the characteristic functions the differences $x_i - x_{o(i)}$ which refer to the stirring process.

We consider the largest among the times $t(i)$; let it be $t(k)$. We then condition on the whole process up to time $t(k)$ and we also specify all the increments of the independent particles after time $t(k)$ except for the increments of the independent particle k. After such conditioning, because of the maximality of $t(k)$, all the characteristic functions are fixed except for the characteristic function of the event

$$|\Delta x_k^0 - C| \leqslant c[\varepsilon^{-2}|t'(k) - t(k)|]^{1/4+a}$$

where C is a constant specified by the conditioning. Since we are also conditioning on ζ, the variable Δx_k^0 is the sum of two quantities; one is the sum of the increments in the time intervals $t - s_i$, $t - s_i + (s_i - s_{i+1})/2$, with $t(k) < t - s_i < t'(k)$ and $t - s_i$ being a time when an a-term is present. These increments are fixed by ζ, while the others have the distribution of a symmetric nearest neighbor random walk with intensity 1 and moving for a time not smaller than $\varepsilon^{-2} \cdot \frac{1}{2}[t'(k) - t(k)]$. Therefore, the contribution from this event is bounded by

$$c|\varepsilon^{-2}[t'(k) - t(k)]|^{-1/4+a}$$

uniformly on x_k and $x_{o(k)}$ and the conditioning on ζ. Note that if $|t'(k) - t(k)| \leqslant \varepsilon^{2-a}$ then the characteristic function referring to this event is no longer present, according to our new notation. However, the contribution of the death of particles k and $o(k)$ can always be bounded by

$$c[\varepsilon^{-2}|t'(k) - t(k)|]^{-1/4+a}[c^{-1}\varepsilon^{-a[1/4-a]}]$$

which holds uniformly on $x_k, x_{o(k)}$, the conditioning on ζ and all $t'(k) - t(k)$. We shall further worsen the above bound by replacing

$t'(k)-t(k)$ by the length of the time interval beginning at time $t(k)$. Since the dependence on x_k and $x_{o(k)}$ has now disappeared, we can iterate the above procedure to the next auxiliary particle, and we keep doing this till all the characteristic functions concerning auxiliary particles have been estimated. We may still be left with characteristic functions involving deaths of only old particles, but now the branching structure of the process has been lost and we can proceed as for the symmetric simple exclusion.[11] Indeed, the argument is very similar to the one presented above, so we just state the result without further comments (*notation*: below and in the following t' stands for $\varepsilon^{-2}t$ and s'_i for $\varepsilon^{-2}s_i$):

$$\begin{aligned}I_\zeta \leqslant c\varepsilon^{-a(N/4)} &\left[\prod_{\lambda_i \in A_1} \frac{1}{(s'_i)^{1/2}} \frac{1}{|t'-s'_i|^{1/4-a}}\right]\left[\prod_{\lambda_i \in A_2} \frac{1}{(s'_i)^{1/2}}\right]\varepsilon^{-2|A|} \\ &\times\left[\prod_{\lambda_i \in B_1} \frac{1}{s_i'^{\,1-a}} \frac{1}{|t'-s'_i|^{1/4-a}}\right]\left[\prod_{\lambda_i \in B_2} \frac{1}{s_i'^{\,1-a}}\right]\varepsilon^{-2|B|} \\ &\times\left[\prod_{\lambda_i \in C_1} \frac{1}{|t'-s'_i|^{\gamma(\lambda_i)}}\right]\left[\prod_{\lambda_i \in C_3} \frac{1}{|s'_i-s'_{i+1}|^{1/4-a}}\right]\end{aligned} \tag{A.25}$$

where $|A|$ and $|B|$ equal the total number of a-terms and b-terms respectively; A_1 refers to a-terms where an old particle dies (so that they describe cases where an old particle dies close to another old particle), A_2 refers to a-terms where no old particle dies, while B_1 refers to b-terms where two old particles die; B_2 classifies terms where either no old particle dies or one old and one auxiliary particle die (the contribution coming from these terms is taken into account by the product over C_3). C_1 refers to c-terms where more than one old particle dies, say $k>1$ old particles die; then $\gamma(\lambda_i)=k/8-ka$; actually, a better estimate holds, namely $k/4-ka$, but the previous one is, on one hand, sufficient for our purposes, and, from on the other, notationally more convenient, as we shall see. The set C_2 collects both c-terms where there are births, but no auxiliary particle is born, or there are deaths involving one old and one auxiliary particle (as for B_2, the contribution from these terms is taken into account by the product over C_3); C_3 covers the remaining cases.

The first factor on the right-hand side of (A.25) compensates for the possibility that some of the time intervals are smaller than ε^{2-a}, in such a case in fact we miss the characteristic functions which give rise to the other factors in (A.25). We have also used (A.7) to bound the a_ε and b_ε coefficients, if in (A.7) $\varepsilon^{-2}t \geqslant 1$. In the opposite case we note that the left-hand side of (A.7) is less than 2. Therefore, $|m_\varepsilon(x,t)-m_\varepsilon(x+1,t)|^2 \leqslant c(t')^{-1+a}$.

Since the estimate (A.25) is uniform on ζ, we can easily derive a bound on I from (A.21) because the different g's are mutually independent. In fact, by well-known properties of simple random walks we have that

$$E(g_{s_j, s_{j+1}}(\lambda_j)) \leqslant c\,|s'_j - s'_{j+1}|^{-1/2} \tag{A.26}$$

(remember that $s' \equiv \varepsilon^{-2}s$). We go back to Eq. (A.17) and we fix one term in the sum over $\lambda_1,..., \lambda_m$. We then consider the contribution to this term coming from the first term on the right-hand side of (A.21). This is bounded by the function K defined below in (A.27). We first need some new notation: we say that s_i *is final* if λ_i belongs to any of the sets $\{B_1, B_2, C_1, C_2\}$. We call $k_1,..., k_q$ the decreasing sequence of all the indices of the final times. Then after some easy manipulation, consisting essentially in increasing the domain of integration of positive functions, we get for the above term the bound

$$K = c\varepsilon^{-aN/4} \prod_{i=1}^{q} F(\lambda_{k_i+1},..., \lambda_{k_{i+1}}; t') \tag{A.27}$$

where for any $i \in \{1,..., q\}$ setting $\{\lambda'_1,..., \lambda'_k\} \equiv \{_{k_i+1},..., \lambda_{k_{i+1}}\}$ gives

$$F(\lambda'_1,..., \lambda'_k; t') = \int_0^{t'} ds'_1 \cdots \int_0^{s'_{k-1}} ds'_k \\ \times \varepsilon^{2|C'|} \left[\prod_{i=1}^{k} (t' - s'_i)^{u_i}\right] \left[\prod_{i=1}^{k} s_i'^{\,v_i}\right] \left[\prod_{i=1}^{k} (s'_i - s'_{i+1})^{w_i}\right] \tag{A.28}$$

where $t' \equiv \varepsilon^{-2}t$, while $|C'|$ denotes the number of indices i such that $\lambda'_i \in C$. Furthermore, $u_i = -1/4 + a$ if $\lambda'_i \in A_1 \cup B_1$, $= \gamma(\lambda'_i)$ if $\lambda'_i \in C_1$, and $=0$ otherwise. $v_i = -1/2$ if $\lambda'_i \in A$ and $= -1 + a$ if $\in B$. Finally, $w_i = -1/2$ if $\lambda'_i \in A$ and $= -1/4 + a$ if $\lambda'_i \in C_3$.

It is not hard to see that

$$F(\lambda'_1,..., \lambda'_k; t') = (t')^{(-1/4+a)(|A'_1| + |B'_1| + |C'_3|)} (t')^{-\|C'_1\| + a|B'_2|} (\varepsilon^2 t')^{|C'|} \tag{A.29}$$

$$\|C'_1\| = \sum_{\lambda'_i \in C'_1} \gamma(\lambda'_i) \tag{A.30}$$

as is intuitively clear by using a scaling argument in (A.28). (We are using the same notational convention as before, so that, for instance, $|A'_1|$ denotes the number of indices i for which $\lambda'_i \in A_1$.)

We refer to ref. 11 for a more detailed analysis of the asymptotics of integrals like those in (A.28) and we proceed in proving (A.1). By multiplying all the F-functions, as required in (A.27), we get the same bound as in (A.29) with *primed sets* replaced by *unprimed ones.* Notice then that each element in A_1 and C_3 corresponds to the death of one old particle, each element in B_1 to the death of two old particles, while

$$\|C_1\| \equiv (1/8 - a)\, |||C_1|||$$

where $|||C_1|||$ is the total number of old particles which die according to C_1. Hence,

$$n = |A_1| + 2|B_1| + |||C_1||| + |C_3|$$

$$|A_1| + |B_1| + \frac{1}{2}\,|||C_1||| + |C_3| = \frac{n}{2} + \frac{1}{2}\,(|A_1| + |C_3|)$$

so that

$$K \leqslant c\varepsilon^{-aN/4}(t')^{-n/8}(t')^{a(|A_1| + |B_1| + |B_2| + |C_3| + |||C_1|||)}(\varepsilon^2 t')^{|C|} \qquad \text{(A.31)}$$

If $|C| > 0$, then $(\varepsilon^2 t')^{|C|} \leqslant \varepsilon^\beta$, so that by choosing a small enough we get

$$K \leqslant c(\varepsilon^{-2} t)^{-n/8}$$

as needed for proving (A.1). If, on the other hand, $|C| = 0$, then the Glauber interaction never acts and the corresponding terms in (A.17) are the same as those obtained for the stirring alone: for these the estimate (A.1) is proven in ref. 11.

We now examine the generic term in the sum in (A.18). We use the same procedure as above, bounding the v-function at time $t - s_N$ by 2^N. We get again the same structure as in (A.27) and (A.28) except for the fact that the last time s_N might not be final according to the definition stated before (A.27). The corresponding λ_N might in fact indicate also any a- or C_2 term. In particular, if this is an a-term, then we miss a factor $[\varepsilon^{-2}(s_N - s_{N+1}]^{-1/2}$ [since we are not iterating an $(N+1)$th time]. Therefore we get the same estimates as in (A.29) and (A.31) with the possible presence of a further ε^{-1} factor. Notice also that in this new case the total number of old particles which die might be less than n. If this number is k, we shall get only a contribution from deaths which go like $(t')^{-k/8}$, besides a divergent factor with exponent proportional to a, the same as before. The point is that there is a lower bound on $|C|$ and the converging factor $(\varepsilon^2 t')^{|C|}$ alone will be enough for proving (A.1), if N is large enough, as specified by (A.16). In fact, we have that

$$|A| + |B| + |C| = N$$

by definition. Furthermore, the number of dead particles is not larger than the total number of particles, which is bounded by $n+4|C|$, because at most four new particles can be created at each time. Therefore

$$|A| + |2|B| \leqslant n+4|C|$$
$$|A| + |B| \leqslant n+4|C|$$
$$|A| + |B| + |C| = N \leqslant n+5|C|$$

By (A.16),

$$\varepsilon^{(\beta/5)(N-n)}\varepsilon^{-1} < \varepsilon^{n/8} \tag{A.32}$$

Furthermore, by taking a small enough, we see that the same inequality holds even in the left-hand side is multiplied by $\varepsilon^{-5aN/4}$, which bounds the divergent factor in (A.31). Therefore (A.18) is also bounded as required by (A.1); hence (A.1) is proven.

We close this appendix by remarking that as in ref. 10 it is possible to extend (A.1) and prove that the same inequality holds for t ranging on the compacts. This extension does not require too much extra work, but since we do not need it in the proof of Theorem 2.1, we omit it.

APPENDIX B

Proof of Lemma 4.2. First notice that (4.1) can be written as

$$m_\varepsilon(x, t; \sigma) = \sum_z \bar{P}_t^\varepsilon(x \to z)\, \sigma(z) + \int_0^t ds \sum_z \bar{P}_{t-s}^\varepsilon(x \to z)\, \bar{g}(z, m_{\varepsilon,s}) \tag{B.1}$$

where $\bar{P}_t^\varepsilon$ is the transition probability of a symmetric random walk which jumps on nearest neighbor sites with intensity $\varepsilon^{-2}+2$ and $\bar{g}$ is obtained from g by dropping the first term on the right-hand side of (4.2).

Proof of (4.11a). For $t \geqslant \varepsilon^{2/5}$ we have

$$\left| \sum_z \bar{P}_t^\varepsilon(x \to z)\, \sigma(z) \right| \leqslant \varepsilon^{98a}$$

because $\|\sigma\| \leqslant \varepsilon^{98a}$. From (4.1)

$$|m_\varepsilon(x, t; \sigma)| \leqslant \|\sigma\| + c \int_0^t ds \sup_z |m_\varepsilon(z, s; \sigma)|^3, \qquad \varepsilon^{2/5} \leqslant t \leqslant 2\varepsilon^{-a}$$

and this integral inequality yields (4.11a) because $\|\sigma\| \leqslant \varepsilon^{98a}$.

Proof of (4.11b). This is a consequence of (B.1). In fact, from (4.11a), $|\bar{g}(z, m_{\varepsilon,s})| \leqslant c\,\|\sigma\|^3$ if $s > \varepsilon^{2/5}$ (and $< c$ if $s \leqslant \varepsilon^{2/5}$). Since for all $b' > 0$ and all n there is c so that

$$\sup_z |\bar{P}_t^\varepsilon(x \to z) - \bar{P}_t^\varepsilon(x \to z+1)| \leqslant c\varepsilon^n \qquad \forall t \geqslant \varepsilon^{-b'} \tag{B.2}$$

then

$$\left|\sum_z \bar{P}_t^\varepsilon(x \to z)\,\sigma(z) - \bar{\sigma}\right| \leqslant c\varepsilon^n$$

Proof of (4.12). We fix k and b as in (4.12) and from (B.1) and (B.2) we get for any n

$$\begin{aligned}&\sup_z |m_\varepsilon(x+1, t; \sigma) - m_\varepsilon(x, t; \sigma)| \\ &\qquad \leqslant c\varepsilon^n + \int_{t-\varepsilon^{-b}/N}^{t} ds \sup_z |\bar{g}(x+1, m_{\varepsilon,s}) - \bar{g}(x, m_{\varepsilon,s})|\end{aligned} \tag{B.3}$$

where $\varepsilon^{-b} \leqslant t \leqslant 2\varepsilon^{-a}$ and N is a positive integer which will be specified below.

By (4.11), then,

$$\sup_z |\bar{g}(x+1, m_{\varepsilon,s}) - \bar{g}(x, m_{\varepsilon,s})| \leqslant c\varepsilon^{2\times 98a} \sup_z |m_\varepsilon(x+1, s; \sigma) - m_\varepsilon(x, s; \sigma)| \tag{B.4}$$

We choose N so that

$$N\varepsilon^{2\times 98a}\varepsilon^{-a} \leqslant \varepsilon^k$$

Then, after N iterations of (B.3), we obtain (4.12) if we had chosen $n \geqslant k$ in (B.3).

Proof of (4.13). Let $\sigma^{(n)}$ be a sequence as in Lemma 4.2. Define

$$\Phi_n(t) = \sup_z |m_\varepsilon(x, t; \sigma) - m_\varepsilon(x, t - n\varepsilon^a r; \sigma^{(n)})|, \qquad n\varepsilon^a r < t \leqslant (n+1)\varepsilon^a r \tag{B.5}$$

for $n = 0,..., \bar{n}$, where $\bar{n}\varepsilon^a r \leqslant 2\varepsilon^{-a}$ and $\bar{n}$ is the largest integer for which this holds. Notice that $\Phi_0(t) = 0$. For $1 \leqslant n \leqslant \bar{n}$ we have, by (B.1),

$$\begin{aligned}\Phi_n(t) &\leqslant \sup_z |\bar{P}^\varepsilon_{t-n\varepsilon^a r}(x \to z)[m_\varepsilon(z, n\varepsilon^a r; \sigma) - \sigma^{(n)}(z)]| \\ &\quad + \int_0^{t-n\varepsilon^a r} \sup_z |\bar{g}(z, m_\varepsilon(\cdot, s + n\varepsilon^a r; \sigma)) - \bar{g}(z, m_\varepsilon(\cdot, s; \sigma^{(n)}))|\end{aligned}$$

We use (B.4) and we have for $t-n\varepsilon^a r \geqslant \varepsilon^{2/5}$

$$\Phi_n(t) \leqslant \Phi_{n-1}(n\varepsilon^a r) + \varepsilon^{\alpha-a} + c\int_{\varepsilon^{2/5}}^{t-n\varepsilon^a r} \varepsilon^{2\times 98a}\Phi_n(s+n\varepsilon^a r) + c\varepsilon^{2/5}$$

This integral inequality yields for $t-n\varepsilon^a r \geqslant \varepsilon^{2/5}$

$$\Phi_n(t) \leqslant \exp\{c\varepsilon^{2\times 98a}(t-n\varepsilon^a r-\varepsilon^{2/5})\}[\Phi_{n-1}(n\varepsilon^a r)+c\varepsilon^{\alpha-a}] \tag{B.6}$$

Hence, setting $c(\varepsilon)=c\varepsilon^{2\times 98a}(\varepsilon^a r-\varepsilon^{2/5})$,

$$\begin{aligned}\Phi_n(t) &\leqslant ce^{c(\varepsilon)}\sum_{k=0}^{n-1} e^{c(\varepsilon)k}\varepsilon^{\alpha-a}\\ &\leqslant c\varepsilon^{\alpha-a}\bar{n} \leqslant c\varepsilon^{\alpha-3a}\end{aligned}$$

and this completes the proof of Lemma 4.2.

Proof of Proposition 4.4. For notational simplicity let us fix the the sign of $\bar{\sigma}$ and let us assume that it is positive. Let $\sigma^{(n)}$, $0\leqslant n\leqslant N$, $\varepsilon^{-300a}\leqslant N\varepsilon^a r\leqslant \varepsilon^{-300a}+T\varepsilon^{-1/2}$, be as in Proposition 4.4. For $n\varepsilon^a r+\varepsilon^{2/5}\leqslant t\leqslant (n+1)\varepsilon^a r$ we can write, using (A.7),

$$\begin{aligned}m_\varepsilon(x,t;\sigma^{(n)}) \geqslant{}& \sum_z \bar{P}^\varepsilon_{\varepsilon^{2/5}}(x\to z)\, m_\varepsilon(z,\varepsilon^{2/5};\sigma^{(n)})\\ &+\int_{\varepsilon^{2/5}+n\varepsilon^a r}^{t} ds\sum_z \bar{P}^\varepsilon_{t-s}(x\to z)\\ &\times[m_\varepsilon(z,s;\sigma^{(n)})^3-\tfrac{9}{8}m_\varepsilon(z,s;\sigma^{(n)})^5]-c\varepsilon\end{aligned} \tag{B.7}$$

Denote by $\gamma_\varepsilon(x,t)$ $(\varepsilon^{2/5}\leqslant t\leqslant \varepsilon^a r)$ the solution to the equation

$$\begin{aligned}\gamma_\varepsilon(x,t)={}&\sum_z \bar{P}^\varepsilon_{t-\varepsilon^{2/5}}(x\to z)[m_\varepsilon(z,\varepsilon^{2/5};\sigma^{(n)})-c\varepsilon]\\ &+\int_{\varepsilon^{2/5}}^{t} ds\sum_z \bar{P}^\varepsilon_{t-s}(x\to z)[\gamma_\varepsilon(z,s)^3-\tfrac{9}{8}\gamma_\varepsilon(z,s)^5]\end{aligned} \tag{B.8}$$

Then $m_\varepsilon(x,t-n\varepsilon^a r;\sigma^{(n)})\geqslant \gamma_\varepsilon(x,t-n\varepsilon^a r)$. On the other hand, (B.8) has the same monotonicity properties as the reaction-diffusion equation (2.7), namely if a solution to (B.8) is not smaller than another solution to (B.8) at some time, then it stays so at all later times. Since

$$m_\varepsilon(x,\varepsilon^{2/5};\sigma^{(n)})\geqslant \sum_z P^\varepsilon_{\varepsilon^{2/5}}(x\to z)\, m_\varepsilon(z,\varepsilon^a r;\sigma^{(n-1)})-\varepsilon^{\alpha-a}-c'\varepsilon^{2/5} \tag{B.9}$$

we deduce that

$$m_\varepsilon(x, \varepsilon^a r; \sigma^{(n)}) \geqslant a_{n+1} \tag{B.10}$$

where a_{n+1} is the value at $\varepsilon^a r$ of the solution to the equation

$$\frac{d}{dt} z = z^3 - \tfrac{9}{8} z^5, \qquad \varepsilon^{2/5} \leqslant t \leqslant \varepsilon^a r \tag{B.11a}$$

$$z(\varepsilon^{2/5}) = a_n - \varepsilon^{\alpha - a} - c'\varepsilon^{2/5} - c\varepsilon \equiv a_n - c\varepsilon^{\alpha - a} \tag{B.11b}$$

(the value of the constant c keeps changing from one equation to another). In an analogous fashion we show that

$$m_\varepsilon(z, (n+1)\varepsilon^a r; \sigma^{(n)}) \leqslant b_{n+1}$$

where b_{n+1} is the value at $\varepsilon^a r$ of the solution to (B.11a) with initial value

$$z(\varepsilon^{2/5}) = b_n + c\varepsilon^{\alpha - a} \tag{B.12}$$

Finally, from (4.12) and (4.14), setting $\bar{n} = \varepsilon^{-2a}$, we have

$$\lim_{\varepsilon \to 0} (b_{\bar{n}} - a_{\bar{n}}) = 0$$

uniformly on r and on the choice of $\sigma^{(n)}$, $n \leqslant \bar{n}$. It is now easy to check that b_N and a_N converge to m^* as $\varepsilon \to 0$; we omit the details.

APPENDIX C

Proof of Lemmas 3.2 and 3.3. We consider only $\eta = 1/4$; the other cases are analogous and easier to study. When we make explicit $\gamma_1^\varepsilon(s)\,\gamma_1^\varepsilon(t)$ we obtain several terms: since they are all similar, we shall explicitly study only one of them, the only one which does not vanish when $\varepsilon \to 0$ [cf., however, the remark after (C.17) and the conclusion of the proof of Lemma 3.3]. So the term we are actually going to consider here is $E(\bar{\gamma}^\varepsilon(s)\,\bar{\gamma}^\varepsilon(t))$, where

$$\bar{\gamma}^\varepsilon = f'(\varepsilon^{-1/4}\bar{\sigma})\varepsilon^{1/4} \sum_{x \in Z_\varepsilon} \sigma(x-1)\,\sigma(x)\,\sigma(x+1) \tag{C.1}$$

The expectation above refers to the process starting from a configuration σ such that $|\bar{\sigma}| \leqslant \varepsilon^{1/4 - 10a}$ and $\|\sigma\| \leqslant \varepsilon^{1/4 - 12a}$, $\varepsilon^{-a} \leqslant s \leqslant t \leqslant 2\varepsilon^{-a}$.

The analysis is elementary since we have already established the main ingredients, i.e., (4.9) and (4.11); however, to take full advantage of them we need to use many times the integration by parts formula which relates the semigroup with generator L^ε to that with generator $\varepsilon^{-2}L_0$ [cf. (2.1)]: this makes the proof lengthy, since we have to estimate separately many different terms. They will always have the structure

$$g(\varepsilon^{-1/4}\bar{\sigma}) \prod_{i=1}^{n} \sigma(x_i)$$

where g is a C^∞ real-valued function bounded with all its derivatives; $\mathbf{x} = (x_1,..., x_n)$ is an n-tuplet of distinct sites in Z_ε (notice that $\bar{\gamma}^\varepsilon$ is a sum of terms having this form). In the sequel we shall simply write g for $g(\varepsilon^{-1/4}\bar{\sigma})$. On the above class of functions our generators act as follows:

$$L_0\left[g \prod_{i=1}^{n} \sigma(x_i)\right] = gL_0\left[\prod_{i=1}^{n} \sigma(x_i)\right] \tag{C.2}$$

while

$$L_G\left[g \prod_{i=1}^{n} \sigma(x_i)\right] = 2L_0\left[g \prod_{i=1}^{n} \sigma(x_i)\right] + \sum_{i=1}^{3} F_i^\varepsilon(g, \mathbf{x}) + \sum_{i=1}^{4} R_i^\varepsilon(g, \mathbf{x}) \tag{C.3}$$

where

$$F_1^\varepsilon(g, \mathbf{x}) = g \sum_{i=1}^{n} \sum_{\Delta} b(\Delta)\, 1(\{\Delta + x_i\} \cap \{\mathbf{x}/\mathbf{x}_i\} = \varnothing)$$
$$\times\left[\prod_{j\neq i} \sigma(x_j)\right]\left[\prod_{x\in\Delta+x_i} \sigma(x)\right] \tag{C.4a}$$

and $b(\Delta) = 0$ unless $\Delta = \{-1, 0, 1\}$, $\{2, 3, 4\}$, in which cases it equals $-1/3$, $3/2$, respectively, while

$$F_2^\varepsilon(g, \mathbf{x}) = g \sum_{i=1}^{n} \sum_{\Gamma} a(\Gamma)\, 1(\{\Gamma + x_i\} \cap \{\mathbf{x}/x_i\} = \varnothing)$$
$$\times\left[\prod_{j\neq i} \sigma(x_j)\right]\left[\prod_{x\in\Gamma+x_i} \sigma(\mathrm{x})\right] \tag{C.4b}$$

and $a(\Gamma) = 0$ unless $\Gamma = \{0, 1, 2, 3, 4\}$, $\{-1, 1, 2, 3, 4\}$, $\{-1, 0, 2, 3, 4\}$, in which cases it equals $-3/4$, $3/8$, $-3/4$, respectively.

Furthermore,

$$F_3^\varepsilon(g, \mathbf{x}) = -2\varepsilon^{3/4} \sum_{d(y,\mathbf{x})>6} c(y, \sigma)\, \sigma(y)\, g_y \prod_{j=1}^{n} \sigma(x_j) \tag{C.4c}$$

$$R_1^\varepsilon(g, \mathbf{x}) = g \sum_{i=1}^{n} \sum_{\Delta} b(\Delta)\, 1(\{\Delta + x_i\} \cap \{\mathbf{x}/x_i\} \neq \varnothing)$$
$$\times\left[\prod_{j\neq i} \sigma(x_j)\right]\left[\prod_{x\in\Delta+x_i} \sigma(x)\right] \tag{C.5a}$$

$$R_2^\varepsilon(g, \mathbf{x}) = g \sum_{i=1}^{n} \sum_{\Gamma} a(\Gamma)\, 1(\{\Gamma + x_i\} \cap \{\mathbf{x}/x_i\} \neq \varnothing)$$

$$\times \left[\prod_{j \neq i} \sigma(x_j)\right] \left[\prod_{x \in \Gamma + x_i} \sigma(x)\right] \tag{C.5b}$$

$$R_3^\varepsilon(g, \mathbf{x}) = g \sum_{i,j} 1(x_i = x_j + 1) \left[-\prod_{u=1}^{n} \sigma(x_u) + \prod_{u \neq i,j} \sigma(x_u) \right] \tag{C.5c}$$

$$R_4^\varepsilon(g, \mathbf{x}) = -2\varepsilon^{3/4} \sum_{d(y,\mathbf{x}) \leqslant 6} g_y c(y, \sigma)\, \sigma(y) \prod_{i=1}^{n} \sigma(x_i) \tag{C.5d}$$

where g_x, $x \in Z_\varepsilon$ is a short hand for the function g' computed as some suitable point in the interval with endpoints $\varepsilon^{-1/4}\bar{\sigma}$ and $\varepsilon^{-1/4}\bar{\sigma} - 2\varepsilon^{3/4}\sigma(x)$ (this comes from a remainder term in a Taylor expansion). Furthermore, $d(y, \mathbf{x})$ denotes the distance of y from $\mathbf{x}$.

The terms $F_1^\varepsilon + R_1^\varepsilon$ and $F_2^\varepsilon + R_2^\varepsilon$ take into account the action of L_G on $\prod \sigma(x_i)$ as if g were not present; the reason for splitting them into R and F functions will become clear later. The term R_3^ε also arises from the action of L_G on $\prod \sigma(x_i)$: the intensity $c(x, \sigma)$ contains a term equal to $1 - \frac{1}{2}\sigma(x+1)\, \sigma(x) - \frac{1}{2}\sigma(x-1)\, \sigma(x)$. This acts on the above product just as $2L_0$ when $|x_i - x_j| > 1$. When this condition is not fulfilled there is a correcting term: R_3^ε.

Because of the presence of the function g, more terms need to be added. In fact, when a spin flips, say at x, then $g(\varepsilon^{-1/4}\bar{\sigma})$ changes into $g(\varepsilon^{-1/4}\bar{\sigma} - 2\varepsilon^{3/4}\sigma(x))$. Expanding to first order, we obtain $F_3^\varepsilon + R_4^\varepsilon$. This explains the logic behind the expressions in (C.4) and (C.5); the actual proof is just computational and it is omitted.

Denoting by $\bar{P}_t^\varepsilon$ the transition probability for the process with generator $(\varepsilon^{-2} + 2)L_0$, we have, using the integration-by-parts formula,

$$E(\bar{\gamma}^\varepsilon(s)\, \bar{\gamma}^\varepsilon(t)) = \varepsilon^{1/2} \sum_{x, y, \mathbf{x}} \bar{P}_{t-s}^\varepsilon(x-1, x, x+1 \to \mathbf{z})$$

$$\times E\Bigg[1(\mathbf{z} \cap \{y-1, y, y+1\} = \varnothing)\, f'(\varepsilon^{-1/4}\bar{\sigma}_s)^2$$

$$\times \left[\prod_{i=1}^{3} \sigma(z_i, s)\right] \left[\prod_{i=-1}^{1} \sigma(y+i, s)\right]\Bigg] + R_0^\varepsilon$$

$$+ \int_s^t ds' \sum_{x, \mathbf{z}} \bar{P}_{t-s'}^\varepsilon(x-1, x, x+1 \to \mathbf{z})$$

$$\times E\left[\bar{\gamma}^\varepsilon(s) \left\{ \sum_{i=1}^{3} F_i^\varepsilon(f', s') + \sum_{i=1}^{4} R_i^\varepsilon(f', s') \right\} \right] \tag{C.6}$$

where R_0^ε is obtained from the first term on the right-hand side of (C.6) by replacing the set $\{\mathbf{z} \cap \{y-1, y, y+1\} = \varnothing\}$ by its complement.

We start by considering the R terms in the integral in (C.6), hereafter referred to as I_j, $j=1,..., 4$, and we are going to prove that they vanish as some positive power of ε when $\varepsilon \to 0$. We have by the Cauchy–Schwarz inequality

$$|I_j| \leqslant c\varepsilon^{1/2} \int_s^t ds' \left| E\left(\sum_{x,y} \prod_{i=-1}^{1} [\sigma(x-1, s)\, \sigma(y-i, s)] \right) \right|^{1/2}$$
$$\times \sum_{\mathbf{z}} \pi_{t-s'}^\varepsilon(\mathbf{z}) \left| E\left(\sum_{x,y} R_j^\varepsilon(f', \mathbf{z}+x)\, R_j^\varepsilon(f', \mathbf{z}+y) \right) \right|^{1/2} \qquad \text{(C.7)}$$

We have used the notation

$$\pi_t^\varepsilon(\mathbf{z}) = \bar{P}_t^\varepsilon(-1, 0, 1 \to \mathbf{z}) \qquad \text{(C.8a)}$$

$$\mathbf{z} + x = (z_1 + x,..., z_n + x) \qquad \text{(C.8b)}$$

and exploited the translational invariance of the stirring process in Z_ε, as well as the fact that f' is bounded.

By (4.11a) and (4.9) the first expectation grows like ε^{-1}, the leading term occurring when $x = y$, so that

$$|I_j| \leqslant c \int_s^t ds' \sum_{\mathbf{z}} \pi_{t-s'}^\varepsilon(\mathbf{z}) \left| E\left(\sum_{x,y} R_j^\varepsilon(f', \mathbf{z}+x)\, R_j^\varepsilon(f', \mathbf{z}+y) \right) \right|^{1/2} \qquad \text{(C.9)}$$

Case $j \leqslant 3$ in (C.9). By definition, $R_j^\varepsilon = 0$ when $j \leqslant 3$ and $\mathbf{z} \notin A$, where A denotes the set of all $\mathbf{z}$ such that at least two sites in $\mathbf{z}$ are at distance less than 6. Furthermore, for $n=3$ each term in R_j^ε contains at least one spin. Hence, if in (C.9) $|y-x| > 6$, then the expectation is over a product of at least two spins. We add and subtract $m_\varepsilon(\cdot, s'; \sigma)$ and expand the product. In this way we obtain sums of products of v-functions and m_ε-functions. By using (4.11a), (4.9), and the assumptions on the initial configuration σ, we find the bound $c\|\sigma\|^2 \leqslant \varepsilon^{2(1/4-12a)}$. Therefore for $j \leqslant 3$

$$|I_j| \leqslant c \int_s^t \sum_{\mathbf{z}} \pi_{t-s'}^\varepsilon(\mathbf{z})\, 1(\mathbf{z} \in A) \{\varepsilon^{-1}\varepsilon^{1/4-12a} + \varepsilon^{-1/2}\} \qquad \text{(C.10)}$$

The term $\varepsilon^{-1/2}$ comes from the sum over $|y-x| \leqslant 6$.

To bound the sum over $\mathbf{z}$ we use the following probability estimate.

Lemma C1. For $k < n$ denote by A_k^n the set of $\mathbf{z} = (z_1,..., z_n)$ having

k distinct pairs of sites such that in each pair the sites are at distance not larger than 6. Then

$$\sum_{\mathbf{z}} P_t^\varepsilon(\mathbf{x} \to \mathbf{z})\, 1(\mathbf{z} \in A_k^n) \leqslant c(\varepsilon^{-2} t)^{-1/2 - (k-1)(1-a)/4} \tag{C.11}$$

uniformly on $\mathbf{x}$ and t on the compacts, while for all $\mathbf{z}$, $\mathbf{x}$, and t

$$P_t^\varepsilon(\mathbf{x} \to \mathbf{z}) \leqslant \prod_{i=1}^{n} \left[\sum_{j=1}^{n} P_t^\varepsilon(x_i \to z_j) \right] \tag{C.12}$$

The inequality (C.11) follows from the coupling between the stirring and the independent processes introduced in Appendix A. We refer to ref. 11 and to the Appendix A in ref. 10. The inequality (C.12) is proven in ref. 12.

From (C.11) we then get ($A \subset A_1^3$ and $t - s \leqslant \varepsilon^{-a}$)

$$|I_j| \leqslant \varepsilon^{1/4 - 12a} \int_s^t ds' \, |t - s'|^{-1/2} \leqslant \varepsilon^{1/4 - 12a - a/2}$$

which completes the analysis of I_j when $j \leqslant 3$.

C1. Analysis of I_4

From (C.9) we have

$$|I_4| \leqslant c \int_s^t ds' \sum_{\mathbf{z}} \pi_{t-s'}^\varepsilon(\mathbf{z}) \varepsilon^{3/4}$$
$$\times [1(\mathbf{z} \notin A)\{\varepsilon^{-2} \varepsilon^{6(1/4 - 12a)} + \varepsilon^{-1}\}^{1/2} + 1(\mathbf{z} \in A)\varepsilon^{-1}] \tag{C.13}$$

because if $\mathbf{z} \notin A$, then $R_4^\varepsilon = \sum_x \tau_x$, where τ_x is the translate by x of τ_0, and τ_0 is a finite sum of terms such that in each of them there is a product of at least three spins. Hence when making explicit the sums in (C.9) we see that for each x there are only finitely many values of y such that $\tau_x \tau_y$ is not a product of spins in six different sites. In this latter case by (4.11a) and (4.9) we get the bound $\varepsilon^{6(1/4 - 12a)} \times \varepsilon^{-2}$. In the other case we simply bound τ_x by a constant, and $c\varepsilon^{-1}$ bounds the number of such terms. If $\mathbf{z} \in A$, we bound R_4^ε by a constant and this gives the last term in (C.13). Finally, the factor $\varepsilon^{3/4}$ in (C.13) comes from the definition of R_4^ε.

From (C.13) and (C.11), therefore, $|I_4| \leqslant c\varepsilon^{1/4 \quad a}$, the factor ε^{-a} coming from the time integral.

We complete the analysis of the R terms in (C.6) with the estimate of R_0^ε as follows.

C2. Estimate of R_0^ε

We have, using the same notation as above,

$$R_0^\varepsilon = \varepsilon^{1/2} \sum_{x,\mathbf{z}} \bar{P}_{t-s}^\varepsilon(x-1, x, x+1 \to \mathbf{z})$$
$$\times \Bigg[1(\mathbf{z} \notin A) E\left(f'^2 \sum_{i=1}^{3} \left\{ \prod_{j \neq i} \sigma(z_j, s) \sum_A e(A) \prod_{w \in A} \sigma(w + z_i, s) \right\} \right)$$
$$+ 1(\mathbf{z} \in A) E\left(f'^2 \sum_y 1(d(y, \mathbf{z}) \leqslant 1) \right) \prod_{i=-1}^{1} \sigma(y-i, s) \prod_{i=1}^{3} \sigma(z_i, s) \Bigg] \tag{C.14}$$

where $e(A) = 0$ unless $A = \{-2, -1\}, \{-1, 1\}, \{1, 2\}$, in which cases it has value 1. The first term in (C.14) is bounded, using Cauchy–Schwartz, by

$$c\varepsilon^{1/2} \sum_x \sum_{i=1}^{3} \sum_A e(A) E\Bigg[\sum_{\mathbf{z},\mathbf{z}'} \pi_{t-s}^\varepsilon(\mathbf{z})\, \pi_{t-s}^\varepsilon(\mathbf{z}')\, 1(\mathbf{z} \notin A)\, 1(\mathbf{z}' \notin A)$$
$$\times \left\{ \prod_{j \neq i} \sigma(z_j + x, s) \prod_{w \in A} \sigma(w + z_i + x, s) \right\}$$
$$\times \left\{ \prod_{j \neq i} \sigma(z_j' + x, s) \prod_{w \in A} \sigma(w + z_i' + x, s) \right\} \Bigg]^{1/2}$$
$$\leqslant c\varepsilon^{1/2} \varepsilon^{-1} \left[\sum_{\mathbf{z},\mathbf{z}'} \pi_{t-s}^\varepsilon(\mathbf{z})\, \pi_{t-s}^\varepsilon(\mathbf{z}') \sum_{k=0}^{3} 1(|\mathbf{z} \cap \mathbf{z}'| = k) \right]^{1/2} [\varepsilon^{(1/4 - 12a)(6-2k)}]^{1/2} \tag{C.15}$$

where we have used (4.9) and (4.11a), as usual, to estimate the product of spins possibly present; actually, the estimate is obviously better when $k = 0$. By (C.12) for any $\mathbf{z}$ it easily follows that for $k \geqslant 1$

$$\sum_{\mathbf{z}'} \pi_{t-s}^\varepsilon(\mathbf{z}')\, 1(|\mathbf{z} - \mathbf{z}'| = k) \leqslant c[\varepsilon^{-2}(t-s)]^{-k/2}$$

Hence, the contribution to R_0^ε from $\{\mathbf{z} \notin A\}$ goes like

$$c \sum_{k=0}^{3} \varepsilon^{-1/2} \varepsilon^{6/8 - 36a} \varepsilon^{-(1/4 - 12a)k + k/2} (t-s)^{-k/4}$$

Hence, the integral over $t - s$ (which appears in Lemmas 3.2 and 3.3) vanishes like $\varepsilon^{1/4 - 36a}$.

Let us now go back to (C.14) and consider the contribution of $\{\mathbf{z} \in \Lambda\}$. This is easy. We can afford to bound f' and σ by constants and we get a bound which goes like

$$\varepsilon^{-1}\varepsilon^{1/2} \sum_{\mathbf{z}} \pi_{t-s}^{\varepsilon}(\mathbf{z} \in \Lambda) \leqslant c\varepsilon^{1/2}(t-s)^{-1/2}$$

[the factor ε^{-1} comes from the sum over x in (C.14)], which, integrated over t, vanishes as $\varepsilon^{1/2-a/2}$.

C3. Analysis of the F Terms

We begin with F_3^ε. Let

$$F_0^\varepsilon(g, \mathbf{x}) = g'(\varepsilon^{-1/4}\bar{\sigma})\varepsilon^{3/4} \sum_{d(y,\mathbf{x})>6} \sum_{\Delta} [a(\Delta) + b(\delta)] \left[\prod_{x \in \Delta + y} \sigma(x)\right]\left[\prod_{i=1}^{n} \sigma(x_i)\right] \tag{C.16}$$

[cf. (C.3)]. We want to show that the contribution of $F_3^\varepsilon - F_0^\varepsilon$ in (C.6) vanishes like some positive power of ε. Let

$$F = -2\varepsilon^{3/4} \sum_{d(w,\mathbf{z})>6} c(w, \sigma)\,\sigma(w)[g_w - g'(\varepsilon^{-1/4}\bar{\sigma})] \prod_{j=1}^{n} \sigma(z_j)$$

and call I the contribution coming from such an F; we are actually interested in the case $g = f'$ and $n = 3$. By the definition of g_w the square bracket term is bounded by $c\varepsilon^{3/4}$. Then, by using Cauchy–Schwarz [as when deriving (C.7)] for each fixed $\mathbf{z}$ and w in F we get

$$\begin{aligned}|I| &\leqslant c \int_s^t ds'\, \varepsilon^{3/2} \sum_{\mathbf{z}} \pi_{t-s}^{\varepsilon}(\mathbf{z}) \sum_{w} E\left[\sum_{d(w,x+\mathbf{z})>6} \sum_{d(w,y+\mathbf{z})>6} \left\{\prod_{i=1}^{3} \sigma(x+z_i, s')\right.\right.\\ &\qquad \left.\left.\times\, \sigma(x+z_i', s')\right\}\right]^{1/2}\\ &\leqslant c \int -s^t\, ds'\, \varepsilon^{1/2}[\varepsilon^{-2}\varepsilon^{6(1/4-12a)} + \varepsilon^{-1}]^{1/2}\\ &\leqslant c\varepsilon^{1/4-a}\end{aligned}$$

by (4.9) and (4.11a), as usual by now.

This proves that we can change g_y into $g'(\varepsilon^{-1/4}\bar{\sigma})$ in F_3^ε, the difference giving vanishing contribution to (C.6). On the other hand, the difference between such a new F_3^ε and F_0^ε is given by

$$g'\varepsilon^{3/2} \sum_{d(y,\mathbf{x})>6} [\sigma(y+1) + \sigma(y-1) - 2\sigma(y)] \prod_{i=1}^{n} \sigma(x_i)$$

This is a telescopic sum, so that only the boundary terms contribute. The resulting expression has then the same structure as R_4^ε and gives the same contribution to (C.6) as R_4^ε, hence it is neglected.

All this proves that we can take the sum in (C.6) which refers to the F functions with i ranging from 0 to 2 (rather than from 1 to 3). The functions F_i^ε, $i=0, 1, 2$, are sum of terms like $g(\varepsilon^{-1/4}\bar{\sigma}) \prod_{i=1}^n \sigma(x_i)$ with $n=5$ for F_1^ε, $n=7$ for F_2^ε, $n=6, 8$ for F_0^ε. We call $\mathbf{w}$ such sites and the relevant quantity to estimate is (g below might either be f' or f'', according to cases)

$$I' = \varepsilon^{1/4} \sum_x E\left[\bar{\gamma}_1^\varepsilon(s)\, g(\varepsilon^{-1/4}\bar{\sigma}_s) \prod_{i=1}^n \sigma(w_i, s')\right] \tag{C.17}$$

Notice that the terms appearing in γ^ε but not in $\bar{\gamma}^\varepsilon$ have this form, so that their analysis can be done similarly to the following one. From (C.17), using the integration-by-parts formula, we get $I' = I_1' + I_2'$, where

$$\begin{aligned} I_1' = \int_s^{s'} du\, \varepsilon^{1/4} \sum_x \sum_{\mathbf{z}} \bar{P}_{s'-u}^\varepsilon(x + \mathbf{w} \to \mathbf{z}) \\ \times E\left[\bar{\gamma}^\varepsilon(s) \left\{\sum_{i=1}^3 F_i^\varepsilon(g, \mathbf{z}) + \sum_{i=1}^4 R_i^\varepsilon(g, \mathbf{z})\right\}\right] \end{aligned} \tag{C.18a}$$

$$I_2' = \varepsilon^{1/4} \sum_x \sum_{\mathbf{z}} \bar{P}_{s'-s}^\varepsilon(x + \mathbf{w} \to \mathbf{z}) E\left[\bar{\gamma}^\varepsilon(s) \prod_{i=1}^n \sigma(z_i, s)\right] \tag{C.18b}$$

C4. Contribution of the *R* Functions to I_1'

If $\mathbf{z} \in A_k^n$, $k \geqslant 2$, since $|R_i^\varepsilon| \leqslant c$ we get the following bound (using Cauchy–Schwarz):

$$\begin{aligned} &c\varepsilon^{-1} \int_s^{s'} du \sum_{\mathbf{z}} \bar{P}_{s'-u}^\varepsilon(\mathbf{w} \to \mathbf{z})\, 1(\mathbf{z} \in A_k^n, k \geqslant 2) \\ &\leqslant c\varepsilon^{-1} \int_s^{s'} du\, [\varepsilon^{-2}(s'-u)]^{-(3/4-a)} \\ &\leqslant c\varepsilon^{1/2-2a-a/4} \end{aligned}$$

If $\mathbf{z} \in A_1^n$, then the same analysis as when $n=3$ applies, since we know that the R functions contain at least one spin, hence these terms, too, vanish like a positive power of ε.

C5. Contribution of the F Functions to I_1'

The functions F_i^ε, $i=1, 2$, have the following structure:

$$\sum_{i=1}^{n} \sigma_\Delta(z_i) \prod_{j\neq i} \sigma(z_j)$$

where

$$\sigma_\Delta(z_i) = \prod_{x\in \Delta + z_i} \sigma(x)$$

and $|\Delta| = 1, 3, 5$ and it is contained in $[-1, 4]$. Furthermore, $\{z_j \cap (\Delta + z_i)\} = \varnothing$. Therefore, their contribution to I_1' is bounded by

$$\begin{aligned} c\sum_x \int_s^{s'} du \sum_i \sum_\Delta E\Bigg[&\sum_{\mathbf{z},\mathbf{z}'} \bar{P}^\varepsilon_{s'-u}(x+\mathbf{w}\to\mathbf{z})\, \bar{P}^\varepsilon_{s'-u}(x+\mathbf{w}\to\mathbf{z}') \\ &\times 1(\forall j\neq i:\quad z_j \cap \{\Delta + z_i\} = \varnothing)\, 1(\forall j\neq i:\quad z_j' \cap \{\Delta + z_i'\} = \varnothing) \\ &\times \sigma_\Delta(z_i, u)\, \sigma_\Delta(z_i', u) \prod_{j\neq i} \sigma(z_i, u)\, \sigma(z_i', u)\Bigg]^{1/2} \end{aligned}$$

which by (C.12), (4.9), and (4.11a) is bounded by

$$c\varepsilon^{-1} \int_s^{s'} du \left[\sum_{k=0}^{n} \varepsilon^{(1/4-12a)(2n-2k)} (\varepsilon^{-2}|s'-u|)^{(k\wedge 3)/2}\right]^{1/2} c\varepsilon^{1/4-61a}$$

since $n\geqslant 5$ (and the last inequality refers to $n=5$). For F_3^ε we proceed analogously: we have an extra factor $\varepsilon^{-1/4}$ $(=\varepsilon^{-1}\varepsilon^{3/4})$, which takes into account the sum over y in (C.4c). Hence, $n\geqslant 6$ and we get the bound

$$\leqslant c\varepsilon^{-1-1/4} \varepsilon^{(1/4-12a)6} \varepsilon^{-a} \leqslant c\varepsilon^{1/4-73a}$$

The estimate for I_2' is very similar to that for R_0^ε and it is omitted.

We have therefore proven that all terms in (C.6) vanish like some positive power of ε except for the first one on the right-hand side of (C.6). Set $\mathbf{w} = (z_1, z_2, z_3, y-1, y, y+1)$; then we have to study

$$J \equiv \varepsilon^{-3/2} E\left(f'(\varepsilon^{-1/4}\tilde{\sigma}_s)^2 \prod_{i=1}^{6} \sigma(w_i, s)\right) \tag{C.19}$$

The factor $\varepsilon^{-3/2}$ comes from the product of the factor $\varepsilon^{1/2}$ present in (C.6) and the extra factor ε^{-2} added to take care of the double sum over x and y in (C.6). Since our estimates will be uniform on $\mathbf{w}$, these sums just produce the diverging factor ε^{-2} that we are considering.

Using the integration-by-parts formula, we get

$$J = \varepsilon^{-3/2} \sum_{\mathbf{z}} \bar{P}_s^\varepsilon(\mathbf{w} \to \mathbf{z}) E\left[\prod_{i=1}^{6} \sigma(z_i, 1) f'(\varepsilon^{-1/4}\bar{\sigma}_1)^2\right]$$
$$+ \varepsilon^{-3/2} \int_1^s ds' \sum_{\mathbf{z}} \bar{P}_{s-s'}^\varepsilon(\mathbf{w} \to \mathbf{z}) E\left[\sum_{i=1}^{3} F_i^\varepsilon(f'^2, s') + \sum_{i=1}^{4} R_i^\varepsilon(f'^2, s')\right] \tag{C.20}$$

The integral term in (C.20) can be studied using the same techniques introduced above; we omit the details and we take for proven that it vanishes as a positive power of ε. To study the first term on the right-hand side of (C.20), we use the following lemma.

Lemma C.2. For any $k \geqslant 2$ and for any n there is c so that

$$\sup_{\mathbf{w},\mathbf{z},\mathbf{z}'} |\bar{P}_s^\varepsilon(\mathbf{w} \to \mathbf{z}) - \bar{P}_s^\varepsilon(\mathbf{w} \to \mathbf{z}')| \leqslant c\varepsilon^n \tag{C.21}$$

where $\mathbf{w} = (w_1, ..., w_k)$ and $s \geqslant \varepsilon^{-a}$, the constant c being independent of s and of the actual choice of $\mathbf{w}$).

Proof. Since $\bar{P}_s^\varepsilon(\mathbf{w} \to \mathbf{z}) = \bar{P}_s^\varepsilon(\mathbf{z} \to \mathbf{w})$ it is enough to estimate $|\bar{P}_s^\varepsilon(\mathbf{z} \to \mathbf{w}) - \bar{P}_s^\varepsilon(\mathbf{z}' \to \mathbf{w})|$. Assume $z_i' = z_i$ for all $i > 1$ (the general case is easily obtained from this by a telescopic sum). Let $\mathbf{x} = (x_1, ..., x_{k+1})$, $x_i = z_i$ for $i = 1, ..., k$ and $x_{k+1} = z_1'$. Then, letting τ be the stopping time introduced before (A.14) and referring to the particles with labels 1 and $k+1$, we have that for all n there is c so that

$$|\bar{P}_s^\varepsilon(\mathbf{z} \to \mathbf{w}) - \bar{P}_s^\varepsilon(\mathbf{z}' \to \mathbf{w})| \leqslant P(\tau > s) \leqslant c\varepsilon^n$$

where $P(\tau > s)$ is the probability that an independent symmetric random walk in Z_ε [and moving with intensity $2(\varepsilon^{-2} + 2)$ jumping on nearest neighbor sites starting from $z_1' - z_1$] does not reach the origin before s. The last inequality follows then from classical estimates on random walks and by the arbitrariness of n.

From (C.21) it easily follows that the first term on the right-hand side of (C.20) differs from

$$E[\varepsilon^{-3/2} f'(\varepsilon^{-1/4}\bar{\sigma}_1)^2 (\bar{\sigma}_1)^6] \tag{C.22}$$

by a quantity which vanishes when $\varepsilon \to 0$. On the other hand, the expression inside the expectation in (C.22) equals the function $x^6 f'(x)^2$, $x = \varepsilon^{-1/4}\bar{\sigma}_1$, which is uniformly bounded, i.e., independently of the value of $\bar{\sigma}_1$; hence, Lemma 3.2 is proven. Notice that we can repeat the whole procedure starting from the function $(\varepsilon^{-1/4}\bar{\sigma})^3$ rather than from $\bar{\gamma}^\varepsilon$ and we would obtain just the same result, since the only contribution comes when we apply the stirring generator in the integration-by-parts formula and after a time of the order of ε^{-a} the initial sites of the spins are completely forgotten. This proves that for some positive β

$$\int_s^{s+\varepsilon^{-a}} dt\, |E[\{\bar{\gamma}_t^\varepsilon - (X_t^\varepsilon)^3\}\{\bar{\gamma}_s^\varepsilon - (X_s^\varepsilon)^3\}]| \leqslant c\varepsilon^\beta \tag{C.23}$$

When replacing $\bar{\gamma}$ by γ we get an additional contribution coming from products of more spins. The same analysis presented above allows us to conclude that the leading contribution behaves as

$$E[\varepsilon^{-3/2} f'(\varepsilon^{-1/4}\bar{\sigma}_1)^2 (\bar{\sigma}_1)^k] \tag{C.24}$$

where $k > 6$. By using Cauchy–Schwarz we find the following bound on (C.24):

$$cE([\bar{\sigma}_1]^{2(k-6)})^{1/2}$$

By (4.9) and (4.10) this vanishes like some positive power of ε, hence, Lemma 3.3 is also proven.

Proof of (3.22). After making explicit the square in (3.22) we obtain a double sum over x and y of products $h(x, t)\, h(y, t)$. The function $h(x, t)$ is obtained by shifting by x the cylinder function $h(0, t)$. We can neglect the cases where the two functions have nondisjoint bases; they vanish like ε because of the normalization in (3.22). By making explicit the product of the two functions we see that this is a sum of products of spins. But we are now in the conditions discussed during the proof of Lemmas 3.2 and 3.3; hence, the expectation of the above product vanishes like a positive power of ε. We omit the details.

ACKNOWLEDGMENTS

We are indebted to D. A. Dawson for pointing out the right scaling properties of the model and for many enlightening discussions. We are also grateful to T. Shiga for useful comments and to G. Jona-Lasinio for showing us the analogies between our model and the stochastic quantization problems.

The final version of this paper has benefitted from useful comments and discussions following seminars on this subject held at UCLA, Princeton, Boulder, Cincinnati, and the Courant Institute.

Part of this work was done during the visits of M. E. V. to Rome, P. C. to IMPA, A. P. to Marseille Luminy, and E. P. to Rutgers and to Boulder, Colorado: we all acknowledge the very kind hospitality we received.

This work was partially supported by CNRPSAITM, CNPq grants 401112/86 and 1.19.10.002/86, and the CNR–CNPq agreement.

REFERENCES

1. F. Baras, G. Nicolis, M. Malek Mansour, and J. W. Turner, *J. Stat. Phys.* **32**:1 (1983).
2. P. Billingsley, *Convergence of Probability Measures* (Wiley, New York, 1968).
3. G. Broggi, L. A. Lugiato, and A. Colombo, *Phys. Rev. A* **32**:2803–2812 (1985).
4. M. Cassandro, A. Galves, E. Olivieri, and M. E. Vares, *J. Stat. Phys.* **35**:603 (1984).
5. D. A. Dawson, *J. Stat. Phys.* **31**:29 (1983).
6. G. F. Dell'Antonio, Small random perturbations and uniqueness of the limit measure for a vector field with a manifold of zeroes, CARR Preprint 6/88.
7. A. De Masi, P. A. Ferrari, and J. L. Lebowitz, *Phys. Rev. Lett.* **55**:19, 1947 (1985).
8. A. De Masi, P. A. Ferrari, and J. L. Lebowitz, *J. Stat. Phys.* **44**:589 (1985).
9. A. DeMasi, C. Kipnis, E. Presutti, and E. Saada, Microscopic structure at the shock in the asymmetric simple exclusion, *Stochastics* (1987), submitted.
10. A. DeMasi, E. Presutti, and E. Scacciatelli, The weakly asymmetric simple exclusion process, *Ann. Inst. Henri Poincaré* (1989).
11. A. DeMasi, E. Presutti, and M. E. Vares, *J. Stat. Phys.* **44**:645 (1986).
12. P. A. Ferrari, E. Presutti, E. Scacciatelli, and M. E. Vares, The symmetric simple exclusion process. I. Probability estimates, Preprint UCSB (October 1987).
13. R. L. Holley and D. W. Stroock, *Kyoto Univ. Res. Inst. Math. Sci. Publ. A* **14**:86 (1978).
14. N. Ikeda and S. Watanabe, *Stochastic Differential Equations and Diffusion Processes* (North-Holland, 1981).
15. G. Jona-Lasinio and P. K. Mitter, *Commun. Math. Phys.* **101**:409–436 (1985).
16. J. L. Lebowitz, E. Orlandi, and E. Presutti, Convergence of stochastic cellular automaton to Burgers' equation: Fluctuations and Stability, *Physica D* (1988).
17. J. L. Lebowitz, E. Presutti, and H. Spohn, Microscopic models of hydrodynamical behavior, *J. Stat. Phys.* **51**:841 (1988).
18. T. M. Liggett, *Interacting Particle Systems* (Springer-Verlag, 1985).
19. C. W. Meyer, G. Ahlers, and D. S. Cannell, *Phys. Rev. Lett.* (1987).
20. D. W. Stroock and S. Varadhan, *Multidimensional Diffusion Processes* (Springer-Verlag, 1979).
21. W. D. Wick, *J. Stat. Phys.* **38**:1005–1025 (1985).