Large deviations for the symmetric simple exclusion process in dimensions $d \geq 3$

J. Quastel[1], F.Rezakhanlou[2] and S.R.S. Varadhan[3]

1.Introduction

A classical theorem of Sanov identifies the entropy function as the rate functional for the large deviations of the empirical measure. More precisely if $X_1, X_2, \ldots, X_n, \ldots$ is a sequence of independent and identically distributed random variables with values in some Polish space D, having a common distribution P, then the strong law of large numbers asserts that, in the weak topology on the space of probability measures on D,

$$\lim_{L\to\infty} \frac{1}{L} \sum_{i=1}^{L} \delta_{X_i} = P \tag{1.1}$$

with probability one. Moreover the probabilities of deviations of the left hand side of (1.1) from P are exponentially small in the size L. Roughly speaking one has

$$\text{Prob}\left[\frac{1}{L}\sum_{i=1}^{L} \delta_{X_i} \;\text{ is close to } Q\right] \simeq \exp\{-LH(Q\,,P)\} \tag{1.2}$$

where $H(\cdot,\cdot)$ is the entropy functional

$$H(Q,P) = \begin{cases} \int \log \frac{dQ}{dP}\, dQ & \text{if } Q << P \\ +\infty & \text{otherwise.} \end{cases} \tag{1.3}$$

The aim of this article is to prove a theorem similar in spirit in the context of an interacting particle system. If we have a system of L interacting particles and denote by $x_1(\cdot), \ldots, x_L(\cdot)$ their trajectories in some finite interval $[0,T]$, we can form the empirical process

$$R_{L,\omega} = \frac{1}{L}\left[\delta_{x_1(\cdot)} + \cdots + \delta_{x_L(\cdot)}\right] \tag{1.4}$$

which will be a random measure on a suitable function space of trajectories on $[0,T]$. As $L \to \infty$, assuming that the model and the scaling have been properly chosen, one can expect a law of large numbers type result asserting the convergence of $R_{L,\omega}$ to a non-random limit P. We are interested in the rate function for large deviations from this

[1] Department of Mathematics, University of California, Davis, CA 95616. Partially supported by NSF grant DMS-9504791

[2] Department of Mathematics, University of California, Berkeley, CA 94720. Partially supported by NSF grant DMS-9424270

[3] Courant Institute, 251 Mercer St., New York, NY 10012. Partially supported by NSF grant DMS-

convergence and our goal is the explicit calculation of the rate function in terms of physical quantities arising out of the model.

Although questions of this sort are relevent for large classes of models, we will restrict our attention in this article to a special class of models known as random walks with symmetric simple exclusion. Our physical space will be the d-dimensional integer lattice modulo N, consisting of N^d lattice sites. The lattice sites will be denoted by letters x, y, etc. We will scale the periodic lattice by $\frac{1}{N}$, obtaining for each point on the lattice a corresponding point on the d-torus and these will be denoted by θ. We will have a certain number L of particles on the lattice and their locations will be denoted by $x_1, \ldots, x_L$. Since the particles move, their locations are functions of t and will be denoted by $x_1(t), \ldots, x_L(t)$. Each particle waits for an exponential time at its current site and then tries to jump to a new site. The new site is picked randomly, with the probability of attempting a jump from x to $x+z$ being given by $p(z)$. The probability distribution $p(\cdot)$ of the jump size is assumed to be symmetric, irreducible and of finite range. The jumps however are not always accomplished. Simple exclusion means that the particle is not allowed to jump to a site that is already occupied. In such a situation the particle does not jump and waits for a new exponential time. Of course, all particles are doing this simultaneously and independently of each other. Because of continuous time there will be no ties to resolve.

If we start our L particles with at most one particle per site, then it will remain so for all times. For any initial configuration $(x_1, \cdots, x_L)$ this defines a stochastic process on the space $D\big[[0,\infty), (Z_N^d)^L\big]$. We will rescale the process by defining

$$y_i(t) = \frac{1}{N} x_i(N^2 t) \tag{1.5}$$

and view $(y_1(t), \cdots, y_L(t))$ as taking values in $D\left[[0,T], \mathbf{T}^d\right]$ for some fixed time T. The corresponding empirical measure then is

$$\hat{R}_{L,\omega} = \frac{1}{L}\left[\delta_{y_1(\cdot)} + \cdots + \delta_{y_L(\cdot)}\right]. \tag{1.6}$$

We let $N \to \infty$ in such a way that

$$\lim_{N\to\infty} \frac{L}{N^d} = \bar{\rho} \tag{1.7}$$

exists and the initial configuration $\xi = (x_1, \ldots, x_L)$ is assumed to behave in such a way that

$$\lim_{N\to\infty} \frac{1}{N^d}\left[\delta_{y_1(0)} + \cdots + \delta_{y_L(0)}\right] = \rho_0(\theta) d\theta \tag{1.8}$$

exists weakly on $\mathbf{T}^d$. Note that because of the exclusion $0 \le \rho_0(\theta) \le 1$ and $\int_{\mathbf{T}^d} \rho_0(\theta) d\theta = \bar{\rho}$.

One can ask if under these conditions $\hat{R}_{L,\omega}$ satisfies a law of large numbers with a non-random weak limit P. It will be more convenient to replace $\hat{R}_{L,\omega}$ by $R_{N,\omega}$ where

$$R_{N,\omega} = \frac{1}{N^d}\left[\delta_{y_1(\cdot)} + \cdots + \delta_{y_L(\cdot)}\right]. \tag{1.9}$$

and in view of (1.7) they are essentially the same.

It is relatively straightforward to prove that the marginal distribution

$$R_{N,\omega,t} = \frac{1}{N^d}\left[\delta_{y_1(t)} + \cdots + \delta_{y_L(t)}\right] \tag{1.10}$$

of the empirical process at time t satisfies a law of large numbers and converges weakly to a non-random limit $\rho(t,\theta)d\theta$ on $\mathbf{T}^d$. The density $\rho(t,\theta)$ is determined as the solution of the linear heat equation

$$\frac{\partial\rho}{\partial t} = \frac{1}{2}\nabla\cdot D\nabla\rho \tag{1.11}$$

with $\rho(0,\theta) = \rho_0(\theta)$. The diffusion matrix D is nothing but the covariance matrix of the distribution $p(\cdot)$ of the jump sizes. The symmetry of $p(\cdot)$ makes the exclusion irrelevent in this context, and the result is basically the central limit theorem for the symmetric random walk. One should note however that this happens because the particles are not distinguishable at this level. On the other hand a law of large numbers also holds for the empirical process $R_{L,\omega}$ and now the particles are tagged and one has as the weak limit an inhomogeneous diffusion process on $\mathbf{T}^d$ with initial density $\rho_0(\theta)$. This diffusion process has a time dependent generator

$$\frac{1}{2}\nabla\cdot S\left(\rho(t\,,x)\right)\nabla + \frac{1}{2}\left[S\left(\rho(t\,,x)\right) - D\right]\frac{(\nabla\rho)(t\,,x)}{\rho(t\,,x)}\cdot\nabla \tag{1.12}$$

where $\rho(t,\theta)$ is given by (1.11). Here $S = S(\rho)$ is the self diffusion coefficient that describes the diffusive behavior of a tagged particle in equilibrium at density ρ. The existence of this diffusion matrix was established in [KV] and the law of large numbers was proved in [R] based on the earlier work in [Q1].

Here a qualitative difference arises between the nearest neighbor case in $d = 1$, i.e. $p(1) = p(-1) = \frac{1}{2}$ in $d = 1$, and all other cases. In the one dimensional nearest neighbor case, particles block each other completely and therefore $S(\rho) \equiv 0$ for $\rho > 0$. In all other cases, $S(\rho) > 0$ for $0 \le \rho < 1$ with $S(\rho) \to D$ as $\rho \to 0$, and to 0 as $\rho \to 1$.

Although the special case of $S(\rho) \equiv 0$ is relatively simple, we shall focus our attention exclusively on the alternate situation. Most of our results are valid in all dimensions. But at some crucial steps we need to know that $S(\rho)$ is Lipschitz continuous in ρ and this is known only when $d \ge 3$ [V2]. So the complete results are valid only when $d \ge 3$.

Now we turn to the large deviation behavior of the empirical process (1.9). As a first step we may look at the large deviation behavior of the empirical distribution (1.10) as measure–valued functions on $[0,T]$. This was carried out in [KOV] where it was shown that a large deviation principle is valid for $R_{N,t,\omega}$. If $q(t,\theta)$ is an arbitrary nonnegative function of t and θ with $\int q(t,\theta)d\theta \equiv \int q(0,\theta)d\theta$, the large deviation rate function for $R_{N,t,\omega}$ to be close to $q(\cdot,\theta)d\theta$ in $[0,T]$ is given by

$$I_{KOV}(q(\cdot,\cdot)) = I_{init}(q(\cdot,\cdot)) + I_{dyn}(q(\cdot,\cdot)). \tag{1.13}$$

Here $I_{init}(q(\cdot,\cdot))$ is the large deviation rate for the initial profile $q(0,\cdot)$ which depends on the choice of the initial measure. If we were to start with a Bernoulli product measure with $\rho_0(x/N)$ as the probability of a site x being occupied then

$$I_{init}(q(\cdot,\cdot)) = \int_{\mathbf{T}^d} h\left(q(0,\theta),\rho_0(\theta)\right) d\theta \tag{1.14}$$

where

$$h(a,b) = a\log\frac{a}{b} + (1-a)\log\frac{1-a}{1-b}. \tag{1.15}$$

The second piece, $I_{dyn}(q)$, is the dynamical part. It is infinite unless $\hat{I}(q(\cdot,\cdot)) < \infty$, where

$$\hat{I}(q(\cdot,\cdot)) = \int_0^T \int_{\mathbf{T}^d} \frac{|\nabla q|^2}{q(1-q)} d\theta dt. \tag{1.16}$$

For such q, one can consider the class $\mathcal{B}_q$ of $b(\cdot,\cdot)$ which satisfy

$$\frac{\partial q}{\partial t} = \frac{1}{2}\nabla\cdot D\nabla q - \nabla\cdot q(1-q)Db$$

and

$$\mathcal{E}[b] = \frac{1}{2}\int_0^T \int_{\mathbf{T}^d} q(1-q)b\cdot Db\, d\theta dt < \infty.$$

Then

$$I_{dyn}(q) = \inf_{b\in\mathcal{B}_q} \mathcal{E}[b]. \tag{1.17}$$

It could happen that $\mathcal{B}_q$ contains no b for which $\mathcal{E}[b] < \infty$ and then $I_{dyn}(q) = \infty$. It turns out that, if it is finite, the infimum in (1.17) is attained at some $b_0 = \nabla\phi$ for some $\phi : [0,T]\times\mathbf{T}^d \to \mathbf{R}$, and such a choice of b is unique.

We are now ready to define the rate function $\mathcal{I}(Q)$ for the empirical process. Given an arbitrary subprobability measure Q on $D\left[[0,T];\mathbf{T}^d\right]$, we can compute the marginal densities $q(t,\theta)$ for $0\le t\le T$. There is the initial rate $I_{init}(q(\cdot,\cdot))$ which depends on how the choice of the initial condition ξ is chosen. $I_{init}(q(\cdot,\cdot))$ would be given by (1.14) for a Bernoulli choice, or, if ξ is chosen deterministically, $I_{init}(q(\cdot,\cdot))$ would be infinite except when $q(0,\cdot) = \rho_0(\cdot)$, in which case it would be zero.

For each $b\in\mathcal{B}_q$ we have a well defined diffusion process P_q^b with one dimensional marginal densities $q(t,\theta)$ and with time dependent generator

$$\mathcal{A}_{q,b} = \frac{1}{2}\nabla\cdot S(q)\nabla + \frac{1}{2}(S(q)-D)\frac{\nabla q}{q}\cdot\nabla + (1-q)Db\cdot\nabla.$$

For each admissible q the existence and uniqueness of such a process has been established in [QV]. One can also verify that the relative entropy $H(Q,P_q^b)$ is either identically infinite for all $b \in \mathcal{B}_q$, or is finite for every member of $\mathcal{B}_q$. The rate function is infinite in the former case. In the latter case a special $b = b_Q$ is singled out in $\mathcal{B}_q$ by the property that

$$E^Q \left[\int_0^T \phi(t, x(t)) \cdot dx(t) \right] = E^{P^{b_Q}} \left[\int_0^T \phi(t, x(t)) \cdot dx(t) \right]$$

for all smooth functions $\phi : [0, T] \times T^d \to R^d$. Note that in general $b_Q \neq b_0$. The full rate function for the large deviations of the empirical process is given by

$$\mathcal{I}(Q) = I_{init}(q(\cdot, \cdot)) + H(Q\,, P^{b_Q}) + \mathcal{E}[b_Q].$$

Let $P_N \in \mathcal{M}_1\left(\mathcal{M}\left(D[[0,T];\mathbf{T}^d]\right)\right)$ be the distribution of the empirical process $R_{N,\omega}$. The topology is the usual topology of weak convergence for measures and the uniform topology on $D[[0,T];\mathbf{T}^d]$. We can now state the main result.

Theorem 1.1. *For $d \geq 3$, the empirical process P_N satisfies the large deviation principle with rate function $\mathcal{I}(Q)$. In particular, $\mathcal{I}$ is lower semicontinuous, has compact level sets. We have: For each closed set $\mathcal{C}$,*

$$\limsup_{N \to \infty} \frac{1}{N} \log P_N(\mathcal{C}) \leq - \inf_{Q \in \mathcal{C}} \mathcal{I}(Q),$$

and for each open set $\mathcal{O}$,

$$\liminf_{N \to \infty} \frac{1}{N} \log P_N(\mathcal{O}) \geq - \inf_{Q \in \mathcal{O}} \mathcal{I}(Q).$$

We will establish the large deviation principle by first studying the finite dimensional marginals of $R_{N,\omega}$ which correspond to the n-color version of the exclusion process. We then use certain superexponential estimates to pass from the finite dimensional distributions to the full empirical process.

Finally, a word about the organization of the material. In section 2, we provide precise statements of results and an outline of the methods of proof. Sections 3-6 deal with the n-color version of the exclusion process and the remaining sections deal with the transition from the finite dimensional marginals to the empirical process.

2. Notation and Summary.

Let Z_N denote the integers modulo N, i.e. $\{1, 2, \ldots, N\}$ with $N+1$ identified periodically with 1. Z_N^d is the d-fold product of Z_N. We have L tagged particles and the positions of these L particles are described by a point $(x_1, x_2, \ldots, x_L)$ in $(Z_N^d)^L$. Since we have simple exclusion, which means that at most one particle can occupy any one site at a given time, our effective state space is the part of $(Z_N^d)^L$ where $x_1, x_2, \ldots, x_L$ are distinct points of Z_N^d.

Our Markov process can be specified by the following infinitesimal generator. If $f(x_1, x_2, \ldots, x_L)$ is a function on the state space,

$$\tag{2.1} \begin{aligned} (L_N f)(x_1, x_2, \ldots, x_L) =& N^2 \sum_z \sum_{i=1}^{L} p(z)(1-\eta(x_i+z)) \\ & \Big(f(x_1, x_2, \ldots, x_i+z, \ldots, x_L) - f(x_1, x_2, \ldots, x_i, \ldots, x_L)\Big) \end{aligned}$$

Here the factor N^2 represents a speeding up or rescaling of time by a factor N^2. The probability distribution $p(\cdot)$ of the jump size is independent of N, is assumed to be symmetric, i.e. $p(z) = p(-z)$, and of finite range which means that $p(z) \neq 0$ only for $z \in F$, for some finite set F. We will also assume that the random walk with transition probabilities $p(\cdot)$ is irreducible. Finally if $d = 1$ we assume that $p(1) < 1/2$ to exclude the nearest neighbor case. The function $\eta(x)$ is the particle count at x

$$\tag{2.2} \eta(x) = \sum_{i=1}^{L} \chi_x(x_i)$$

Because our occupied sites are distinct, $\eta(x)$ is either 0 or 1 and the factor $(1 - \eta(x+z))$ in (2.1) indicates that a jump from x_i to $x_i + z$ of the i^{th} particle is allowed only when $\eta(x_i + z) = 0$, i.e, the site $x_i + z$ is free.

Since the generator is symmetric in the particles we can reduce the state space to just the set of occupied sites ignoring the labels of particles. The state space can then be the set of functions

$$\tag{2.3} \eta : Z_N^d \to \{0, 1\}$$

where $\eta(x)$ is given by (2.2) and tells us which sites are occupied. The reduced generator is given by

$$\tag{2.4} (L_N f)(\eta) = N^2 \sum_x \sum_z p(z)\eta(x)(1 - \eta(x+z)) \left(f(\eta^{x,x+z}) - f(\eta)\right)$$

where $\eta^{x,y}$ is the new configuration obtained by interchanging the situation at the sites x and y.

We may also decide that there are n types of particles distinguished by colors and may want to keep track of just colors and not individual labels. We then define

$$\tag{2.5} \eta_j : Z_N^d \to \{0, 1\}, \qquad \eta_j(x) = \sum_{i=1}^{L} \chi_x^j(x_i),$$

where χ_x^j counts the number of particles of color j at site x. Of course

$$\sum_{j=1}^{n} \eta_j(x) = \eta(x). \tag{2.6}$$

If we denote the typical element of the state space by

$$\zeta = (\eta_1, \ldots, \eta_n)$$

then the n-color generator is given by

$$(L_N^n)(\zeta) = N^2 \sum_x \sum_z p(z)\eta(x)(1 - \eta(x+z)) \left(f(\zeta^{x,x+z}) - f(\zeta)\right).$$

or in a more symmetric form by,

$$(L_N^n)(\zeta) = \frac{N^2}{2} \sum_x \sum_y p(y-x)[\eta(x)(1-\eta(y)) + (1-\eta(x))\eta(y)] \left(f(\zeta^{x,y}) - f(\zeta)\right) \tag{2.7}$$

Here $\zeta^{x,y}$ is the configuration obtained by interchanging the situation at x and y without changing the color of particles, if any, at these sites. Of course if there are particles at both x and y the prefactor kills this transition. If both the sites x and y are empty, then nothing happens. So the only things that can happen is that a particle jumps from x to y or from y to x depending on where the particle is, and the colors of particles never change in our model. More formally one can define $\zeta^{x,y}$ by

$$\begin{aligned} \zeta^{x,y}(u) &= \zeta(u) \quad \text{if } u \neq x \text{ or } y; \\ \zeta^{x,y}(x) &= \zeta(y); \\ \zeta^{x,y}(y) &= \zeta(x). \end{aligned} \tag{2.8}$$

We now imbed Z_N^d in $\mathbf{T}^d$, the d-dimensional torus by the map $y = x/N$. The L particles will evolve as a stochastic process

$$y(t) = (y_1(t), \cdots, y_L(t)). \tag{2.9}$$

The η process can be described by the measure–valued process

$$\nu_N(t) = \frac{1}{N^d} \sum_{i=1}^{L} \delta_{y_i(t)}, \tag{2.10}$$

on $\mathbf{T}^d$ of total mass $\frac{L}{N^d}$. If we denote the empirical process by

$$R_N = \frac{1}{N^d} \sum_{i=1}^{L} \delta_{y_i(\cdot)} \tag{2.11}$$

then $\nu_N(t)$ is the marginal at time t of R_N. The ζ process can, in a similar manner, be described by the measure–valued process with n components

$$\tilde{\nu}_N(t) = \left\{\nu_N^j(t);\, 1 \le j \le n\right\} \tag{2.12}$$

of total masses L_j/N^d, where L_j is the number of particles of color j. The measures $\nu_N^j(t)$ are defined by restricting the sum in (2.10) to particles of color j. Let us choose an arbitrary $T < \infty$ and fix it. We will consider all processes in the time interval $[0,T]$. $\nu_N(\cdot)$ is then viewed as a stochastic process with values in $D\left[[0,T],\mathcal{M}\right]$ of measure–valued functions on $[0,T]$. The large deviation theory for $\nu_N(\cdot)$ was established in [KOV]. Since we outlined it in the introduction we will not repeat it here. Our first task is to extend the large deviation results to $\tilde{\nu}_N(\cdot)$ of n-tuples of measure–valued functions on $[0,T]$.

In [KV] the diffusive behavior of a tagged particle in equilibrium was established for the simple exclusion process with self diffusion coefficients given by a density dependent matrix $S(\rho)$, which is a positive semidefinite symmetric matrix of size $d \times d$. Of course $S(\rho)$ really depends on ρ with $S(\rho) \to D$ as $\rho \to 0$ and $S(\rho) \to 0$ as $\rho \to 1$.

Let us denote by $\mathcal{C}$ the set of possible color densities, which is the convex set in $\mathbf{R}^n$ consisting of $\tilde{\rho} = (\rho_1, \rho_2, \ldots, \rho_n)$ such that $\rho_i \ge 0$ for $i = 1, 2, \ldots, n$ and $\rho = \sum_{i=1}^n \rho_i \le 1$.

We define the following matrices A and χ which are really to be thought of as $n \times n$ matrices corresponding to color types whose entries are $d \times d$ matrices that correspond to coordinates in $\mathbf{R}^d$. If j, j' are two color types then

$$\begin{aligned} A_{j,j'} &= A^{(1)}_{j,j'} + A^{(2)}_{j,j'} \\ &= S(\rho)\left(\rho_j \delta_{j,j'} - \frac{\rho_j \rho_{j'}}{\rho}\right) + D\left(\rho_j \rho_{j'} \frac{(1-\rho)}{\rho}\right) \end{aligned} \tag{2.13}$$

Each entry of $A_{j,j'}$, as a $d \times d$ matrix, is a linear combination of $S(\rho)$ and D with weights that have been given in (2.13). Similarly

$$\chi_{j,j'} = I\left(\frac{\delta_{j,j'}}{\rho_j} + \frac{1}{(1-\rho)}\right) \tag{2.14}$$

where I is the identity matrix of size $d \times d$. We will denote the inner products in $\mathbf{R}^n$, $\mathbf{R}^d$ and $\mathbf{R}^{nd}$ by $\langle\, ,\, \rangle$, $\cdot$ and $\langle\, ;\, \rangle$ respectively. We remark that if we define the entropy function

$$h(\tilde{\rho}) = \sum_{i=1}^n \rho_i \log \rho_i + (1-\rho)\log(1-\rho)$$

then

$$\lambda(\tilde{\rho}) = (\nabla h)(\tilde{\rho})$$

and χ is the Hessian

$$\chi = (\nabla \lambda)(\tilde{\rho}) = (\nabla\nabla h)(\tilde{\rho})$$

Let us note that although χ is undefined when $\rho = 1$, A and $A\chi$ are well defined on $\mathcal{C}$. In [Q1] it was shown that the empirical measures $\tilde{\nu}_N(\cdot)$ satisfy a law of large numbers. Assuming that the initial configurations, whether they are random or deterministic, satisfy

$$\tilde{\nu}_N(\cdot) \Rightarrow \tilde{\rho}^0(\theta)d\theta \tag{2.15}$$

in probability, in the weak topology on $\mathbf{T}^d$, it was shown that for all $t > 0$,

$$\tilde{\nu}_N(\cdot) \Rightarrow \tilde{\rho}(t,\theta)d\theta \tag{2.16}$$

The limit $\tilde{\rho}(t,\theta)$ is characterized as the unique solution of

$$\frac{\partial \tilde{\rho}}{\partial t} = \frac{1}{2}\nabla \cdot A\chi\nabla\tilde{\rho} \tag{2.17}$$

with initial condition

$$\tilde{\rho}(0,\theta) = \tilde{\rho}^0(\theta) \tag{2.18}$$

The equations (2.17) and (2.18) have the following interpretation. The sum $\rho(t,\theta) = \sum_{j=1}^n \rho_j(t,\theta)$ satisfies the heat equation

$$\frac{\partial \rho}{\partial t} = \frac{1}{2}\nabla \cdot D\nabla\rho \tag{2.19}$$

with initial condition

$$\rho(0,\theta) = \rho^0(\theta) = \sum_1^n \rho_j^0(\theta) \tag{2.20}$$

and, given the solution of (2.19), each component of (2.17) can be recast as

$$\frac{\partial \rho_j}{\partial t} = \mathcal{A}_\rho^* \rho_j \tag{2.21}$$

with $\rho_j(0,\theta) = \rho_j^0(\theta)$, where

$$\mathcal{A}_\rho^* f = \frac{1}{2}\nabla \cdot S(\rho(t,\theta))\nabla f + \frac{1}{2}\nabla \cdot [(D - S(\rho(t,\theta))] \frac{\nabla\rho(t,\theta)}{\rho(t,\theta)} f. \tag{2.22}$$

We will now describe the large deviation theory for the empirical measure–valued process $\tilde{\nu}_N(\cdot)$. If P_N is the probability distribution of $\tilde{\nu}_N(\cdot)$ it has a standard large deviation property with the rate function $I^n(\tilde{\rho}(\cdot,\cdot))$ defined in (2.33) for functions

$$\tilde{\rho}(\cdot) : [0,T] \times \mathbf{T}^d \to \mathcal{C} \tag{2.23}$$

such that $\int \rho_j(t,\theta)d\theta$ is independent of t for each j. Given a function $\mathbf{q}(\theta)$ on $\mathbf{T}^d$ taking values in symmetric positive semidefinite $n \times n$ matrices whose entries are $d \times d$ matrices and a function $\tilde{r}(\cdot) : \mathbf{T}^d \to \mathbf{R^n}$ with mean zero, we define

$$
(2.24) \qquad ||\tilde{r}||^2_{-1,\mathbf{q}(\cdot)} = \sup_{\Phi} \left\{ 2 \int_{\mathbf{T}^d} \langle \tilde{r}(\theta) \,, \Phi(\theta) \rangle \, d\theta - \int_{\mathbf{T}^d} \langle \nabla \Phi(\theta) \,; \mathbf{q}(\theta) \nabla \Phi(\theta) \rangle \, d\theta \right\}
$$

where the supremum is over C^∞ maps Φ from $\mathbf{T}^d$ to $\mathbf{R}^n$. Let us also define

(2.25)
$$
\hat{I}^n(\tilde{\rho}(\cdot,\cdot)) = \sup_{\Phi} \left\{ 2 \int_0^T \int_{\mathbf{T}^d} \langle (A\chi)(\tilde{\rho}) \nabla \tilde{\rho} \,; \nabla \Phi \rangle \, dt d\theta - \int_0^T \int_{\mathbf{T}^d} \langle \nabla \Phi \,; A(\tilde{\rho}) \nabla \Phi \rangle \, dt d\theta \right\}
$$

where the supremum is now taken over C^∞ functions on $[0,T] \times \mathbf{T}^d$ with values in $\mathbf{R}^n$. When it makes sense, we can choose $\Phi = \lambda(\tilde{\rho})$, so that $\nabla \Phi = \chi \nabla \tilde{\rho}$, and obtain

$$
(2.26) \qquad \hat{I}^n(\tilde{\rho}(\cdot,\cdot)) = \int_0^T \int_{\mathbf{T}^d} \langle \nabla \tilde{\rho} \,; \chi A \chi(\tilde{\rho}) \nabla \tilde{\rho} \rangle \, dt d\theta
$$

One can explicitly calculate $\chi A \chi$ and the quantity in (2.26) is finite if and only if

$$
\hat{I}(\tilde{\rho}(\cdot,\cdot)) + \sum_{j=1}^n \hat{I}_j(\tilde{\rho}(\cdot\,,\cdot)) < \infty,
$$

where $\hat{I}$ is given by

$$
(2.27) \qquad \hat{I}(\rho(\cdot\,,\cdot)) = \int_0^T \int_{\mathbf{T}^d} \frac{|\nabla \rho|^2}{\rho(1-\rho)} d\theta dt
$$

as in (1.16) and for $1 \le j \le n$,

$$
(2.28) \qquad \hat{I}_j(\tilde{\rho}(\cdot\,,\cdot)) = \int_0^T \int_{\mathbf{T}^d} (1-\rho) \frac{|\nabla \rho_j|^2}{\rho_j} d\theta dt.
$$

The equivalence follows from the following properties which were established in [V2]:

$$
(2.29) \qquad D - S(\rho) \le c\rho
$$

and

$$
(2.30) \qquad \delta(1-\rho) I \le S(\rho) \le (1-\rho) D
$$

where δ and c are two positive constants. The dynamical rate function $I^n_{dyn}(\tilde{\rho}(\cdot,\cdot))$ is defined to be $+\infty$ unless $\hat{I}^n(\tilde{\rho}(\cdot,\cdot)) < \infty$ and in such a case,

$$
(2.31) \qquad I^n_{dyn}(\tilde{\rho}(\cdot,\cdot)) = \frac{1}{2} \int_0^T \left\| \frac{\partial \tilde{\rho}}{\partial t} - \frac{1}{2} \nabla \cdot A \chi \nabla \tilde{\rho} \right\|^2_{-1, A(\tilde{\rho}(t,\cdot))} dt.
$$

If the initial conditions are deterministic and the law of large numbers holds initially with profile $\tilde{\rho}^0(\cdot)$ then

$$I^n_{init}(\tilde{\rho}(\cdot,\cdot)) = \begin{cases} 0 & \text{if } \tilde{\rho}(0,\cdot) = \tilde{\rho}^0(\cdot); \\ +\infty & \text{otherwise.} \end{cases} \tag{2.32}$$

For other suitable initial conditions we pick the corresponding rate function for the initial profile and this shall be our $I^n_{init}(\tilde{\rho}(\cdot,\cdot))$. The full rate function is given by

$$I^n(\tilde{\rho}(\cdot,\cdot)) = I^n_{init}(\tilde{\rho}(\cdot,\cdot)) + I^n_{dyn}(\tilde{\rho}(\cdot,\cdot)). \tag{2.33}$$

A major step towards the full large deviation principle for the empirical process is the following theorem proved in sections 3-6.

Theorem 2.1. *Assuming that the initial distribution satisfies a large deviation principle with a rate funcion I^n_{init}, the measures P_N on the space $E_n = D\left[[0,T];\mathcal{M}(\mathbf{T}^d)\right]^n$ satisfy a large deviation principle with rate function I^n given in (2.33). Moreover, I^n is lower semicontinuous, has compact level sets and is supported on the continuous maps in E_n. More precisely,*

$$\limsup_{N\to\infty} \frac{1}{N^d} \log P_N(C) \leq - \inf_{\tilde{\rho}(\cdot,\cdot)\in C} I^n(\tilde{\rho}(\cdot,\cdot)) \tag{2.34}$$

for closed $C \subset E_n$. For $G \subset E_n$ open,we have

$$\liminf_{N\to\infty} \frac{1}{N^d} \log P_N(G) \geq - \inf_{\tilde{\rho}(\cdot,\cdot)\in G} I^n(\tilde{\rho}(\cdot,\cdot)) \tag{2.35}$$

The transition from the above theorem to the large deviation result for the empirical process is in principle straightforward and is a matter of careful calculation and identification. We first prove a superexponential estimate for the empirical process $R_{N,\omega}$ that allows us to interchange limits freely and provides a cover for various technical points along the route. The rate function for the empirical process is then calculated by computing rate functions for finite dimensional projections. We can divide the torus $\mathbf{T}^d$ into a finite number of cells and the large deviation principle can be reduced to a large deviation principle for the proportions of particles that lie in a specified sequence of cells at time points $0 < t_1 < t_2 < \cdots < T$. This is a weak form of tagging and can be reduced to a problem with a finite but large number of colors by assigning different colors to particles with different histories. When the dust settles, it is a matter of identification of the rate function by chasing through the limiting procedure as the number of cells in the partition and the number of inspection times become infinite. This is carried out in section 9. The big surprise is the reemergence of entropy in the final formula and the physical interpretation that the rate function has because of it. This is not evident in the formula for the large deviation rate for n-colors. Finally let us remark that it suffices to prove the large deviation principle for deterministic initial conditions. Since we can integrate at the end with respect to the initial distribution, any large deviation property of the initial distribution can be incorporated at the very end and contributes a term I^n_{init}.

3. Superexponential estimates.

The proof of the large deviation principle for the n-color model follows essentially the methods of [Q1] and [Q2]. There are however some serious problems as $\rho \to 1$. In [Q1] while establishing the law of large numbers, these issues could be avoided because the total density ρ evolves according to the heat equation, and is bounded away from 1 for $t > 0$. In the context of large deviations, the density $\rho(t, \theta)$ can be 1 in a region of space and time, and still have a finite rate function. So the question of ρ being 1 in a region has to be faced.

In addition to the operator L_N^n for the evolution of our basic n-color process on Z_N^d we will consider the generator L^n on the entire lattice Z^d. We will also have occasion to consider the operator L_K^n which will operate in a finite box $|x| \le K$ disallowing any jumps ever to the exterior of the box. The evolution in the box with a given number of particles of each color can have certain pathological ergodicity properties. For example there are irreducible symmetric random walks involving a single particle in the plane that have no way of getting out of a bad corner in a cube of any size. Also, if the box is full, nothing can move and every color configuration is frozen: The ergodic invariant measures are clearly all the degenerate distributions with mass 1 at an arbitrary color configuration. On the other hand, if all nearest neighbor jumps can occur with positive probability, and if there is at least one empty site, then for boxes with sides larger than a fixed number that depends only on the number n of colors, the ergodic invariant measures are precisely the uniform distributions on all compatible configurations. In fact, it is always true that, for some ℓ_0, any invariant distibution for a box of size ℓ with at least one empty site, when restricted to smaller box of size $\ell - \ell_0$, is a mixture of $\pi_{\ell-\ell_0,\tilde{\rho}}$, the uniform distributions on configurations in the smaller box with given particle numbers for each color.

Let $f(\zeta)$ be a function that is constant on each ergodic component of the configurations in a box of size l, and let us fix the densities of each color in the box. This may not determine f but there will be functions $f^+(\tilde{\rho})$ and $f^-(\tilde{\rho})$ such that $f^+ \ge f \ge f^-$. Throughout this section and the next the number of colors n will be fixed and so we will drop all references to it in our notation.

It is convenient to symmetrize the operator to some extent and for this we introduce the definition

$$a_{x,y}(\zeta) = [\eta(x)(1 - \eta(y)) + \eta(y)(1 - \eta(x)] \tag{3.1}$$

and rewrite our operators in the form

$$L = \frac{1}{2} \sum_{x,y} p(y - x) a_{x,y}(\zeta)[F(\zeta^{x,y}) - F(\zeta)]$$

$$L_K = \frac{1}{2} \sum_{x,y \in B_K} p(y - x) a_{x,y}(\zeta)[F(\zeta^{x,y}) - F(\zeta)]$$

where B_K denotes the box $|x| \le K$.

There are two basic superexponential estimates, the one block and two block estimates that are important and have been proved in [Q1]. The one block estimate is:

Theorem 3.1. *Suppose $f(t,\theta,\zeta)$ is a smooth function of t and θ on $[0,t]\times\mathbf{T}^d$ which is a local function depending on a fixed box B and which vanishes identically in t and θ when the box is full. Let us denote by $\hat{f}(t,\theta,\tilde{\rho})$ the expected value of f under the Bernoulli measure $\pi_{\tilde{\rho}}$. Then*

$$\limsup_{K\to\infty}\limsup_{N\to\infty}\frac{1}{N^d}$$
$$\log E^{P_N}\left[\exp\Big\{\int_0^T\sum_x\Big|\frac{1}{(2K+1)^d}\sum_{y:|y-x|\le K}f(t,\frac{y}{N},\tau_y\,\zeta(t))-\hat{f}(t,\frac{x}{N},\tau_x\,\bar{\zeta}_K)\Big|dt\Big\}\right]\le 0.$$

Here $\bar{\zeta}_K$ refers to the average of $\zeta(x)$ over the box B_K and τ_x is the shift by x. The two block estimate is:

Theorem 3.2. *Suppose $f(t,\theta,\tilde{\rho},\tilde{\rho}')$ is a continuous function on $[0,T]\times\mathbf{T}^d\times\mathcal{C}\times\mathcal{C}$ that vanishes identically on $[0,T]\times\mathbf{T}^d$ when $\rho=\rho'=1$ or $\tilde{\rho}=\tilde{\rho}'$. Then*

$$\limsup_{\epsilon\to 0}\limsup_{K\to\infty}\limsup_{N\to\infty}\frac{1}{N^d}\log E^{P_N}\left[\exp\Big\{\int_0^T\sum_x f(t,\frac{x}{N},\tau_x\bar{\zeta}_{N\epsilon},\tau_x\bar{\zeta}_K)dt\Big\}\right]\le 0.$$

We now turn to some of the details specific to nongradient systems. Suppose $v(\zeta)$ is a local function depending only on configurations in some finite box B_l, which has mean zero with respect to every invariant measure for the evolution restricted to that box. This will of course force the function to vanish for every configuration that fills the box. We denote the class of such functions by $\mathcal{N}_l$. A function defined on a box B_l may have this property not for B_l but only for a larger box $B_{l'}$ in which case it will be in $\mathcal{N}_{l'}$. We denote by $\mathcal{N}$, the increasing union of $\mathcal{N}_l$ over l.

For any $v\in\mathcal{N}_l$ we form the sum

$$\sum_{|x|\le K-l}v(\tau_x\zeta)=V_K(\zeta)$$

and it defines a function on B_K that belongs to $\mathcal{N}_K$. We define the quantity $f_K(\zeta)$, for configurations of the box B_K that belong to the support of any extremal invariant measure to equal $\big\langle(-L_K)^{-1}V_K,V_K\big\rangle$ relative to that invariant measure. On the set of configurations where every site is filled we may take f_K to be zero. The function f_K of course will be constant on each ergodic component. We should expect it to be purely a function of densities $\tilde{\rho}$ of various colors in the box. However because of the possible breakdown in ergodicity we can only bound it above and below by such functions ${f_K}^{\pm}$. For $v\in\mathcal{N}$, it is easy to get an upper bound for f_K by writing the dual variational formula

$$f_K=\big\langle(-L_K)^{-1}V_K,V_K\big\rangle=\sup_u\Big\{2E^{\zeta}[u(\cdot)V_K(\cdot)]-\langle -L_K\,u,u\rangle\Big\}\tag{3.3}$$

We note that if $v \in \mathcal{N}_l$, it can be written as $L_K w$ for some B_l measurable w. Then

$$V_K(\zeta) = \sum_{|x| \le K-l} L_{B+x} \tau_x w$$

and it is straightforward to perform a summation by parts and obtain, using Schwarz's inequality,

$$f_K \le C(w)(2K+1)^d$$

Here $C(w)$ can be bounded by

$$C(w) \le C|B| \sup_{x,\zeta} E^\zeta\{[\tau_x w]^2\}$$

where x ranges over allowable translates i.e. $|x| \le K - l$, and E^ζ ranges over all extremal invariant measures for the box of size K. In particular

$$\limsup_{K\to\infty} \frac{1}{(2K+1)^d} \sup_\zeta f_K(\zeta) \le C(w)$$

A subset $\mathcal{V}_l \subset \mathcal{N}_l$ of these functions v are of the form

$$\sum_r u_r(\eta) v_r(\zeta) \tag{3.2}$$

for some functions u_r that are "color blind" and functions v_r that are linear functions of the color variables $\zeta(x)$. This class of functions is preserved under the action of the generator. In [Q1] it is shown that the spectral gap of the generator, acting on this smaller class of functions, on a box of side length l with either periodic or reflecting boundary conditions, decays as l^{-2}. In the future, for the most part, we will restrict v to the class $\mathcal{V}$ which is defined as the (increasing) union of $\mathcal{V}_l$ over l. It was shown in [Q1] that

$$\lim_{K\to\infty} \frac{1}{(2K+1)^d} f_K{}^{\pm}(\tilde\rho) = [v, v]_{\tilde\rho} \tag{3.3}$$

exists locally uniformly on $\mathcal{C}^0$ defined as the subset of the color densities $\mathcal{C}$ where $\rho < 1$.

To carry out these calculations we need to introduce certain Hilbert spaces. For any $\tilde\rho$ in $\mathcal{C}^0$ let $\pi_{\tilde\rho}$ be the Bernoulli measure on the configurations in Z^d. Let $g(\zeta)$ be a local function and consider the formal sum

$$G(\zeta) = \sum_{x \in Z^d} g(\tau_x \zeta). \tag{3.4}$$

While the sum G clearly does not exist, the quantities

$$(\nabla_{0,z} G)(\zeta) = G(\zeta^{0,z}) - G(\zeta) = \xi_z \tag{3.5}$$

exist and involve only a finite sum. We consider the collection $\{\xi_z\}$ as some sort of gradient and view them as elements in a Hilbert space $H^0_{\tilde\rho}$ with respect to the norm

$$\langle\!\langle \xi, \xi \rangle\!\rangle_{\tilde\rho} = \frac{1}{4} E^{\pi_{\tilde\rho}} \left[\sum_z p(z) a_{0,z}(\zeta) |\xi_z|^2 \right]. \tag{3.6}$$

A gradient $\{\xi_z\}$ created in this way (3.5) will possess certain natural algebraic properties. For example, if for some k a finite sequence $(x_i, x_i + z_i)$, $i = 1, \dots, k$ of bonds forms a closed loop in Z^d, then $\sum_{i=1}^k \tau_{x_i} \xi_{z_i} = 0$. One can then define abstract gradients $\{\xi_z\}$ that possess these same algebraic properties. Under the same norm, this will define a slightly larger Hilbert space $H_{\tilde\rho}$ that is exactly nd dimensions larger than $H^0_{\tilde\rho}$ and a relative (non-orthogonal) complement to $H^0_{\tilde\rho}$ is spanned by $\boldsymbol{\Psi} = \{\psi^{i,j}\}$,where

$$\psi^{i,j}_z = \langle z, e_i \rangle \left[\eta^j(z) - \eta^j(0) \right] \tag{3.7}$$

for $1 \le i \le d$ and $1 \le j \le n$.

Now we discuss how to compute $[v, v]_{\tilde\rho}$. The general v is in fact a linear combination of three basic types and we discuss the computation type by type. If we have a function v of the form

$$v = Lg \tag{3.8}$$

where g is a local function which is linear in color, i.e. of the form (3.2), then one can compute $[v, v]_{\tilde\rho}$ as the norm in $H_{\tilde\rho}$ of the image ξ of v through the map $g \to G \to \xi$ defined in (3.4), (3.5) and one finds that

$$[v, v]_{\tilde\rho} = \langle\!\langle \xi, \xi \rangle\!\rangle_{\tilde\rho} \tag{3.9}$$

On the other hand we can have a v which under an appropriate shift takes the form,

$$v = \sum_z a_{0,z}(\zeta) p(z) [F(\zeta^{0,z}) - F(\zeta)] \tag{3.10}$$

where F is a local function linear in color. Let us briefly consider the space $\bar H_{\tilde\rho}$ of functions $\{h_z\}$ that do not necessarily satisfy the algebraic relations, but equipped with the same norm

$$\langle\!\langle h, h \rangle\!\rangle_{\tilde\rho} = \frac{1}{4} E^{\pi_{\tilde\rho}} \left[\sum_z p(z) a_{0,z}(\zeta) |h_z|^2 \right]. \tag{3.11}$$

Such an $\{h_z\}$ can be projected orthogonally into $H_{\tilde\rho}$ to get a $\{\xi_z\}$. If v is of the form (3.10) then for any $\tilde\rho \in \mathcal{C}^0$ the way to compute the value of $[v, v]_{\tilde\rho}$ is as follows: First associate with v the following $\{h_z\}$ in $\bar H_{\tilde\rho}$:

$$h_z(\zeta) = -2[F(\zeta^{0,z}) - F(\zeta)]. \tag{3.12}$$

Then project $\{h_z\}$ into $H_{\tilde\rho}$ and obtain $\{\xi_z\}$. Finally $[v,v]_{\tilde\rho}$ is calculated as $\langle\langle\xi,\xi\rangle\rangle_{\tilde\rho}$. One can verify that this is consistent with the map $g \to G \to \xi$ defined earlier.

The "current" term $\mathbf{W}$ given by

$$W^{i,j} = \frac{1}{2}\sum_z a_{0,z}(\zeta)p(z)\langle z, e_i\rangle(\eta_j(z)-\eta_j(0)) \tag{3.13}$$

is easily seen to be imbedded in this manner in $H_{\tilde\rho}$ as $\mathbf{\Psi} = \{\psi_z^{i,j}\}$ where the $\psi_z^{i,j}$ are defined in (3.7) for $1 \le i \le d$ and $1 \le j \le n$.

The following calculation was carried out in [Q1]: For $\tilde\rho$ in $\mathcal{C}^0$ and any matrix B

$$\inf_{\xi\in H^0_{\tilde\rho}} E^{\pi_{\tilde\rho}}[\frac{1}{4}\sum_z a_{0,z}(\zeta)p(z)(\langle B\,;\mathbf{\Psi_z}\rangle-\xi_z)^2] = \frac{1}{2}\langle B\,;A(\tilde\rho)B\rangle \tag{3.14}$$

Although the result was proved for $\tilde\rho$ in the interior $\mathcal{C}^0$ where $\rho < 1$, we have an easy estimate that all the norms concerned for v in $\mathcal{V}$ are bounded by a multiple of $1-\rho$. In fact, everything extends to the boundary easily except the following family of density gradients $\nabla\zeta$ which are *not* in $\mathcal{V}$:

$$\nabla_e\zeta = \{\eta_j(e_i)-\eta_j(0);\quad 1\le i\le d, 1\le j\le n\} \tag{3.15}$$

It was shown in [Q1] that these can be imbedded in $H_{\tilde\rho}$ as well, for $\rho < 1$, as

$$\mathbf{f} = \{f_{i,j}\} \tag{3.16}$$

and in fact belong to the orthogonal complement of $H^0_{\tilde\rho}$ providing us a natural nd dimensional subspace orthogonal to $H^0_{\tilde\rho}$. A simple calculation using L_K^{-1} establishes

$$[\langle b\,;\mathbf{W}\alpha\rangle, \langle b'\,;(\nabla\zeta)\alpha'\rangle]_{\tilde\rho} = \left\langle b\otimes\alpha\,;\chi^{-1}(\tilde\rho)\; b'\otimes\alpha'\right\rangle$$

Since χ is nonsingular this proves that modulo $H^0_{\tilde\rho}$

$$\mathbf{\Psi} = \frac{1}{2}A\chi\,\mathbf{f} \tag{3.18}$$

which essentially captures the diffusion law.

Finally there are some large deviation results that are proved in [Q1] that are restated here for convenience. Suppose v is a function is class $\mathcal{N}$. Let $Q(\tilde\rho)$ be a continuous function on $\mathcal{C}$ that satisfies

$$\limsup_{\substack{K\to\infty\\ \zeta\sim\tilde\rho}} \frac{1}{(2K+1)^d} f_K(\zeta) \le Q(\tilde\rho)$$

In particular if $v\in\mathcal{V}$ we can take $Q(\tilde\rho) = \|v\|^2_{\tilde\rho} = [v,v]_{\tilde\rho}$ which defines a continuous function on $\mathcal{C}^0$ and because it is dominated by $1-\rho$ it extends continuously to $\mathcal{C}$.

Theorem 3.3. *If we denote by $\hat{E}$ the expectation with respect to any equilibrium process then*

$$\limsup_{K\to\infty}\limsup_{N\to\infty} \quad \frac{1}{N^d}\log\ \hat{E}[\exp\{N\int_0^T \sum_x \tau_x\ v\ ds - \int_0^T \sum_x \tau_x Q(\bar{\zeta}_K)\}] \leq 0$$

where $\bar{\zeta}_K$ is the average of ζ over the box of side length $(2K+1)$.

Remark 1. Since the density of any deterministic initial condition is bounded by C^{N^d} for some constant C, for arbitrary initial conditions the above theorem is valid with a universal constant C replacing zero on the right as a bound for the limsup.

Remark 2. The theorem is valid in exactly the same form if v is replaced by $v(t, \frac{x}{N}, \tau_x\zeta)$ provided it depends continuously on t and θ as they range over $[0,T]\times\mathbf{T}^d$ and takes values in $\mathcal{V}_l$ for some value of l.

Remark 3. The right hand side of (3.14) is a continuous function of $\tilde{\rho}\in\mathcal{C}^0$, the subset of color densities with $\rho<1$. If we limit the infimum to ξ that come from a g that is local and defined in a box of size l and then increase l, we get a nondecresing sequence of continuous functions that converges on $\mathcal{C}^0$ to a continuous limit. By Dini's theorem the convergence is uniform. It now follows easily that the minimizers ξ_z that are defined on different spaces, have the property

$$E^{P_{\tilde{\rho}}}\Big[\sum_z a_{0,z}(\zeta)p(z) < B_z(\zeta)\,;\,\xi_z > \Big]$$

are continuous on $\mathcal{C}^0$ provided B_z are local functions.

4. Upper and Lower Bounds for the n-color system

Like many large deviation results, the upper bound depends on our ability to estimate $E[\exp\{N^d F_\alpha(\zeta(\cdot))\}]$ for a suitable family of functions F_α. It is important that the functions F_α have representations in terms of the measure-valued process $\tilde{\nu}(\cdot)$ that makes them bounded and continuous in the topology of weak (as measures on $\mathbf{T}^d$) uniform convergence on the interval $[0,T]$. Although we can, and will, use many of the exponential martingales that are available, we cannot do so directly because of the above restriction. The plan for the upper bound is then to exhibit a family of functionals $F_\alpha(\zeta(\cdot))$ that are representable as bounded continuous functionals $G_\alpha(\tilde{\nu}(\cdot))$ of the measure valued process $\tilde{\nu}(\cdot)$ and have the property

$$\limsup_{n\to\infty} \frac{1}{N^d} \log E[\exp\{N^d F_\alpha(\zeta(\cdot))\}] \leq 0$$

This will then give immediately an upper bound for compact sets with the rate function $I(\tilde{\rho}(\cdot\,,\cdot))$, if we can show that $I(\tilde{\rho}(\cdot\,,\cdot))$ is lower semicontinuous and satisfies

$$I(\tilde{\rho}(\cdot\,,\cdot)) \leq \bar{I}(\tilde{\rho}(\cdot\,,\cdot)) = \sup_\alpha G_\alpha(\tilde{\rho}(\cdot))$$

We will often misuse the notation and not distinguish between F and G. If we have superexponential tightness estimates, then the upper bound will extend from compact sets to closed sets. The lower semicontinuous functional I should be shown to have compact level sets.

The lower bounds depend on establishing laws of large numbers, in hydrodynamic scaling, for certain weak perturbations and controlling by the Girsanov formula the relative entropies of the processes that go with these perturbations. Since the methods for establishing laws of large numbers are fairly close to the ones for providing upper bounds for large deviations, we prefer to establish the upper bounds for a class of models that includes the weak perturbations that we will use for the lower bound.

Our basic infinitesimal generator L_N^n is modified by a time dependent lower order term A_N^n given by

$$A_N^n f(\zeta) = \frac{N}{2} \sum_{x,z} a_{x,x+z}(\zeta) p(z) b(t, \frac{x}{N}, z, \tau_x\zeta)(f(\zeta^{x,x+z}) - f(\zeta)) \tag{4.1}$$

The new term b provides a correction to the old jump rate $p(z)$, and the modified jump rate is of the form $p(z)(1 + \frac{1}{N} b(t, \frac{x}{N}, z, \tau_x\zeta))$. The functions $b(t,\theta,z,\cdot)$ are all assumed to be local functions of ζ that depend on the configuration in some fixed finite box. Moreover, for each jump size z, and each configuration ζ, they are assumed to be continuous functions of t and θ on $[0,T] \times \mathbf{T}^d$. We will also assume that they have a uniform bound.

We will denote by $P_{N,b}$ the measure corresponding to the n-color process on $[0,T]$ with generator $L_N^n + A_N^n$ and with arbitrary initial conditions. $E_{N,b}$ will denote expectation with respect to $P_{N,b}$. We will continue to denote by the same $P_{N,b}$ the induced measure on the space of measure–valued functions $\tilde{\nu}(\cdot)$ on $[0,T]$. If b is 0 it will be omitted. An elementary calculation with the Girsanov formula for the Radon-Nikodym derivative of $P_{N,b}$ with respect to P_N provides an upper bound for any L_p norm of this Radon-Nikodym derivative. In particular, since we have assumed a uniform bound for b, one can easily prove the following lemma.

Lemma 4.1. *Uniformly with respect to all initial conditions, we have*

$$\lim_{p\to 1}\limsup_{N\to\infty}\frac{1}{N^d}\log E_N[|\frac{dP_{N,b}}{dP_N}|^p]\leq 0.$$

Remark. We could also take the initial condition to be some reasonable equilibrium distribution. Changing over from such an equilibrium to any initial condition again involves a Radon-Nikodym derivative, for the initial condition, that clearly is bounded uniformly by $\exp[CN^d]$. In particular superexponential estimates obtained in equilibrium for the unperturbed process remain valid for all initial conditions and for all of our perturbations by a simple application of Hölder's inequality.

We will prove in section 7 a superexponential tightness estimate for the family Q_N of measures on the space of processes. This will automatically yield a similar estimate for $P_{N,b}$. We denote by $D[[0,T],\mathcal{M}]$ the space of measure-valued maps with discontinuities of the first kind only. E_n denotes the space of n-tuples of such maps on which the measures $P_{N,b}$ live and where the large deviation results are to be established. Moreover by the simple exclusion property we have an apriori bound of 1 for the total density. The rate function for the large deviation will then live on the space $\mathcal{D}_n$ of weakly continuous L_1 maps $\tilde\rho$ from $[0,T]\times\mathbf{T}^d$ into $\mathcal{C}=\{(\rho_\infty,\dots,\rho_\setminus),\rho_\rangle\geq\prime,\rho=\sum_{\rangle=\infty}^{\setminus}\rho_\rangle\leq\infty$. We will need to prove certain preliminary estimates in order to limit the support of the rate function. We define $\mathcal{D}_n^0$ as those maps for which $\hat I^n(\tilde\rho(\cdot\,,\cdot))$, given by (2.26), is finite or equivalently $\hat I(\tilde\rho(\cdot\,,\cdot))$ and $\{\hat I_j(\tilde\rho(\cdot\,,\cdot));1\leq j\leq n\}$, given by (2.27) and (2.28), are finite.

We define the rate function $I_{dyn}^{n,b}(\tilde\rho(\cdot\,,\cdot))$ by

(4.2)
$$I_{dyn}^{n,b}(\tilde\rho(\cdot\,,\cdot))=\begin{cases}\frac{1}{2}\int_0^T\|\frac{\partial\tilde\rho}{\partial t}-\frac{1}{2}\nabla\cdot(A\chi)(\tilde\rho)\nabla\tilde\rho-\nabla\cdot H\|^2_{-1,A(\tilde\rho(t,\cdot))}\,dt & \text{if}\quad \tilde\rho(\cdot\,,\cdot)\in\mathcal{D}_n^0\\ =+\infty & \text{if}\quad \tilde\rho(\cdot\,,\cdot)\notin\mathcal{D}_n^0.\end{cases}$$

Here $H=H_{i,j}(t\,,x,\,\tilde\rho)$ depends in a nontrivial way on the perturbation b. In fact $H=\hat{\mathbf{a}}[\mathbf{b}]$ given by (4.51), but for now we do not need the specific form. All we need is that H is continuous in $\tilde\rho$ and satisfies the bound

$$\sup_{t,x,j}|H_{i,j}(t\,,x\,,\tilde\rho)|\leq C\rho_i(1-\rho). \tag{4.3}$$

For $\tilde\rho(\cdot\,,\cdot)\in\mathcal{D}_n^0$, by duality, $I_{dyn}^{n,b}(\tilde\rho(\cdot\,,\cdot))$ has a variational formula,

(4.4)
$$I_{dyn}^{n,b}(\tilde\rho(\cdot\,,\cdot))=\sup_\Phi\Big\{\int_0^T\int_{\mathbf{T}^d}\Big[-\Big\langle\frac{\partial\Phi}{\partial t}\,,\tilde\rho\Big\rangle+\frac{1}{2}\langle\nabla\Phi\,;(A\chi)(\tilde\rho)\nabla\tilde\rho\rangle+\langle\nabla\Phi\,;H\rangle\Big]\,dt\,dx$$
$$-\frac{1}{2}\int_0^T\int_{\mathbf{T}^d}\langle\nabla\Phi\,;A(\tilde\rho)\nabla\Phi\rangle\;dt\,dx\Big\}$$

which makes sense for $\tilde\rho\in\mathcal{D}_n^0$, because χ^{-1} is bounded.

The space of functions $\tilde\rho$ with a uniform bound on their rate functions enjoy certain compactness properties that will be used in several occasions.

Lemma 4.2. *Let ρ_n and R_n be two sequences of functions such that for all n,*

$$\int_0^T \int \left\{ |\nabla\rho_n|^2 + (1-\rho_n)|\nabla R_n|^2 \right\} dx dt + \int_0^T \|\frac{\partial \rho_n}{\partial t}\|_{-1}^2 + \|\frac{\partial R_n}{\partial t}\|_{-1}^2 dt \le C.$$

Suppose $0 \le \rho, R \le 1$ and that $\rho_n \rightharpoonup \rho$, $R_n \rightharpoonup R$ weakly in $L^2([0,T]\times \mathbf{T}^d)$. Then $\rho_n \to \rho$ and $(1-\rho_n)R_n \to (1-\rho)R$ strongly in $L^1([0,T]\times \mathbf{T}^d)$.

Proof: Let $\phi(x)$ be a smooth mollifier and put $\rho_n^\delta = \rho_n * \phi^\delta$, where $\phi^\delta(x) = \delta^{-d}\phi(\frac{x}{\delta})$ and the convolution is in the x variable. The gradient estimate on ρ_n implies

$$\lim_{\delta\to 0} \sup_n \int_0^T \int |\rho_n^\delta - \rho_n| dx dt = 0. \tag{4.5}$$

On the other hand, the H_{-1} bound on the $\partial_t \rho_n$ implies that for each δ, the full gradient $(\partial_t \rho_n^\delta, \nabla \rho_n^\delta)$ is uniformly bounded in n. From this and the weak convergence of ρ_n to ρ we deduce that ρ_n^δ converges strongly to $\rho^\delta = \rho * \phi^\delta$. We can now let δ go to zero and use (4.5) to conclude the strong convergence of ρ_n to ρ.

For any integrable function g, let $u_n(g,t,x) = \int R_n(t,y)g(y,t)\phi^\delta(x-y)dy$. If g is smooth, the H_{-1} bound on $\partial_t R_n$ and the first assertion of the lemma imply the L^1 convergence of $u_n(g,t,x)$ to $u(g,t,x) = \int R(y,t)g(y,t)\phi^\delta$. Since

$$\int_0^T \int |u_n(g') - u_n(g)| dx dt \le \int_0^T \int |g' - g| dx dt,$$

we also have that $u_n(g)$ converges strongly to $u(g)$ for every integrable g. Repeating the same reasoning we deduce that if g_n converges to g in L^1 sense, then $u_n(g_n)$ converges to $u(g)$ in L^1 sense too. We now choose $g_n = 1-\rho_n$. Therefore we have the following L^1 convergence:

$$\lim_{n\to\infty} [(1-\rho_n)R_n] * \phi^\delta = [(1-\rho)R] * \phi^\delta.$$

On the other hand, we have a uniform bound on the L^2 norm of the spatial gradient of the sequence $(1-\rho_n)R_n$ which allows us to send δ to zero and drop the convolution with ϕ^δ.

Theorem 4.3. *There are constants C_1, C_2 and C_3 such that the rate function $I_{dyn}^{n,b}(\tilde\rho(\cdot,\cdot))$ defined by (4.3) satisfies the bounds*

$$\hat I^n(\tilde\rho(\cdot\,,\cdot)) \le C_1 + C_2 I_{dyn}^{n,b}(\tilde\rho(\cdot\,,\cdot)) + C_3 \int_0^T \int_{\mathbf{T}^d} \left\langle H\,; A^{-1}(\tilde\rho) H \right\rangle \,dt\,dx \tag{4.6}$$

and

$$\int_0^T \|\frac{\partial\tilde\rho}{\partial t}\|_{-1,A(\tilde\rho(t\,,\cdot))}^2 dt \le C_1 + C_2 I_{dyn}^{n,b}(\tilde\rho(\cdot\,,\cdot)) + C_3 \int_0^T \int_{\mathbf{T}^d} \left\langle H\,; A^{-1}(\tilde\rho) H \right\rangle \,dt\,dx \tag{4.7}$$

Moreover the function $I_{dyn}^{n,b}$ is a lower semicontinuous function with compact level sets. Finally, for the measures $P_{N,b}$ the large deviation rate function is infinite outside $\mathcal{D}_n^0$.

Proof: The function $h(\tilde\rho) = \sum_i \rho_i \log \rho_i + (1-\rho)\log(1-\rho)$ is bounded on $\mathcal{C}$ and an easy calculation of

$$\frac{\partial}{\partial t}\int_{\mathbf{T}^d} h(\tilde\rho(t,x))dx$$

followed by an integration by parts yields the bound (4.6) and (4.7) follows from (4.6). The bounds (4.6) and (4.7) are sufficient to establish compact level sets. We next prove the lower semicontinuity of $I_{dyn}^{n,b}(\tilde\rho(\cdot\,,\cdot))$. Let $\tilde\rho^k(t\,,\cdot)) \Rightarrow \tilde\rho(t\,,\cdot))$ uniformly in t. Under the assumption that $I_{dyn}^{n,b}(\tilde\rho^k(\cdot\,,\cdot))$ is uniformly bounded by some constant, we will establish that $A\chi(\tilde\rho^k(\cdot\,,\cdot))\nabla\tilde\rho^k$, $A(\tilde\rho^k(\cdot\,,\cdot))$ and $H(\tilde\rho^k(\cdot,\cdot))$ converge weakly in $L^2([0,T]\times\mathbf{T}^d)$ to $A\chi(\tilde\rho(\cdot\,,\cdot))$, $A(\tilde\rho(\cdot\,,\cdot))$ and $H(\tilde\rho(\cdot,\cdot))$ respectively. Then we can take the limit in (4.4) and the variational formula will establish the lower semicontinuity. The formula for $A(\tilde\rho)$ is given in (2.13). An explicit calculation yields

$$[A\chi(\tilde\rho)\nabla\tilde\rho]_i = S(\rho)\nabla\rho_i + \frac{\rho_i}{\rho}(D - S(\rho))\nabla\rho. \tag{4.8}$$

By observation of these various functions it is clear that we need: 1. Strong convergence of the total densities $\rho^k(\cdot\,,\cdot)$ to $\rho(\cdot\,,\cdot)$, 2. Weak convergence of $(1-\rho^k)\nabla\rho_i^k$ to $(1-\rho)\nabla\rho_i$, 3. Strong convergence of $\rho_i^k(1-\rho^k)$ to $\rho_i(1-\rho)$. By (4.3) we have a uniform bound on $\langle H\,; A^{-1}H\rangle$. Let us note that the bounds (4.6) and (4.7) imply uniform bounds on

$$\int_0^T\int_{\mathbf{T}^d}\frac{|\nabla\rho^k|^2}{\rho^k(1-\rho^k)}\,dt\,dx \quad\text{and}\quad \int_0^T \|\frac{\partial\rho^k}{\partial t}\|_{-1}^2\,dt$$

implying 1. While we do not have quite the same bounds on the densities of individual colors, we do have bounds on

$$\int_0^T\int_{\mathbf{T}^d}\frac{(1-\rho^k)}{\rho^k}|\nabla\rho_i^k|^2\,dt\,dx \quad\text{and}\quad \int_0^T \|\frac{\partial\rho_i^k}{\partial t}\|_{-1}^2\,dt$$

implying 2 and 3.

We now turn to proving that the large deviation rate for $P_{N,b}$ is infinite outside $\mathcal{D}_n^0$. Part of it, namely that $\hat I$ given by (2.27) has to be finite was established in [KOV]. We now concentrate on (2.28). Let e be a coordinate direction and let $f = \tau_e F - F$ be the difference between a local function F and its shift in one of the coordinate directions by one step. We shall assume that the local function F depends on a finite box, is linear in color, and vanishes identically for all configurations that fill that box. It is easily seen that f, which is then defined in a slightly larger box, is from the class $\mathcal{V}$ and can in fact be represented in the form $f = L_B\, g$ for some g and and some box B. If f vanishes identically on a set A of configurations inside the box B which is invariant under permutations inside B, such a set will be an ergodic component under L_B and g can be taken to be zero on A.

Moreover, if f is bounded by 1, there is a constant C that depends only on the size of the box B that will bound g.

In the next few steps, as we proceed with the reasoning, this constant may change but will always be replaced by another constant that only depends on the size of the box B. We shall consider the expectation $\hat{E}$ in equilibrium :

$$\hat{E}[\exp\{N\int_0^T\sum_x \Phi(t,\frac{x}{N})\tau_x f(\zeta(t))dt - C\int_0^T\sum_x \tau_x\chi_A(\zeta(t))|\Phi(t,\frac{x}{N})|^2dt\}]. \tag{4.9}$$

Because f is of the form $f = L_B g$, the variational quantity

$$\sup_\phi\Big\{E^\phi[N\sum_x\Phi(t,\frac{x}{N})\tau_x f(\zeta) - C\int_0^T\sum_x\tau_x\chi_A(\zeta(t))|\Phi(t,\frac{x}{N})|^2dt \\ - \frac{N^2}{4}\sum_{x,y}a_{x,y}(\zeta)p(y-x)\frac{1}{\phi}|\nabla_{x,y}\sqrt{\phi}|^2]\Big\}$$

is seen to be nonpositive for a suitable choice of C. This yields a bound of 1, by an application of Feynman- Kac formula, to the quantity (4.9). After performing a summation by parts we can apply Theorems 3.1 and 3.2. The remarks about the Radon-Nikodym derivatives now allow us to construct functions of the form

$$F_\alpha(\zeta(\cdot)) = \int_0^T\frac{1}{N^d}\Big[\sum_x(\nabla_e\Phi)(t,\frac{x}{N})\hat{F}(\tau_x\bar{\zeta}_{N\epsilon}) - C\sum_x\hat{\chi}_A(\tau_x\bar{\zeta}_{N\epsilon})|\Phi(t,\frac{x}{N})|^2\Big]dt$$

that depend on $\alpha = \{\Phi,\epsilon\}$, that are weakly continuous and correspond to the functionals

$$G_\alpha(\tilde{\rho}(\cdot)) = \int_0^T\int_{\mathbf{T}^d}\Big[(\nabla_e\Phi)(t,\theta)\hat{F}(\tilde{\rho}_\epsilon(t,\theta)) - C\,\hat{\chi}_A(\tilde{\rho}_\epsilon(t,\theta))|\Phi(t,\theta)|^2\Big]\,dt\,d\theta$$

Here $\tilde{\rho}_\epsilon$ represents the convolution of $\tilde{\rho}$ with the normalized indicator function of a box of size ϵ on the torus $\mathbf{T}^d$. The functionals have the property that for some fixed $0<\lambda<1$ and $C<\infty$, that are independent of α

$$\limsup_{\epsilon\to 0}\limsup_{N\to\infty}\frac{1}{N^d}\log E\left[\exp\left\{N^d\,\lambda\,F_\alpha(\zeta(\cdot))\right\}\right]\le C$$

for every Φ. This provides a lower bound for the large deviation rate $\bar{I}$.

$$\bar{I}(\tilde{\rho}(\cdot,\cdot)) \ge \sup_\Phi\lim_{\epsilon\to 0}\{\lambda\,G_\alpha(\tilde{\rho}_\epsilon(\cdot,\cdot)) - C\} \tag{4.10}$$

Now we take F to be $F^l_j = \chi_l(\zeta)\eta_j(0)$ where χ_l is the indicator of those configurations with at least one empty site in the box of side $2l+1$ around the origin. While we will let $l\to\infty$ later on, for the present proof we only need some l so that for all i,j, $\nabla_{e_i}F^l_j$ are in $\mathcal{V}_{l'}$ for some l' and then hold l and l' fixed. F^l_j vanishes unless there is at least one empty site and at least one particle of color j in the bigger box of size l'. Taking A to be the set of these configurations we get the inequality $\hat{\chi}_A(\tilde{\rho}) \leq C\rho_j(1-\rho)$. Moreover it is seen that $\hat{F}^l_j(\tilde{\rho}) = E_{\pi_{\tilde{\rho}}}[F^l_j] = \rho_j c_l(\rho)$ where $c_l(\rho)$ and $\frac{dc_l(\rho)}{d\rho}$ behave like $(1-\rho)$ and 1 as $\rho \to 1$. Since we know (2.27) holds, now (2.28) follows as well.

We now turn towards proving upper bounds. We need to prove them only at those points that are in $\mathcal{D}^0_n$. If $U(t,\zeta)$ is any function of time and the n–color configuration then

$$\begin{aligned}\exp\{U(t,\zeta(t)) - U(0,\zeta(0)) - &\int_0^t U_s(s,\zeta(s))ds \\ &- N^2 \int_0^t (\exp[-U]\ [L^n_N + A^n_N]\ \exp[U])(\zeta(s))ds\}\end{aligned} \tag{4.11}$$

is a martingale with respect to the n-color process $P_{N,b}$. We can, without loss of generality, assume that the initial condition is deterministic and leads to a profile $\tilde{\rho}^0$ of densities at time 0. The basic U's that we want to look at are of the form

$$U(t,\zeta) = \sum_x \left\langle \Phi(t,\frac{x}{N}),\zeta(x)\right\rangle - \frac{1}{N}\sum_x g(t,\frac{x}{N},\tau_x\zeta) \tag{4.12}$$

where Φ is a continuously differentiable function on $[0,T]\times \mathbf{T}^d$ with values in $\mathbf{R}^n$ and g is a local function depending continuously on $[0,T]\times\mathbf{T}^d$. We will later pick g to be of the form $g = \sum_{i,j}(\nabla\Phi)_{i,j}g_{i,j}$. In addition to Φ and g and the functions F^l_j that depend on the box size l, we will have certain auxiliary parameters: four small positive numbers $\epsilon_1,\epsilon_2,\epsilon,\delta$, and a continuous map J from $\mathcal{C}$ into $N\times d$ matrices. We will assume that J vanishes if $\rho = 1$.

Assume that F is a local function defined on some finite box, is linear in color, and is zero if the box is full. Let us define $\nabla_e F = \tau_e F - F$ where e is a coordinate direction and $\bar{F}_{x,k}$ as the average of $\tau_y F$ for y in a box of side $2k+1$ centered at x. For any function G we define the lattice gradient in the dirction e as

$$\nabla_{N\epsilon,e}G = \frac{1}{2\epsilon}[\tau_{N\epsilon e}G - \tau_{-N\epsilon e}G]. \tag{4.13}$$

Remark. Although we use ∇ to denote different types of objects like ∇, ∇_e, $\nabla_{x,y}$ and $\nabla_{N\epsilon,e}$ it should be clear from the context what it denotes.

The next three lemmas are stated for the scalar case, with functions φ and h. They are typically one of the components of J, Φ or $\nabla\Phi$ and as a component of J, h will be zero on $\rho = 1$. The local function F will be linear in color and vanish if all sites are filled.

Lemma 4.4. *For any F, any coordinate direction e and for any $\epsilon_1 > 0$*

$$\begin{aligned}\limsup_{\epsilon\to 0}\limsup_{\epsilon_2\to 0}\limsup_{N\to\infty}\frac{1}{N^d}\log\hat{E}[\exp\{\int_0^T &\sum_x \varphi(t,\frac{x}{N})h(\bar{\zeta}_{x,N\epsilon_1}) \\ &[N\nabla_e(\tau_x F) - \nabla_{N\epsilon,e}\bar{F}_{x,N\epsilon_2}]dt\}] \leq 0\end{aligned}$$

Proof: We rewrite

$$\sum_x \varphi(t,\frac{x}{N})h(\bar{\zeta}_{x,N\epsilon_1})[N\nabla_e(\tau_x F)-\nabla_{N\epsilon,e}\bar{F}_{x,N\epsilon_2}] = \sum_x c_{N,\epsilon,\epsilon_1,\epsilon_2}(t,\frac{x}{N},\zeta)N\tau_x f_e$$
$$= \sum_x c_{N,\epsilon,\epsilon_1,\epsilon_2}(t,\frac{x}{N},\zeta)NL_{B_x}[\tau_x g]$$

where $f_e=\tau_e F - F = L_B\, g$ for some box B. Since $c_{N,\epsilon,\epsilon_1,\epsilon_2}(t,\frac{x}{N},\zeta)$ is unaltered by any rearrangement of particles inside B_x, it is easy to estimate the left hand side of (4.14) by

$$\limsup_{\epsilon\to 0}\limsup_{\epsilon_2\to 0}\limsup_{N\to\infty}\frac{1}{N^d}\sup_\zeta C\sum_x |c_{N,\epsilon,\epsilon_1,\epsilon_2}(t,\frac{x}{N},\zeta)|^2$$

which is easily computed to be 0.

Lemma 4.5. *Let $\hat{F}(\tilde{\rho})$ be the expected value of F under the product measure $\pi_{\tilde{\rho}}$. Then for any $\epsilon<\frac{1}{2}\epsilon_1$ and coordinate direction e,*

$$\limsup_{\epsilon_2\to 0}\limsup_{N\to\infty}\frac{1}{N^d}\log\hat{E}[\exp\{\int_0^T\sum_x|\nabla_{N\epsilon,e}\bar{F}_{x,N\epsilon_2}$$
$$-\frac{1}{2\epsilon}[\hat{F}(\bar{\zeta}_{x+N\epsilon e,N\epsilon_2})-\hat{F}(\bar{\zeta}_{x-N\epsilon e,N\epsilon_2})]|dt\}]\le 0$$

Proof: Since $\epsilon>0$ is fixed, it is sufficient to establish

$$\limsup_{\epsilon_2\to 0}\limsup_{N\to\infty}\frac{1}{N^d}\log\hat{E}[\exp\{\int_0^T\sum_x|\bar{F}_{x,N\epsilon_2}-\hat{F}(\bar{\zeta}_{x,N\epsilon_2})|dt\}]\le 0$$

which is an easy consequence of Theorems 3.1 and 3.2.

Lemma 4.6. *For any F,φ and h,*

$$\limsup_{\substack{\epsilon\to 0\\ k\to\infty}}\limsup_{N\to\infty}\frac{1}{N^d}\log\hat{E}[\exp\{\int_0^T\sum_x\varphi(t,\frac{x}{N})[h(\bar{\zeta}_{x,N\epsilon})-h(\bar{\zeta}_k)](N\tau_x\nabla_e F)dt\}]\le 0$$

Proof: We can apply Schwarz's inequality, and an application of Feynman-Kac formula and the variational formula yields

$$\limsup_{\substack{\epsilon\to 0\\ k\to\infty}}\limsup_{N\to\infty}\frac{1}{N^d}\log\hat{E}[\exp\{\int_0^T\sum_x\varphi(t,\frac{x}{N})[h(\bar{\zeta}_{x,N\epsilon})-h(\bar{\zeta}_k)](N\tau_x\nabla_e F)$$
$$-C\sum_x[h(\bar{\zeta}_{x,N\epsilon})-h(\bar{\zeta}_k)]^2dt\}]\le 0$$

We are therefore left with proving

$$\limsup_{\substack{\epsilon\to 0\\ k\to\infty}} \limsup_{N\to\infty} \frac{1}{N^d} \log \hat{E}[\exp\{C\int_0^T \sum_x [h(\bar{\zeta}_{x,N\epsilon}) - h(\bar{\zeta}_k)]^2 dt\}] \le 0$$

which follows from Theorem 3.2.

Although the Lemmas 4.4, 4.5 and 4.6 are only stated for $\hat{E}$ they are valid for $E_{N,b}$ as well.

Recall that for v in $\mathcal{V}$ we defined

$$\||v\||^2(\tilde{\rho}) = [v,v]_{\tilde{\rho}}. \tag{4.14}$$

If v depends on other parameters so will $\||v\||$.

Lemma 4.7. *Let $v(\tilde{\rho},\cdot)$ be in some fixed $\mathcal{V}_l$ for every $\tilde{\rho}$ in $\mathcal{C}$ and be continuous on $\mathcal{C}$. Then for sufficiently large k, $v(\bar{\zeta}_k,\zeta)$ is in $\mathcal{N}$. Consequently, so is $v_k(\tilde{\rho},\zeta) = v(\bar{\zeta}_k,\zeta) - v(\tilde{\rho},\zeta)$ and if we define $f_K(k,\tilde{\rho},\zeta)$ as $\langle(-L_K)^{-1}(\sum_{|x|\le K-k} v_k), (\sum_{|x|\le K-k} v_k\rangle$, there exist $\delta_k \to 0$ satisfying*

$$\limsup_{\substack{K\to\infty\\ \zeta\sim\tilde{\rho}}} \frac{1}{(2K+1)^d} f_K(\zeta) \le \delta_k.$$

Proof: We can write $v(\tilde{\rho},\cdot) = L_B w(\tilde{\rho},\cdot)$ with w depending continuously on $\tilde{\rho}$ as well. If $k > l$, where l is the linear size of the box B, then $v_k(\bar{\zeta}_{x,k},\tau_x\zeta) = L_{B+x}[w(\bar{\zeta}_{x,k},\tau_x\zeta) - v(\tilde{\rho},\tau_x\zeta)]$ and δ_k can be taken to be

$$\sup_{\tilde{\rho}} E^{\pi_{\tilde{\rho}}}\big\{\sup_{\zeta}[w(\bar{\zeta}_{0,k},\zeta) - w(\tilde{\rho},\zeta)]^2\big\}$$

which goes to 0.

We now calculate carefully the exponent in the exponential martingale (4.7), making the choice $g = \sum_{i,j}(\nabla\Phi)_{i,j} g_{i,j}$, and paying particular attention to various powers of N. We denote by Θ the exponent and write

$$\Theta = \Theta_1 + \Theta_2 + \Theta_3 + \Theta_4 - \Theta_5 + \Theta_6 \tag{4.15}$$

where

$$\begin{aligned}\Theta_1 =& \sum_x \Big\langle \Phi(T,\frac{x}{N}), \zeta(T,x)\Big\rangle - \sum_x \Big\langle \Phi(0,\frac{x}{N}), \zeta(0,x)\Big\rangle \\ &- \int_0^T \sum_x \Big\langle \Phi_t(t,\frac{x}{N}), \zeta(t,x)\Big\rangle\, dt\end{aligned} \tag{4.16}$$

$$\Theta_2 = N\int_0^T \sum_x \Big\langle \nabla\Phi(t,\frac{x}{N}); \tau_x \mathbf{W}(t) + \tau_x L\mathbf{g}(t,\frac{x}{N},\cdot)\Big\rangle\, dt \tag{4.17}$$

Here $\mathbf{W}$ has components

$$W_{i,j} = \frac{1}{2}\sum_y a_{0,y}(\zeta)p(y)(y,e_i)[\eta_j(y)-\eta_j(0)] \tag{4.18}$$

and L is applied to each component $g_{i,j}$ of $\mathbf{g}$ keeping t and x fixed.

$$\begin{aligned}\Theta_3 = \frac{1}{2}\int_0^T \sum_x \langle \nabla\Phi(t,\frac{x}{N});\tau_x\{\sum_z a_{0,z}(\zeta)p(z)b(t,\frac{x}{N},z,\zeta)\times \\ [z\otimes(\zeta(z)-\zeta(0))+\nabla_{0,z}\mathbf{G}(t,\frac{x}{N},\zeta)]\}\rangle dt\end{aligned} \tag{4.19}$$

where $\mathbf{G}(t,\theta,\zeta)=\sum_x \tau_x \mathbf{g}(t,\theta,\zeta)$.

$$\Theta_4 = \frac{1}{4}\int_0^T \sum_x\sum_z a_{x,x+z}(\zeta)\left\langle (z\otimes z)\nabla^2\Phi,\zeta(x+z)-\zeta(x)\right\rangle dt \tag{4.20}$$

Let us note that every term of (4.20) above has mean zero under every $\pi_{\tilde\rho}$.

$$\Theta_5 = \int_0^T \sum_x q(t,\frac{x}{N},\tau_x\zeta)dt \tag{4.21}$$

with

$$q(t,\theta,\zeta) = \frac{1}{4}\sum_y a_{0,y}(\zeta)p(y)\left\langle \nabla\Phi\,,[y\otimes(\zeta(y)-\zeta(0))+\nabla_{0,y}\mathbf{G}(t,\frac{x}{N},\cdot)]\right\rangle^2 . \tag{4.22}$$

One can verify that

$$E_{\pi_{\tilde\rho}}[q(t,\theta,\cdot)] = \|\!|\, \langle\nabla\Phi;\mathbf{W}+L\mathbf{G}\rangle \,|\!\|^2(t,\theta,\tilde\rho) \tag{4.23}$$

where $\|\!|\cdot|\!\|$ is as defined in (4.14). We combine all the remaining terms as,

$$\Theta_6 = o(N^d).$$

The last term is the error term. The sources are the exponential expansion, Taylor expansion for Φ and the slow variation in x when we pretend that $\mathbf{g}$ does not depend on θ in the definition of $\mathbf{G}$.

Let us now define a family of functions depending on $\delta>0$, $0<\lambda<\infty$, Φ , ϵ_1, ϵ,ϵ_2, J and l:

$$\mathbf{F} = \mathbf{F}_1+\mathbf{F}_2+\mathbf{F}_3+\mathbf{F}_4-(1+\delta)\mathbf{F}_5-2\lambda^2\mathbf{F}_6 \tag{4.24}$$

where

$$\begin{aligned}\mathbf{F}_1 = \Theta_1 = \sum_x \langle \Phi(T, \frac{x}{N}), \zeta(T,x)\rangle - \sum_x \langle \Phi(0, \frac{x}{N}, \zeta(0,x)\rangle \\ - \int_0^T \sum_x \langle \Phi_t(t, \frac{x}{N}), \zeta(t,x)\rangle dt\end{aligned} \tag{4.25}$$

$$\mathbf{F}_2 = \int_0^T \sum_x \langle \nabla\Phi; J(\bar{\zeta}_{x,N\epsilon_1})(\frac{1}{2\epsilon})[\hat{F}^l(\bar{\zeta}_{x+N\epsilon e,N\epsilon_2}) - \hat{F}^l(\bar{\zeta}_{x-N\epsilon e,N\epsilon_2})]\rangle dt \tag{4.26}$$

$$\mathbf{F}_3 = \frac{1}{2}\int_0^T \sum_x \langle \nabla\Phi(t, \frac{x}{N}); \mathbf{a}[\mathbf{b},\mathbf{g}](t, \frac{x}{N}, \zeta_{x,N\epsilon_1})\rangle dt. \tag{4.27}$$

Here $\mathbf{a}[\mathbf{b},\mathbf{g}]$ has components

$$\begin{aligned}&(\mathbf{a}[\mathbf{b},\mathbf{g}])_{i,j}(t,\theta,\tilde{\rho}) \\ &= \frac{1}{2}E^{\pi_{\tilde{\rho}}}[\sum_z a_{0,z}(\zeta)p(z)b(t,\theta,z,\zeta)\{\langle z, e_i\rangle (\eta_j(z) - \eta_j(0)) + \nabla_{0,z}G_{i,j}(t,\theta,\zeta)\}]\end{aligned} \tag{4.28}$$

where G and $\nabla_{0,z}G$ are defined in terms of g in (3.4), (3.5).

$$\mathbf{F}_4 = 0 \tag{4.29}$$

$$\mathbf{F}_5 = \int_0^T \sum_x \| \langle \nabla\Phi, \mathbf{W} + L\mathbf{g}\rangle \|^2(t, \frac{x}{N}, \zeta_{x,N\epsilon_1})dt \tag{4.30}$$

and

$$\mathbf{F}_6 = \int_0^T \sum_x \| \left\langle \nabla\Phi; \mathbf{W} + L\mathbf{g} - J(\tilde{\rho})\nabla_{\mathbf{e}}F^l\right\rangle \|^2(t, \frac{x}{N}, \zeta_{x,N\epsilon_1})dt \tag{4.31}$$

with $\nabla_{\mathbf{e}}F^l$ representing entries $\nabla_{e_i}F^l_j$. The notation here is that $\|.\|$ is a function of t,θ and $\tilde{\rho}$, and $\frac{x}{N}$ and $\zeta_{x,N\epsilon_1}$ are substituted for the last two.

Let us denote by

$$\mathbf{F}_1 - \Theta_1 = \mathbf{K}_1 = 0 \tag{4.32}$$

$$\mathbf{F}_2 - \Theta_2 - 2\lambda^2\mathbf{F}_6 = \mathbf{K}_2 \tag{4.33}$$

$$\mathbf{F}_3 - \Theta_3 = \mathbf{K}_3 \tag{4.34}$$

$$\mathbf{F}_4 - \Theta_4 = -\Theta_4 = \mathbf{K}_4 \tag{4.35}$$

$$(1+\delta)(\Theta_5 - \mathbf{F}_5) = \mathbf{K}_5 \tag{4.36}$$

We write

$$\mathbf{F} = \Theta_1 + \Theta_2 + \Theta_3 + \Theta_4 - (1+\delta)\Theta_5 + \mathbf{K}_1 + \mathbf{K}_2 + \mathbf{K}_3 + \mathbf{K}_4 + \mathbf{K}_5 \tag{4.37}$$

Theorem 4.8. *For any $\delta > 0$ there exist a $\lambda = \lambda(\delta) > 0$, depending only on δ, such that $\lambda(\delta) \to \infty$ as $\delta \to 0$ and*

$$\limsup_{N\to\infty} \frac{1}{N^d} \log E_{N,b}[\exp\{\mathbf{F}\}] \le C(\lambda, \epsilon, \epsilon_1, \epsilon_2) + \sigma(\lambda) \tag{4.38}$$

where

$$\lim_{\lambda\to\infty} \sigma(\lambda) = 0 \tag{4.39}$$

and for every positive λ,

$$\limsup_{\epsilon_1\to 0} \limsup_{\epsilon\to 0} \limsup_{\epsilon_2\to 0} C(\lambda, \epsilon, \epsilon_1, \epsilon_2) \le 0 \tag{4.40}$$

Proof. We estimate $E_{N,b}[\exp\{\mathbf{F}\}]$ by Hölder's inequality with a factor of $(1+\delta)$ multiplying the five Θ terms and a corresponding λ multiplying each one of the $\mathbf{K}$ terms.

The first of these is basically the exponential martingale with two provisos. We have $(1+\delta)\Phi$ instead of Φ. The term Θ_6 which is uniformly negligible is missing. Therefore

$$\limsup_{N\to\infty} \frac{1}{N^d} \log E_{N,b}[\exp\{(1+\delta)(\Theta_1 + \Theta_2 + \Theta_3 + \Theta_4) - (1+\delta)^2\Theta_5\}] \le 0. \tag{4.41}$$

It remains to show only that

$$\limsup_{N\to\infty} \frac{1}{N^d} \log E_{N,b}[\exp\{\lambda \mathbf{K}_i\}] \le C(\lambda, \epsilon_1, \epsilon, \epsilon_2) + \sigma(\lambda) \tag{4.42}$$

for $i = 1,2,3,4$ and 5, with C and σ satisfying (4.40) and (4.39). We go from $E_{N,b}$ to $\hat{E}$ by Holder's inequality using $p = \frac{\lambda}{\lambda-1}$ for the Radon-Nikodym derivative and $q = \lambda$ for the various exponentials. We therefore need only prove that

$$\limsup_{N\to\infty} \frac{1}{N^d} \log \hat{E}\ [\exp\{\lambda^2 \mathbf{K}_i\}] \le C(\lambda, \epsilon_1, \epsilon, \epsilon_2) \tag{4.43}$$

for some C satisfying (4.40). Since $\mathbf{K}_1 = 0$ this is trivial for $i = 1$. For $i = 3, 4$ and 5 this is an easy consequence of the one and two block estimates Theorems 3.1 and 3.2. We now concentrate on $\mathbf{K}_2$.

$$\begin{aligned}\mathbf{K}_2 = &\int_0^T \sum_x \langle \nabla\Phi; J(\bar{\zeta}_{x,N\epsilon_1})(\frac{1}{2\epsilon})[\hat{F}^l(\bar{\zeta}_{x+N\epsilon,N\epsilon_2}) - \hat{F}^l(\bar{\zeta}_{x-N\epsilon,N\epsilon_2})]\rangle dt \\ &- \int_0^T \sum_x \langle \nabla\Phi; [\tau_x \mathbf{W} + (L\mathbf{g})(t, \frac{x}{N}, \tau_x\zeta)]\rangle dt \\ &- 2\lambda^2 \int_0^T \sum_x |||\langle \nabla\Phi; \mathbf{W} + L\mathbf{g} - J(\tilde{\rho})\nabla_{\mathbf{e}} F^l\rangle|||^2(t, \frac{x}{N}, \zeta_{x,N\epsilon_1}) dt\end{aligned} \tag{4.44}$$

is written as the sum of four terms $\mathbf{K}_{21} + \mathbf{K}_{22} + \mathbf{K}_{23} + \mathbf{K}_{24}$ given by

$$\begin{aligned}\mathbf{K}_{21} = &\int_0^T \sum_x \langle \nabla\Phi; J(\bar{\zeta}_{x,N\epsilon_1})(\frac{1}{2\epsilon})[\hat{F}^l(\bar{\zeta}_{x+N\epsilon,N\epsilon_2}) - \hat{F}^l(\bar{\zeta}_{x-N\epsilon,N\epsilon_2})]\rangle dt \\ &- \int_0^T \sum_x \langle \nabla\Phi; J(\bar{\zeta}_{x,N\epsilon_1}) \nabla_{N\epsilon,\mathbf{e}}[\bar{F}^l_{x,N\epsilon_2}]\rangle dt\end{aligned} \tag{4.45}$$

$$\begin{aligned}\mathbf{K}_{22} = &\int_0^T \sum_x \langle \nabla\Phi; J(\bar{\zeta}_{x,N\epsilon_1}) \nabla_{N\epsilon,\mathbf{e}}[\bar{F}^l_{x,N\epsilon_2}]\rangle dt \\ &- \int_0^T \sum_x \langle \nabla\Phi; J(\bar{\zeta}_{x,N\epsilon_1})(N\nabla_e)(\tau_x F^l)\rangle dt\end{aligned} \tag{4.46}$$

$$\begin{aligned}\mathbf{K}_{23} = &\int_0^T \sum_x \langle \nabla\Phi; J(\bar{\zeta}_{x,N\epsilon_1})(N\nabla_e)(\tau_x F^l)\rangle dt \\ &- \int_0^T \sum_x \langle \nabla\Phi; J(\bar{\zeta}_{x,k})(N\nabla_e)(\tau_x F^l)\rangle dt\end{aligned} \tag{4.47}$$

$$\begin{aligned}\mathbf{K}_{24} = &\int_0^T \sum_x \langle \nabla\Phi; [J(\bar{\zeta}_{x,k})(N\nabla_e)(\tau_x F^l) - \tau_x \mathbf{W} - (L\mathbf{g})(t, \frac{x}{N}, \tau_x\zeta)\rangle dt \\ &- 2\lambda^2 \int_0^T \sum_x |||\langle \nabla\Phi; \mathbf{W} + L\mathbf{g} - J(\tilde{\rho})\nabla_{\mathbf{e}} F^l\rangle|||^2(t, \frac{x}{N}, \zeta_{x,N\epsilon_1}) dt\end{aligned} \tag{4.48}$$

For $\mathbf{K}_{21}$, (4.43) is a consequence of Lemma 4.7. For $\mathbf{K}_{22}$ it follows from lemma 4.3 and Lemma 4.7 takes care of $\mathbf{K}_{23}$. Finally it follows from theorems 3.1,3.2 and 3.3 that

$$\limsup_{k\to\infty,\epsilon_1\to 0} \limsup_{N\to\infty} \frac{1}{N^d} \log \hat{E}[\exp\{2\lambda^2 \mathbf{K}_{24}\}] \le 0 \tag{4.49}$$

We have to apply Holder's inequlaity pull the exponents apart. We have allowed for an extra factor of 2 in $\mathbf{K}_{24}$ and the other terms that are linear in F can absorb any constant. We have now established a lower bound for the n-color "rate function" $\hat{I}$.

Theorem 4.9. *The rate function $\hat{I}(\tilde{\rho}(\cdot,\cdot))$ has the following lower bound.*

$$
\begin{aligned}
\hat{I}(\tilde{\rho}(\cdot,\cdot)) \geq & \int_{\mathbf{T}^d} \langle \Phi(T,\theta), \tilde{\rho}(T,\theta)) \rangle \, d\theta - \int_{\mathbf{T}^d} \langle \Phi(0,\theta), \tilde{\rho}(0,\theta)) \rangle \, d\theta \\
& - \int_0^T \int_{\mathbf{T}^d} \langle \Phi_t(t,\theta); \tilde{\rho}(t,\theta)) \rangle \, dtd\theta \\
& + \int_0^T \int_{\mathbf{T}^d} \left\langle \nabla \Phi(t,\theta); (J \nabla \hat{F}^l)(\tilde{\rho}(t,\theta)) \right\rangle dtd\theta \\
& + \int_0^T \int_{\mathbf{T}^d} \langle \nabla \Phi(t,\theta); \mathbf{a}[\mathbf{b},\mathbf{g}](t,\theta,\tilde{\rho}(t,\theta)) \rangle dtd\theta \\
& - (1+\delta) \int_0^T \int_{\mathbf{T}^d} \||\langle \nabla \Phi; \mathbf{W} + L\mathbf{g} \rangle \||^2 (t,\theta,\tilde{\rho}(t,\theta)) dtd\theta \\
& - \lambda^2 \int_0^T \int_{\mathbf{T}^d} \|| \left\langle \nabla \Phi; \mathbf{W} + L\mathbf{g} - J(\tilde{\rho}) \nabla_e F^l \right\rangle \||^2 (t,\theta,\tilde{\rho}(t,\theta)) dtd\theta \\
& - \sigma(\lambda)
\end{aligned}
\tag{4.50}
$$

¿From this point we pick $J(\tilde{\rho})$ to be of the form $h(\rho)\ \frac{1}{2}A(\tilde{\rho})\ \chi(\tilde{\rho})$ where h is a cutoff function vanishing near $\rho = 1$. First we remark that because of Lemma 4.7 we can assume that $\mathbf{g}$ can depend on $\tilde{\rho}$ as well. Calculations in "gradient space " allows us to optimize with respect to $\mathbf{g}$ which essentially projects $\mathbf{W}$ into $\frac{1}{2}A\chi$ times the image of the density gradient $\mathbf{f}$. This projection is carried out explicitly in the Hilbert space $H_{\tilde{\rho}}$. If we denote by $\tilde{\mathbf{\Psi}}$ and $\tilde{\mathbf{f}}$ the images in $H_{\tilde{\rho}}$ of $\mathbf{W}$ and $\mathbf{f}$ respectively then $\tilde{\mathbf{\Psi}} = \frac{1}{2}A(\tilde{\rho})\chi(\tilde{\rho})\tilde{\mathbf{f}}$ modulo $H^0_{\tilde{\rho}}$ and while the minimizing g may not exist there will be a unique minimizing ξ in $H^0_{\tilde{\rho}}$. So the functions $\mathbf{a}[\mathbf{b},\mathbf{g}](t,\theta,\tilde{\rho})$ will have a limit along any minimizing sequence of g's and the limit can be calculated as

$$
\hat{\mathbf{a}}[\mathbf{b}]_{i,j}(t,\theta,\tilde{\rho}) = \frac{1}{2} E^{\pi_{\tilde{\rho}}} \{ \sum_z a_{0,z}(\zeta) p(z) b(t,\theta,z,\zeta) [\langle z, e_i \rangle (\eta_j(z) - \eta_j(0)) + \xi_z] \}
\tag{4.51}
$$

which, by Remark 3 at the end of Section 3, is a continuous function of t, θ and $\tilde{\rho}$ for our choice of b.

A simple calculation provides a new lower bound.

$$
\begin{aligned}
\hat{I}(\tilde{\rho}(\cdot,\cdot)) \geq & \int_{\mathbf{T}^d} \langle \Phi(T,\theta), \tilde{\rho}(T,\theta)) \rangle \, d\theta - \int_{\mathbf{T}^d} \langle \Phi(0,\theta), \tilde{\rho}(0,\theta) \rangle \, d\theta \\
& - \int_0^T \int_{\mathbf{T}^d} \langle \Phi_t(t,\theta), \tilde{\rho}(t,\theta) \rangle \, dt d\theta \\
& + \int_0^T \int_{\mathbf{T}^d} \left\langle \nabla\Phi(t,\theta); (\frac{1}{2} hA\chi(\nabla \mathbf{F}^l)(\tilde{\rho}(t,\theta)) \right\rangle dt d\theta \\
& + \int_0^T \int_{\mathbf{T}^d} \langle \nabla\Phi(t,\theta); \hat{\mathbf{a}}[\mathbf{b}](t,\theta,\tilde{\rho}(t,\theta)) \rangle \, dt d\theta \\
& - \frac{(1+\delta)}{2} \int_0^T \int_{\mathbf{T}^d} \langle \nabla\Phi(t,\theta); A(\tilde{\rho}(t,\theta)) \nabla\Phi(t,\theta) \rangle \, dt d\theta \\
& - \frac{\lambda^2}{4} \int_0^T \int_{\mathbf{T}^d} || \left\langle \nabla\Phi; [A\chi\mathbf{f} - hA\chi\mathbf{f}^l] \right\rangle ||^2_{H_{\tilde{\rho}}}(t,\theta,\tilde{\rho}(t,\theta)) dt d\theta \\
& - \sigma(\lambda)
\end{aligned} \tag{4.52}
$$

We next show that we can let $l \to \infty$ and $\mathbf{f}^l$ converges nicely to $\mathbf{f}$ away from $\rho = 1$. Because of the cutoff h this limit can be taken. Then the cutoff is removed allowing h to tend to 1. The estimates (4.6) and (4.7) provided in Theorem 4.3 allow us to show that the term that is multiplied by λ goes to zero for any λ. Finally we let $\delta \to 0$ and $\lambda \to \infty$ and obtain

$$
\begin{aligned}
\hat{I}(\tilde{\rho}(\cdot,\cdot)) \geq & \int_{\mathbf{T}^d} \langle \Phi(T,\theta), \tilde{\rho}(T,\theta)) \rangle \, d\theta - \int_{\mathbf{T}^d} \langle \Phi(0,\theta), \tilde{\rho}(0,\theta)) \rangle \, d\theta \\
& - \int_0^T \int_{\mathbf{T}^d} \langle \Phi_t(t,\theta), \tilde{\rho}(t,\theta)) \rangle \, dt d\theta \\
& + \int_0^T \int_{\mathbf{T}^d} \left\langle \nabla\Phi(t,\theta); (\frac{1}{2} A\chi)(\tilde{\rho}(t,\theta)) \nabla\tilde{\rho}(t,\theta)) \right\rangle dt d\theta \\
& + \int_0^T \int_{\mathbf{T}^d} \langle \nabla\Phi(t,\theta); \hat{\mathbf{a}}[\mathbf{b}](t,\theta,\tilde{\rho}(t,\theta)) \rangle \, dt d\theta \\
& - \frac{1}{2} \int_0^T \int_{\mathbf{T}^d} \langle \nabla\Phi(t,\theta); A(\tilde{\rho}(t,\theta)) \nabla\Phi(t,\theta) \rangle \, dt d\theta.
\end{aligned} \tag{4.53}
$$

Taking the supremum over Φ we obtain

$$
\hat{I}(\tilde{\rho}(\cdot,\cdot)) \geq I^{n,b}_{dyn}(\tilde{\rho}(\cdot,\cdot)) = \frac{1}{2} \int_0^T || \frac{\partial\tilde{\rho}}{\partial t} - \frac{1}{2} \nabla \cdot A\chi(\tilde{\rho}) \nabla\tilde{\rho} - \nabla \cdot \hat{\mathbf{a}}[\mathbf{b}] ||^2_{-1,A(\tilde{\rho})} dt \tag{4.54}
$$

Theorem 4.10. *The measures $P_{N,b}$ have a large deviation upper bound with the rate function $I^{n,b}_{dyn}(\tilde{\rho}(\cdot,\cdot))$. More precisely on the space E_n the family $P_{N,b}$ satisfies the large deviation upper bound so that for every closed set C*

$$
\limsup_{N\to\infty} \frac{1}{N^d} \log P_{N,b}[C] \leq - \inf_{\tilde{\rho}(\cdot,\cdot)\in C_0} I^{n,b}_{dyn}(\tilde{\rho}(\cdot,\cdot)) \tag{4.55}
$$

where $C_0 = C \cap \{\tilde\rho(\cdot,\cdot) : \tilde\rho(0,\cdot) = \tilde\rho_0(\cdot)\}$ are those profiles in C that are compatible with the initial profile at time 0.

In particular we can take $b \equiv 0$ and we obtain the upper bounds for P_N. A further consequence is the law of large numbers for the sequence $P_{N,b}$:

Theorem 4.11. *For any b the family $P_{N,b}$ is tight on the space E_n and any weak limit of $P_{N,b}$ will be concentrated in the subset of D_n^0 that are weak solutions of*

$$\frac{\partial\tilde\rho}{\partial t} - \frac{1}{2}\nabla\cdot A\chi(\tilde\rho)\nabla\tilde\rho = \nabla\cdot\hat{\mathbf{a}}[\mathbf{b}] \tag{4.56}$$

with initial condition $\tilde\rho_0$.

We will now attempt to prove the large deviation lower bounds for P_N using as our tool Theorem 4.9. Suppose $\mathbf{H}(t,\theta)$ is a continuous map from $[0,T]\times\mathbf{T}^d$ into $\mathbf{R}^n\otimes\mathbf{R}^d$ and suppose we are given a profile $\tilde\rho(t,\theta)$ that satisfies

$$\frac{\partial\tilde\rho}{\partial t} - \frac{1}{2}\nabla\cdot A\chi(\tilde\rho)\nabla\tilde\rho = \nabla\cdot A(\tilde\rho)\mathbf{H} \tag{4.57}$$

in the weak sense. Let us suppose moreover that the above equation has a unique solution $\tilde\rho$ in this class $\mathcal{D}_n^0$ for the given $\mathbf{H}$ and initial profile $\tilde\rho_0$. We look for functions $b(t,\theta,z,\zeta)$ such that $\hat{\mathbf{a}}[\mathbf{b}] = A(\tilde\rho(t,\theta))\mathbf{H}$. We may denote this class by $\mathcal{B}_{\tilde\rho,\mathbf{H}}$. For each b in this class the perturbed process $P_{N,b}$ will have a certain relative entropy with respect to P_N and we call this $Ent(N,b)$. We now show

Theorem 4.12.

$$\inf_{b\in\mathcal{B}_{\tilde\rho,\mathbf{H}}}\limsup_{N\to\infty}\frac{1}{N^d}Ent(N,b) \le \frac{1}{2}\int_0^T\int_{\mathbf{T}^d}\langle\mathbf{H}(t,\theta), A(\tilde\rho(t,\theta))\mathbf{H}(t,\theta)\rangle\, dt d\theta \tag{4.58}$$

Proof. If we make the choice

$$b(t,\theta,z,\zeta) = \langle\mathbf{H}, z\otimes(\zeta(z)-\zeta(0))\rangle \tag{4.59}$$

then a straightforward calculation yields,

$$\begin{aligned}\hat{\mathbf{a}}[\mathbf{b}](t,\theta,\tilde\rho) &= \frac{1}{2}E^{\pi_{\tilde\rho}}[\sum_z p(z)a_{0,z}(\zeta)\langle\mathbf{H}; z\otimes(\zeta(z)-\zeta(0))\rangle(z\otimes(\zeta(z)-\zeta(0))+\xi_z)] \\ &= \frac{1}{2}E^{\pi_{\tilde\rho}}[\sum_z p(z)a_{0,z}(\zeta)\langle\mathbf{H};\psi_z\rangle(\psi_z+\xi_z)]\end{aligned} \tag{4.60}$$

where $\{\xi_z\}$ is the projection of $\{\psi_z\} = z\otimes(\zeta(z)-\zeta(0))$ onto $H^0_{\tilde\rho}$. From (3.14) this is seen to equal $A(\tilde\rho)\mathbf{H}$. Actually we could make the choice

$$b(t,\theta,z,\zeta) = \langle\mathbf{H}; z\otimes(\zeta(z)-\zeta(0))\rangle + \nabla_{0,z}G \tag{4.61}$$

for some $G = \sum_x g(\tau_x \zeta)$ and due to the orthogonality of $\{\psi_z + \xi_z\}$ and $\nabla_{0,z}G$, $\hat{\mathbf{a}}[\mathbf{b}]$ will be unaffected by this modification. On the other hand, using the Girsanov formula the relative entropy $Ent(N, b)$ can be calculated explicitly and the answer depends on the modification $g = g(t, \theta, \zeta)$ used. An easy computation yields

(4.62)
$$\limsup_{N\to\infty} \frac{1}{N^d} Ent(N, b) \le \int_0^T \int_{\mathbf{T}^d} \frac{1}{4} E^{\pi_{\tilde\rho(t,\theta)}} \left[\sum_z a_{0,z}(\zeta) p(z) \left(\langle \mathbf{H}; \psi_z \rangle - \nabla_{0,z} G \right)^2 \right] dt d\theta.$$

By the variational formula (3.14) for A, minimization over g gives

$$\frac{1}{2} \int_0^T \int_{\mathbf{T}^d} \langle \mathbf{H}(t,\theta); A(\tilde\rho(t,\theta)) \mathbf{H}(t,\theta) \rangle \, dt d\theta \tag{4.63}$$

which is the required result.

Let us describe a class $\mathcal{S}$ of profiles. For a given initial profile $\tilde\rho_0(\cdot)$, these will satisfy the following: $\tilde\rho(0,\theta) = \tilde\rho_0(\theta)$ on $\mathbf{T}^d$. $\tilde\rho(\cdot,\cdot)$ is an infinitely differentiable function of t and θ in the open interval $(0,T] \times \mathbf{T}^d$ and avoids all the boundaries of $\mathcal{C}$ i.e. ρ_j and $(1-\rho)$ are all bounded away from zero so long as t remains positive. Finally for some $t_0 > 0$ that is allowed to depend on the choice of $\tilde\rho(\cdot,\cdot)$ the profile $\tilde\rho(\cdot,\cdot)$ satisfies weakly the equation

$$\frac{\partial \tilde\rho}{\partial t} = \frac{1}{2} \nabla \cdot A \chi \nabla \tilde\rho \tag{4.64}$$

in the interval $[0, t_0]$.

In section 5 we will prove that each $\tilde\rho(\cdot,\cdot)$ in $\mathcal{S}$ is the unique weak solution in $\mathcal{D}_n^0$ of an equation

$$\frac{\partial \tilde\rho}{\partial t} = \frac{1}{2} \nabla \cdot A \chi \nabla \tilde\rho + \nabla \cdot A(\tilde\rho(t,\cdot)) \nabla \mathcal{H} \tag{4.65}$$

for some continuously differentiable $\mathcal{H}$, that of course vanishes for sufficiently small t, i.e $t < t_0$.

In section 6, we will prove that $\mathcal{S}$ is dense in $\mathcal{D}_n^0$ in the proper sense. More precisely we will prove that for any $\tilde\rho(\cdot,\cdot)$ in $\mathcal{D}_n^0$ such that $I_{dyn}^n(\tilde\rho(\cdot,\cdot)) < \infty$ there exists a sequence $\tilde\rho_k(\cdot,\cdot)$ in $\mathcal{S}$ such that, they all have the same initial profile $\tilde\rho_0(\cdot)$,

$$\lim_{k\to\infty} \tilde\rho_k(\cdot,\cdot) = \tilde\rho_k(\cdot,\cdot) \quad \text{in} \quad C[[0,T], \mathcal{M}]$$

and

$$\liminf_{k\to\infty} I_{dyn}^n(\tilde\rho_k(\cdot,\cdot)) \le I_{dyn}^n(\tilde\rho(\cdot,\cdot)).$$

Once these uniqueness and approximation results are established the lower bounds for the large deviation principle follows.

Theorem 4.13. *For any deterministic initial conditions compatible with any given profile $\tilde\rho_0(\cdot)$ the large deviation principle holds for P_N with rate function*

$$I(\tilde\rho(\cdot,\cdot)) = \frac{1}{2} \int_0^T || \frac{\partial \tilde\rho}{\partial t} - \frac{1}{2} \nabla \cdot A \chi(\tilde\rho(t,\cdot)) \nabla \tilde\rho(t,\cdot) ||_{-1,A(\tilde\rho(t,\cdot))}^2 dt \tag{4.66}$$

on those $\tilde\rho(\cdot,\cdot)$ that are compatible with the initial profile $\tilde\rho_0(\cdot)$ at time 0.

5. Uniqueness for driven systems

The aim of this section is to prove a theorem regarding uniqueness of weak solutions for a system of equations of the form

$$\frac{\partial \tilde{\rho}}{\partial t} = \frac{1}{2} \nabla \cdot A\chi(\tilde{\rho}(t,\cdot))\nabla\tilde{\rho}(t,\cdot) + \nabla \cdot A(\tilde{\rho}(t,\cdot))\mathbf{H} \tag{5.1}$$

with initial conditions

$$\tilde{\rho}(0,\cdot) = \tilde{\rho}_0(\cdot), \tag{5.2}$$

for a matrix $\mathbf{H}_{i,j}(t,x)$ each of whose components are uniformly bounded. Since we will be dealing exclusively with functions defined on $[0,T] \times \mathbf{T}^d$, in this section as well as the next one we will find it more convenient to denote by x and y points on the torus $\mathbf{T}^d$. Our uniqueness theorem will be for $\tilde{\rho}(\cdot,\cdot)$ in the class $\mathcal{D}_n^0$ of functions $\tilde{\rho}(\cdot,\cdot)$ mapping $[0,T] \times \mathbf{T}^d$ into $\mathcal{C}$, that satisfy the inital condition (5.2), the regularity conditions

$$\begin{aligned} &\int_0^T \int_{\mathbf{T}^d} \frac{|\nabla \rho|^2}{(1-\rho)} dx dt < \infty \\ &\int_0^T \int_{\mathbf{T}^d} (1-\rho) \frac{|\nabla \rho_j|^2}{\rho_j} dx dt < \infty \end{aligned} \tag{5.3}$$

or, equivalently,

$$\int_0^T \int_{\mathbf{T}^d} \langle \nabla\tilde{\rho}; \chi A \chi \nabla \tilde{\rho} \rangle \, dx dt < \infty \tag{5.4}$$

and the equation (5.1) in weak form: For all smooth $\Phi : [0,T] \times \mathbf{T}^d \to \mathbf{R}^n$,

$$\begin{aligned} &\int_{\mathbf{T}^d} \langle \Phi(T,\cdot), \tilde{\rho}(T,\cdot)\rangle dx - \int_{\mathbf{T}^d} \langle \Phi(0,\cdot), \tilde{\rho}(0,\cdot)\rangle dx - \int_0^T \int_{\mathbf{T}^d} \langle \Phi_t(t,\cdot), \tilde{\rho}(t,\cdot)\rangle dx dt \\ &\quad + \frac{1}{2} \int_0^T \int_{\mathbf{T}^d} \langle \nabla\Phi(t,\cdot); A\chi(\tilde{\rho}(t,\cdot))\nabla\tilde{\rho}(t,\cdot)\rangle dx dt \\ &\quad + \int_0^T \int_{\mathbf{T}^d} \langle \nabla\Phi(t,\cdot); A(\tilde{\rho}(t,\cdot))\mathbf{H}(t,\cdot)\rangle dx dt = 0 \end{aligned} \tag{5.5}$$

Our uniqueness theorem is a little bit complicated to state. We will have a much smaller class $\mathcal{D}_n^1 \subset \mathcal{D}_n^0$ of solutions $\tilde{R}$ of (5.1) in the interval $[0,T]$ that satisfy in addition

$$\sup_{x,t} |\nabla \tilde{R}(t,x)| < \infty \tag{5.6}$$

and for some $\delta > 0$,

$$\begin{aligned} &\sup_{j,x,t} R_j(t,x) \le (1-\delta) \\ &\inf_{j,x,t} R_j(t,x) \ge \delta \end{aligned} \tag{5.7}$$

Before we state and prove our uniqueness theorem we will establish a preliminary estimate on $\rho = \sum_{i=1}^n \rho_i$ for any solution $\tilde\rho = (\rho_1, \dots, \rho_n)$ of (5.1). Let $\tilde\rho$ be a weak solution of (5.1) and $u = 1 - \rho$. Then u is a weak solution of an equation of the form

$$\frac{\partial u}{\partial t} = \frac{1}{2}\nabla \cdot D\nabla + \nabla \cdot bu \tag{5.8}$$

where $b_i = -\sum_{j=1}^n \rho_j \sum_{i'=1}^d D_{i,i'} H_{i',j}$ is uniformly bounded in absolute value.

Lemma 5.1. *Let $u(t,x)$ be the density at time t of a diffusion process on the torus $\mathbf{T}^d$ with the backward generator*

$$\frac{1}{2}\nabla \cdot D\nabla + b \cdot \nabla \tag{5.9}$$

where $\|b\|_\infty < \infty$ and with initial density $u(0,x)$ bounded uniformly away from 0. Then for each $T > 0$ there exists $\delta > 0$ depending only on $(T, \|b\|_\infty)$ such that for all $t \in [0,T]$, $u(t,x) \geq \delta$.

Proof. It is more convenient, for computational reasons, to extend b periodically in space and look at the same problem on $\mathbf{R}^d$. The lower bound on $\mathbf{T}^d$ can be easily deduced from any reasonable lower bound on $\mathbf{R}^d$.

Let $q(0,x,t,y)$ be the transition density on $\mathbf{R}^d$ of the process with drift $b \equiv 0$. By the Cameron-Martin-Girsanov formula we have an explicit representation for p in terms of q and the Brownian bridge $Q_t^{x,y}$ (with covariance D) on paths $z(s)$ starting at x at time 0 and conditioned to end at y at time t,

$$p(0,x,t,y) = q(0,x,t,y) E^{Q_t^{x,y}}\left[\exp\left(\int_0^t \langle D^{-1}b, dz(s)\rangle - \frac{1}{2}\int_0^t \langle D^{-1}b\ , b\rangle ds\right)\right] \tag{5.10}$$

The second term in the exponent in uniformly bounded. By Jensen's inequality and a straightforward stochastic calculus computation,

$$\begin{aligned} E^{Q_t^{x,y}}\left[\exp\left(\int_0^t \langle D^{-1}b\ , dz(s)\rangle\right)\right] &\geq \exp\left(E^{Q_t^{x,y}}\left[\int_0^t \langle D^{-1}b\ , dz(s)\rangle\right]\right) \\ &= \exp\left(E^{Q_t^{x,y}}\left[\int_0^t \frac{\langle D^{-1}b\ , y - z(s)\rangle}{(t-s)} ds\right]\right) \end{aligned} \tag{5.11}$$

By Schwarz's inequality,

$$E^{Q_t^{x,y}}|y - z(s)| \leq \left[E^{Q_t^{x,y}}|y - z(s)|^2\right]^{1/2} = \left[|x-y|^2\left(\frac{t-s}{t}\right)^2 + \frac{s(t-s)}{t}\right]^{1/2} \tag{5.12}$$

and

$$\begin{aligned} &|E^{Q_t^{x,y}}\int_0^t \frac{\langle D^{-1}b\ , y - z(s)\rangle}{t-s} ds| \\ &\leq C\|b\|_\infty \int_0^t \frac{1}{(t-s)}\left[|x-y|\left(\frac{t-s}{t}\right) + \left(\frac{s(t-s)}{t}\right)^{1/2}\right] ds \\ &\leq \|b\|_\infty\left[C_1|x-y| + C_2\right] \end{aligned} \tag{5.13}$$

for some constants that could possibly depend on T. This proves the lemma.

We are now ready to state and prove the main uniqueness theorem.

Theorem 5.2. *Suppose that the function* $\mathbf{H}$ *is bounded on* $[0,T] \times \mathbf{T}^d$ *and equation (5.1) admits a solution* $\tilde{R}$ *that is in* $\mathcal{D}_n^1$*. Then any weak solution of (5.1) from the class* $\mathcal{D}_n^0$ *is identical to* $\tilde{R}$*.*

Proof. Suppose we have two weak solutions: a solution $\tilde{R}$, in $\mathcal{D}_n^1$ and another $\tilde{\rho}$ in $\mathcal{D}_n^0$. We use the notation $A = A(\tilde{R})$, $a = A(\tilde{\rho})$, $X = \chi(\tilde{R})$, $\chi = \chi(\tilde{\rho})$; $\Lambda = \lambda(\tilde{R})$ is the chemical potential corresponding to $\tilde{R}$ given by $\Lambda_j = \log \frac{R_j}{(1-R)}$ for $1 \leq j \leq n$. Let

$$F(\lambda) = \log\left(1 + \sum_{i=1}^{n} e^{\lambda_i}\right) \tag{5.14}$$

be the free energy, and

$$h(\tilde{\rho}) = \sum_{i=1}^{n} \rho_i \log \rho_i + (1-\rho)\log(1-\rho) \tag{5.15}$$

be the corresponding entropy functional. Note that the relative entropy is given by

$$h(\tilde{\rho}; \tilde{R}) = h(\tilde{\rho}) - \tilde{\rho} \cdot \Lambda + F(\Lambda) \tag{5.16}$$

Because $\tilde{R}$ is a solution in $\mathcal{D}_n^1$, $1 - R$ and R_j are bounded away from 0 and since $\tilde{\rho}$ has the same initial value as $\tilde{R}$, $1 - \rho_0$ is bounded away from 0 as well. By lemma 5.1 and (5.8), for our weak solution $\tilde{\rho}$ from the class $\mathcal{D}_n^0$, $1 - \rho$ is bounded away from zero uniformly in the interval $[0, T]$. A computation gives

$$\begin{aligned}
&\int_{\mathbf{T}^d} \left[h(\tilde{\rho}) - \tilde{\rho} \cdot \Lambda + F(\Lambda)\right] dx \Big|_0^T \\
&= -\int_0^T \int_{\mathbf{T}^d} \Big[\Big\langle \left(\chi \nabla \tilde{\rho} - X \nabla \tilde{R}\right); \left(\tfrac{1}{2} a \chi \nabla \tilde{\rho} + a\mathbf{H}\right) \Big\rangle \\
&\quad + \Big\langle \nabla\left(\tilde{R} - \tilde{\rho}\right); X\left(\tfrac{1}{2} A X \nabla \tilde{R} + A\mathbf{H}\right) \Big\rangle \\
&\quad + \Big\langle \nabla \tilde{R}\; X' \cdot \left(\tilde{R} - \tilde{\rho}\right); \left(\tfrac{1}{2} A X \nabla \tilde{R} + A\mathbf{H}\right) \Big\rangle \Big] \, dx dt
\end{aligned} \tag{5.17}$$

Completing the square and defining E by

$$\chi^{-1} = X^{-1} + X^{-1}(X' \cdot (\tilde{R} - \tilde{\rho}))X^{-1} + E \tag{5.18}$$

this becomes

(5.19)

$$\int_0^T \int_{\mathbf{T}^d} \Bigg[-\frac{1}{2} \left| (\chi a \chi)^{1/2} \nabla \tilde{\rho} - \tfrac{1}{2} (\chi a \chi)^{-1/2} (\chi a X + XAX) \nabla \tilde{R} \right|^2$$

$$- \tfrac{1}{8} \left\langle \nabla \tilde{R}; (XAX - Xa\chi) (\chi a \chi)^{-1} (XAX - \chi a X) \nabla \tilde{R} \right\rangle$$

$$- \left\langle (\chi a \chi)^{1/2} \nabla \tilde{\rho} - \frac{1}{2} (\chi a \chi)^{-1/2} (\chi a X + XAX) \nabla \tilde{R}; (\chi a \chi)^{-1/2} (\chi a - XA) \mathbf{H} \right\rangle$$

$$- \left\langle \nabla \tilde{R}; \frac{1}{2} (XAX - Xa\chi) (\chi a \chi)^{-1} (\chi a - XA) \mathbf{H} \right\rangle$$

$$+ \frac{1}{2} \left\langle \nabla \tilde{R}; XEXAX \nabla \tilde{R} \right\rangle$$

$$+ \left\langle \nabla \tilde{R}; (XEXA) \mathbf{H} \right\rangle \Bigg] dxdt$$

By Schwarz's inequality the sum of the first four terms in the integrand is less than or equal to a fixed constant times

$$\langle \mathbf{H}; (a\chi - AX) (\chi a \chi)^{-1} (\chi a - XA) \mathbf{H} \rangle \tag{5.20}$$

This is quadratic in $\tilde{R} - \tilde{\rho}$ with a bounded constant because ρ and R are bounded away from 1. The same is true of the last two terms because E is quadratic in $\tilde{R} - \tilde{\rho}$. We can therefore find a constant C depending only on $\|\nabla \tilde{R}\|_\infty$, the lower bound δ for R_j, $1 - R$, and $1 - \rho$, and the Lipschitz norm of $S(\rho)$ such that the integrand in (5.19) is bounded by

$$C \left[h(\tilde{\rho}) - \tilde{\rho} \cdot \Lambda + F(\Lambda) \right]. \tag{5.21}$$

The uniqueness follows from Gronwall's inequality.

The main drawback of this uniqueness theorem is that it requires uniform bounds (5.6) and (5.7) on $\nabla \tilde{R}$, $\tilde{R}$ and $(1 - R)$. It forces the initial condition to satisfy these bounds uniformly in x. In particular if we consider the class of profiles $\tilde{\rho}(\cdot, \cdot)$ that have a given "bad" initial profile $\tilde{\rho}_0(\cdot)$, it may contain no profiles from the class $\mathcal{D}_n^1$. This can remedied by allowing solutions of (5.1) with $\mathbf{H} \equiv 0$. Suppose $\mathbf{H}$ is identically zero for $t \in [0, \delta]$ for some $\delta > 0$, and is uniformly bounded on $[\delta, T] \times \mathbf{R}^d$. Then, because a solution of (5.1) with $\mathbf{H} \equiv 0$ is unique in the class $\mathcal{D}_n^0$, we would have uniqueness in the time interval $[0, \delta]$. Moreover the solution $\tilde{\rho}(\delta, \cdot)$ at time δ will have a bounded gradient and $\rho_j(\delta, \cdot)$ as well as $(1 - \rho(\delta, \cdot))$ will be bounded away from zero. One could then expect the uniqueness to hold up to time T because a solution would exist on $[\delta, T]$ in the class $\mathcal{D}_n^1$.

We formally state this as a theorem.

Theorem 5.3. *Suppose $\tilde{R}$ is in the class $\mathcal{D}_n^0$ and solves (5.1) with a bounded $\mathbf{H}$ that satisfies $\mathbf{H} \equiv 0$ on $[0, \delta]$ for some $\delta > 0$. Suppose in addition, for $t \geq \delta$, $\tilde{R}$ is a solution satisfying the analogues of (5.6) and (5.7), i.e. a solution in the class $\mathcal{D}_n^1$ for the interval $[\delta, T]$. Then any weak solution $\tilde{\rho}$ of (5.1), belonging to the class $\mathcal{D}_n^0$ in $[0, T]$, and having the same initial value as $\tilde{R}$ in $\mathcal{D}_n^0$, coincides with $\tilde{R}$ in the interval $[0, T]$.*

6. Approximation theorems for the n-color rate function

The initial profile $\tilde{\rho}_0(\cdot)$ will be arbitrary and fixed throughout this section. We denote by $\bar{\mathcal{E}}_n$ profiles that satisfy the initial condition

$$\tilde{\rho}(0,\cdot) = \tilde{\rho}_0(\cdot) \tag{6.1}$$

that are weakly continuous maps from $[0,T] \times \mathbf{T}^d \to \mathcal{C}$, and have a finite value for the rate function

$$I_{dyn}(\tilde{\rho}) = \frac{1}{2}\int_0^T \|\frac{\partial \tilde{\rho}}{\partial t} - \frac{1}{2}\nabla \cdot A\chi(\tilde{\rho})\nabla\tilde{\rho}\|^2_{-1,A(\tilde{\rho})} dt \tag{6.2}$$

In particular they satisfy (5.3), or equivalently (5.4). We denote by $\mathcal{E}_n^0$ the class of evolving profiles $\tilde{\rho}(\cdot,\cdot)$ that are weak solutions of,

$$\frac{\partial \tilde{\rho}}{\partial t} = \frac{1}{2}\nabla \cdot A\chi(\tilde{\rho})\nabla\tilde{\rho} + \nabla \cdot A(\tilde{\rho})\mathbf{H} \tag{6.3}$$

with the initial condition (6.1) for some bounded $\mathbf{H}$, and have the following additional properties: for some $\delta > 0$ depending on $\tilde{\rho}(\cdot,\cdot)$

$$\mathbf{H} \equiv 0 \quad on \quad [0,\delta] \tag{6.4}$$

$$\sup_{x\in\mathbf{T}^d, t\in[\delta,T]} |\nabla\tilde{\rho}(t,x)| < \infty \tag{6.5}$$

$$\inf_{x\in\mathbf{T}^d, t\in[\delta,T]} (1-\rho(t,x)) > 0 \tag{6.6}$$

$$\inf_{1\le j\le n,\ x\in\mathbf{T}^d, t\in[\delta,T]} \rho_j(t,x)) > 0. \tag{6.7}$$

$\tilde{\rho}(\cdot,\cdot)$ will then be a weak solution in the class $\mathcal{D}_n^0$ on $[0,T]$ and a weak solution in the class $\mathcal{D}_n^1$ on $[\delta,T]$. In order to complete the proof of the lower bound of section 4 we need to prove

Theorem 6.1. *For each fixed initial condition $\tilde{\rho}_0(\cdot)$, $\mathcal{E}_n^0$ is properly dense in $\bar{\mathcal{E}}_n$. That is, for any profile $\tilde{\rho}$ in $\bar{\mathcal{E}}_n$ with $I_{dyn}(\tilde{\rho}) < \infty$, there exists a sequence $\tilde{\rho}_k$ in $\mathcal{E}_n^0$ with*

$$\lim_{k\to\infty} \tilde{\rho}_k = \tilde{\rho} \tag{6.8}$$

uniformly on $[0,T]$ in the weak topology and

$$\lim_{k\to\infty} I_{dyn}(\tilde{\rho}_k) = I_{dyn}(\tilde{\rho}) \tag{6.9}$$

The proof of Theorem 6.1 is carried out in several steps each of which is formulated as a Theorem. Let us define $\mathcal{E}_n^2$ to be the subset of those $\tilde{\rho}(\cdot,\cdot) \in \bar{\mathcal{E}}_n$ that are weak solutions of

$$\frac{\partial \tilde{\rho}}{\partial t} = \frac{1}{2}\nabla \cdot A\chi(\tilde{\rho})\nabla\tilde{\rho} \tag{6.10}$$

in the interval $0 \le t \le \delta$ for some $\delta > 0$.

Theorem 6.2. *The set $\mathcal{E}_n^2$ is properly dense in $\bar{\mathcal{E}}_n$.*

Proof. Let $\rho \in \bar{\mathcal{E}}_n$ and consider the following construction. We have the solution $\tilde{R}$ of (6.10) that starts from our initial data $\tilde{\rho}_0(\cdot)$. For $\delta > 0$ let us define $\tilde{\rho}_\delta$ by the following formula

$$
\begin{aligned}
\tilde{\rho}_\delta(t,\cdot) &= \tilde{R}(t,\cdot) \qquad &for \quad 0 \le t \le \delta \\
&= \tilde{R}(2\delta - t,\cdot) \quad &for \quad \delta \le t \le 2\delta \\
&= \tilde{\rho}(t-2\delta,\cdot) \quad &for \quad 2\delta \le t \le T
\end{aligned}
\tag{6.11}
$$

where $\tilde{R}$ is the solution of (6.10) with initial condition $\tilde{\rho}_0$. It is clear that $\tilde{\rho}_\delta \to \tilde{\rho}$ as $\delta \to 0$ and we now have to estimate $I_{dyn}(\tilde{\rho}_\delta)$ as $\delta \to 0$. Because of lower semicontinuity only the upper bound has to be established. $I_{dyn}(\tilde{\rho}_\delta)$ is given as in (6.1) by an integral from 0 to T

$$
\frac{1}{2}\int_0^T \|\frac{\partial \tilde{\rho}_\delta}{\partial t} - \frac{1}{2}\nabla \cdot A\chi(\tilde{\rho}_\delta)\nabla\tilde{\rho}_\delta\|^2_{-1,A(\tilde{\rho}_\delta)}dt
\tag{6.12}
$$

as the sum over the two intervals $\int_0^{2\delta} + \int_{2\delta}^T$. The second of these is clearly the same as the integral over $[0, T-2\delta]$ in (6.2) and is dominated by $I_{dyn}(\tilde{\rho})$. We are therefore left with the task of estimating

$$
\frac{1}{2}\int_0^{2\delta} \|\frac{\partial \tilde{\rho}_\delta}{\partial t} - \frac{1}{2}\nabla \cdot A\chi(\tilde{\rho}_\delta)\nabla\tilde{\rho}_\delta\|^2_{-1,A(\tilde{\rho}_\delta)}dt
\tag{6.13}
$$

and showing that it goes to 0 as $\delta \to 0$. Since $\tilde{\rho}_\delta$ is a solution (6.10) in the interval $[0,\delta]$, the integrand is zero on that interval. Since the profile $\tilde{\rho}_\delta$ in $[\delta, 2\delta]$ is the profile in $[0,\delta]$ run backwards in time, estimating (6.13) reduces to estimating

$$
\frac{1}{2}\int_0^{\delta} \|\frac{\partial \tilde{R}}{\partial t} + \frac{1}{2}\nabla \cdot A\chi(\tilde{R})\nabla\tilde{R}\|^2_{-1,A(\tilde{R})}dt
\tag{6.14}
$$

or equivalently

$$
\frac{1}{2}\int_0^{\delta} \|\nabla \cdot A\chi(\tilde{R})\nabla\tilde{R}\|^2_{-1,A(\tilde{R})}dt = \frac{1}{2}\int_0^{\delta}\int_{\mathbf{T}^d} \langle \nabla\tilde{R}, (\chi A\chi)(\tilde{R})\nabla\tilde{R}\rangle dx dt.
\tag{6.15}
$$

It is now clear that the quantity in (6.15) tends to 0 as δ tends to 0.

If $\mathbf{H} = 0$ in an interval $[0,\delta]$ then the solution $\tilde{\rho}$ of (6.10) will have the following properties. For each δ' satisfying $0 < \delta' < \delta$, $\rho = \sum_j \rho_j$ is a solution of the heat equation with constant coefficients D and is therefore smooth and bounded away from 0 and 1 in the interval $[\delta', \delta]$. Once ρ is determined ρ_j are solutions of the linear equation

$$
\frac{\partial \rho_j}{\partial t} = \frac{1}{2}\nabla \cdot S(\rho(t,x))\nabla\rho_j + \frac{1}{2}\nabla\cdot(D - S(\rho(t,x)))\frac{\nabla\rho}{\rho}\rho_j
\tag{6.16}
$$

with $S(\rho(t,x))$ being Lipschitz and uniformly elliptic. From elliptic regularity theory this will provide us a uniform positive lower bound for ρ_j as well as a uniform upper bound on $\nabla\rho_j$ in the interval $[\delta_1,\delta]$.

Let us next define a class $\mathcal{E}_n^1$ of profiles $\tilde\rho(\cdot,\cdot)$ in $\mathcal{E}_n^2$ that possess the following properties. There are positive numbers $0<\delta_1<\delta_2<T$ such that in $[0,\delta_1]$, $\tilde\rho$ is a solution of (6.10) with initial data $\tilde\rho_0$, and in the class $\mathcal{D}_n^1$ in the interval $[\delta_1,\delta_2]$. Finally there is a positive constant α such that for every color j, $\rho_j(t,x)\ge\alpha\rho(t,x)$ uniformly on $[\delta_1,T]\times\mathbf{T}^d$. We then have,

Theorem 6.3. *The class $\mathcal{E}_n^1$ is properly dense in $\mathcal{E}_n^2$.*

Proof. Any given profile $\tilde\rho$ in $\mathcal{E}_n^2$ is an actual solution of (6.10) in some interval $[0,\delta]$. Let us take a smooth function $\alpha(t)$ that is equal to zero for $t\le\frac{1}{4}\delta$, equal to one for $\frac{3}{4}\delta\le t\le T$ and goes from zero to one in between. We define

$$\tilde\rho_k(t,\cdot)=(1-\frac{\alpha(t)}{k})\ \tilde\rho(t,\cdot)+\frac{\alpha(t)}{k}\ \frac{\tilde c}{c}\ \rho(t,\cdot) \tag{6.16}$$

Here $\tilde c$ is the vector of total densities

$$c_j=\int_{\mathbf{T}^d}\rho_j(t,x)dx \tag{6.17}$$

and

$$c=\sum_j c_j=\int_{\mathbf{T}^d}\rho(t,x)dx \tag{6.18}$$

Because of conservation of color and particles, these are independent of t. For every k, $\tilde\rho_k$ is in the class $\mathcal{E}_n^1$. We are now left with proving

$$\limsup_{k\to\infty} I_{dyn}(\tilde\rho_k)\le I_{dyn}(\tilde\rho) \tag{6.19}$$

The quantity that we need to estimate is one of two types. When we compute the time derivative in

$$\tilde e_k=\frac{\partial\tilde\rho_k}{\partial t}-\frac{1}{2}\nabla\cdot A\chi(\tilde\rho_k)\nabla\tilde\rho_k \tag{6.20}$$

we may either differentiate $\alpha(\cdot)$ or $\tilde\rho_k(\cdot,\cdot)$. They lead to the following terms:

$$\tilde e_k=\tilde e_k^{(1)}+\tilde e_k^{(2)} \tag{6.21}$$

where

$$\begin{aligned}\tilde e_k^{(1)}&=-\frac{1}{k}\ \alpha'(t)\ \tilde\rho+\frac{\tilde c}{c}\ \frac{1}{k}\ \alpha'(t)\ \rho\\&=\frac{1}{k}\alpha'(t)\ [\frac{\tilde c}{c}\ \rho-\tilde\rho\ ]\end{aligned} \tag{6.22}$$

Note that $\alpha'(t)$ is nonzero only in a time interval where ρ_j and $(1-\rho)$ are bounded away from zero, and $\|A^{-1}(\tilde{\rho}_k)\| \le C$ on that interval. On the other hand $[\frac{\tilde{c}}{c}\rho - \tilde{\rho}]$ is a vector of bounded mean zero functions. Therefore

$$\lim_{k\to\infty} \int_0^T \|\tilde{e}_k^{(1)}(t,\cdot)\|^2_{-1,A(\tilde{\rho}_k)} dt = 0 \tag{6.23}$$

We now turn our attention to $\tilde{e}_k^{(2)}$.

$$\begin{aligned} \tilde{e}_k^{(2)} &= (1-\frac{m(t)}{k})(\frac{\partial\tilde{\rho}}{\partial t} - \frac{1}{2}\nabla A\chi(\tilde{\rho})\nabla\tilde{\rho}) + (\frac{m(t)}{k})(\frac{\partial\rho}{\partial t} - \frac{1}{2}\nabla D\nabla\rho) \\ &= (1-\frac{m(t)}{k})\nabla\cdot\mathbf{G} + (\frac{\tilde{c}}{c}\frac{m(t)}{k})\nabla G \end{aligned} \tag{6.24}$$

for some $\mathbf{G} = \{G^j\}$ and $G = \sum_j G^j$. Moreover
(6.25)

$$I_{dyn}(\tilde{\rho}) = \frac{1}{2}\int_0^T \int_{\mathbf{T}^d} \left(\sum_j \frac{\langle S^{-1}(\rho)G^j, G^j\rangle}{\rho_j} + \frac{\langle [D^{-1} - (1-\rho)S^{-1}(\rho)]G, G\rangle}{\rho(1-\rho)} \right) dx dt < \infty$$

Since
(6.26)

$$I_{dyn}(\tilde{\rho}_k) \le \frac{1}{2}\int_0^T \int_{\mathbf{T}^d} \left(\sum_j \frac{\langle S^{-1}(\rho)G_k^j\ , G_k^j\rangle}{\rho_{k,j}} + \frac{\langle [D^{-1} - (1-\rho)S^{-1}(\rho)]G_k\ , G_k\rangle}{\rho(1-\rho)} \right) dx dt$$

where

$$G_k^j = (1-\frac{m(t)}{k})G^j + \frac{c_j}{c}\ \frac{m(t)}{k}\ G \quad \text{and} \quad G_k = \sum_j G_k^j = G, \tag{6.27}$$

and

$$\rho_k^j = (1-\frac{m(t)}{k})\rho_j + (\frac{m(t)}{k})\rho, \tag{6.28}$$

it is straightforward to see that (6.19) is valid.

The last step is

Theorem 6.4. *The class $\mathcal{E}_n^0$ is properly dense in $\mathcal{E}_n^1$.*

Proof. For a given $\tilde{\rho}(\cdot,\cdot)$ in $\mathcal{E}_n^1$ there are $0 < \delta_1 < \delta_2 < T$ with properties listed earlier. Let us pick a function $\epsilon(t)$ that is zero for $t \le \delta_1$ and increases smoothly to one over the interval $[\delta_1, \delta_2]$ remaining equal to one on $[\delta_2, T]$. We define $\epsilon_k(t) = \frac{1}{k}\epsilon(t)$ and take

$$\tilde{\rho}_k(t,\cdot) = \tilde{\rho}(t,\cdot) * \phi_{\epsilon_k(t)}(\cdot) \tag{6.29}$$

where ϕ_ϵ is the heat kernel on $\mathbf{T}^d$ at time ϵ^2 which we use as a mollifier. Clearly $\tilde{\rho}_k(\cdot,\cdot)$ is in $\mathcal{E}_n^0$ and in order to complete our program we need only show that

$$\limsup_{k\to\infty} I_{dyn}(\tilde{\rho}_k) \le I_{dyn}(\tilde{\rho}) \tag{6.30}$$

Before we take up the proof of theorem 6.4 we will establish lemmas 6.5-6.8.

Lemma 6.5. *Let $f \in W^{1,2}(\mathbf{T}^d)$. The following families of functions are uniformly integrable:*

$$
(6.31) \qquad
\begin{aligned}
&(i) \quad g_a(x) = \frac{|f(x+a) - f(x)|^2}{|a|^2}, \quad a \in \mathbf{R}^d; \\
&(ii) \quad h_\epsilon(x) = \epsilon^{-2} \int_{\mathbf{T}^d} \int_{\mathbf{T}^d} |f(x+y) - f(x+z)|^2 \phi_\epsilon(y)\phi_\epsilon(z) dydz, \quad \epsilon > 0; \\
&(iii) \quad \tilde{h}_\epsilon(x) = \int_{\mathbf{T}^d} \int_{\mathbf{T}^d} |f(x+y) - f(x+z)|^2 \frac{|\nabla \phi_\epsilon|^2}{\phi_\epsilon}(y)\phi_\epsilon(z) dydz, \quad \epsilon > 0.
\end{aligned}
$$

Here $|\cdot|$ is used for distance on $\mathbf{R}$.

Proof. Recall that a family of functions $\{g_\gamma\}_{\gamma \in \Gamma}$ is uniformly integrable if and only if there exists a function $\Psi : \mathbf{R} \to \mathbf{R}_+$ such that

$$
(6.32) \qquad
\begin{aligned}
&(a) \quad \lim_{|x| \to \infty} \frac{\Psi(x)}{|x|} = \infty \\
&(b) \quad \sup_{\gamma \in \Gamma} \int \Psi(g_\gamma(x))\, dx < \infty.
\end{aligned}
$$

Furthermore if g is any function in L^1 then there exists an increasing, convex function Ψ satisfying (a) such that $\Psi(|g|) \in L^1$. Choose such a Ψ for $|\nabla f|^2$. By Jensen's inequality,

$$
(6.33) \qquad
\begin{aligned}
\int \Psi(g_a(x))dx &= \int \Psi\left(\frac{|f(x+a) - f(x)|^2}{|a|^2}\right) dx \\
&\leq \int \Psi\left(\left[\int_0^1 |\nabla f(x+ta)|dt\right]^2\right) dx \\
&\leq \int \int_0^1 \Psi\left(|\nabla f(x+ta)|^2\right) dtdx \\
&= \int \Psi\left(|\nabla f|^2\right) dx.
\end{aligned}
$$

The last term has a bound independent of a. Therefore the family $\{g_a\}$, $a \in \mathbf{T}^d$ is uniformly integrable. In a similar fashion,

(6.34)

$$
\begin{aligned}
\int \Psi\left(h_\epsilon(x)\right) dx &= \int \Psi\left(\epsilon^{-2} \int\int |f(x+y) - f(x+z)|^2 \phi_\epsilon(y)\phi_\epsilon(z)dydz\right) dx \\
&= \int \Psi\left(\int_{\mathbf{R}^d}\int_{\mathbf{R}^d} \frac{|f(x+y) - f(x+z)|^2}{|y-z|^2} \left\{\frac{|y-z|^2}{\epsilon^2}\phi_\epsilon(y)\phi_\epsilon(z)\right\} dydz\right) dx
\end{aligned}
$$

and
(6.35)

$$
\begin{aligned}
&\int \Psi\left(\tilde{h}_\epsilon(x)\right) dx \\
&\quad = \int \Psi\left(\int_{\mathbf{T}^d}\int_{\mathbf{T}^d} |f(x+y)-f(x+z)|^2 \frac{|\nabla\phi_\epsilon|^2}{\phi_\epsilon}(y)\phi_\epsilon(z)dydz\right) dx \\
&\quad \leq \int \Psi\left(\int_{\mathbf{R}^d}\int_{\mathbf{R}^d} \frac{|f(x+y)-f(x+z)|^2}{|y-z|^2}\left\{|y-z|^2\frac{|\nabla\phi_\epsilon|^2}{\phi_\epsilon}(y)\phi_\epsilon(z)\right\} dydz\right) dx
\end{aligned}
$$

An easy computation shows that there is a constant C independent of ϵ such that

$$
\int_{\mathbf{R}^d}\int_{\mathbf{R}^d} |y-z|^2 \frac{|\nabla\phi_\epsilon|^2}{\phi_\epsilon}(y)\phi_\epsilon(z)dydz \leq C
$$

The lemma follows by Jensen's inequality. The next lemma is the following.

Lemma 6.6. *Suppose that r and ρ are measurable functions mapping $\mathbf{T}^d$ into $[0,1]$ such that*

$$
\int_{\mathbf{T}^d} (1-\rho)\frac{|\nabla r|^2}{r} dx < \infty, \tag{6.36}
$$

and

$$
\int_{\mathbf{T}^d} \frac{|\nabla\rho|^2}{1-\rho} dx < \infty. \tag{6.37}
$$

Then

$$
\lim_{\epsilon\to 0}\int_{\mathbf{T}^d} (1-\rho_\epsilon)\frac{|\nabla r_\epsilon|^2}{r_\epsilon} dx = \int_{\mathbf{T}^d} (1-\rho)\frac{|\nabla r|^2}{r} dx. \tag{6.38}
$$

Proof. The strategy is to bound the integrand in terms of objects which are uniformly integrable. First of all note that

$$
\begin{aligned}
&(1-\rho_\epsilon)|\nabla r_\epsilon|^2 \\
&= \left[\int_{\mathbf{T}^d} \sqrt{1-\rho_\epsilon}(x)\nabla r(x+y)\phi_\epsilon(y)dy\right]^2 \\
&= \left[\int_{\mathbf{T}^d} \sqrt{1-\rho_\epsilon}(x)\; r(x+y)\nabla\phi_\epsilon(y)dy\right]^2 \\
&\leq 2\left[\int_{\mathbf{T}^d} \sqrt{1-\rho}(x+y)r(x+y)\nabla\phi_\epsilon(y)dy\right]^2 \\
&\quad + 2\left[\int_{\mathbf{T}^d} \left(\sqrt{1-\rho_\epsilon}(x)-\sqrt{1-\rho}(x+y)\right) r(x+y)\nabla\phi_\epsilon(y)dy\right]^2.
\end{aligned} \tag{6.39}
$$

The assumptions of the lemma imply in particular that

$$\int_{\mathbf{T}} \frac{\left|\nabla\left(r\sqrt{1-\rho}\right)\right|^2}{r} dx < \infty. \tag{6.40}$$

Therefore

$$\frac{\left[\int_{\mathbf{T}^d} \sqrt{1-\rho}(x+y) r(x+y) \nabla\phi_\epsilon(y) dy\right]^2}{\int_{\mathbf{T}^d} r(x+y)\phi_\epsilon(y)dy} = \frac{\left|\left(\nabla\left(r\sqrt{1-\rho}\right)\right)_\epsilon\right|^2}{r_\epsilon}(x) \tag{6.41}$$

is uniformly integrable. Now we consider the final term of (6.39). By Schwarz's inequality,

$$\begin{aligned} &\frac{\left[\int_{\mathbf{T}^d}\left(\sqrt{1-\rho_\epsilon}(x) - \sqrt{1-\rho}(x+y)\right) r(x+y)\nabla\phi_\epsilon(y)dy\right]^2}{\int_{\mathbf{T}^d} r(x+y)\phi_\epsilon(y)dy} \\ &\leq \int_{\mathbf{T}^d} \left|\sqrt{1-\rho_\epsilon}(x) - \sqrt{1-\rho}(x+y)\right|^2 \frac{|\nabla\phi_\epsilon|^2}{\phi_\epsilon}(y)dy. \end{aligned} \tag{6.42}$$

By Jensens's inequality,

$$\begin{aligned} &\left|\sqrt{1-\rho_\epsilon}(x) - \sqrt{1-\rho}(x+y)\right|^2 \\ &\leq 2\left|\sqrt{1-\rho_\epsilon}(x) - \left(\sqrt{1-\rho}\right)_\epsilon(x)\right|^2 + 2\left|\left(\sqrt{1-\rho}\right)_\epsilon(x) - \sqrt{1-\rho}(x+y)\right|^2 \\ &\leq 2\int_{\mathbf{T}^d}\int_{\mathbf{T}^d} \left|\sqrt{1-\rho}(x+y) - \sqrt{1-\rho}(x+z)\right|^2 \phi_\epsilon(y)\phi_\epsilon(z)dydz \\ &+ 2\int_{\mathbf{T}^d} \left|\sqrt{1-\rho}(x+y) - \sqrt{1-\rho}(x+z)\right|^2 \phi_\epsilon(z)dydz. \end{aligned} \tag{6.43}$$

Therefore

$$\begin{aligned} &\frac{\left[\int_{\mathbf{T}^d}\left(\sqrt{1-\rho_\epsilon}(x) - \sqrt{1-\rho}(x+y)\right) r(x+y)\nabla\phi_\epsilon(y)dy\right]^2}{\int_{\mathbf{T}^d} r(x+y)\phi_\epsilon(y)dy} \\ &\leq 2\int_{\mathbf{T}^d}\int_{\mathbf{T}^d} \left|\sqrt{1-\rho}(x+y) - \sqrt{1-\rho}(x+z)\right|^2 \phi_\epsilon(y)\phi_\epsilon(z)dydz \int_{\mathbf{T}^d} \frac{|\nabla\phi_\epsilon|^2}{\phi_\epsilon}(y)dy \\ &+ 2\int_{\mathbf{T}^d}\int_{\mathbf{T}^d} \left|\sqrt{1-\rho}(x+y) - \sqrt{1-\rho}(x+z)\right|^2 \frac{|\nabla\phi_\epsilon|^2}{\phi_\epsilon}(y)\phi_\epsilon(z)dydz. \end{aligned} \tag{6.44}$$

By assumption $\sqrt{1-\rho} \in W^{1,2}$ and therefore, by the previous lemma, these last two terms are uniformly integrable.

Lemma 6.7. *Suppose that $I_0(\tilde{\rho}(\cdot,\cdot)) < \infty$ and, for some $\alpha > 0$, $\rho_j(t,x) \geq \alpha\, \rho(t,x)$ for all $j = 1, ..., m$, $t \in [0,T]$ and $x \in \mathbf{T}^d$. Let*

$$\rho_{j,\epsilon}(t,x) = \int_{\mathbf{T}^d} \rho_j(t,x+y)\phi_\epsilon(y)dy. \tag{6.45}$$

Then, as $\epsilon \to 0$,

$$I_0(\tilde{\rho}_\epsilon(\cdot,\cdot)) \to I_0(\tilde{\rho}(\cdot,\cdot)). \tag{6.46}$$

Proof. We write $I_0(\tilde{\rho}(\cdot,\cdot))$ explicitly as

$$\int_0^T \int_{\mathbf{T}^d} \sum_{j=1}^n \frac{\nabla\rho_j \cdot S(\rho)\nabla\rho_j}{\rho_j} + \frac{\nabla\rho \cdot D\nabla\rho}{\rho(1-\rho)} - \frac{\nabla\rho \cdot S(\rho)\nabla\rho}{\rho}\, dxdt. \tag{6.47}$$

Note that the third integrand is bounded pointwise by either of the previous two. Therefore the whole integral is finite only if each of the three integrals is finite seperately. Recall

$$\delta(1-\rho)I \le S(\rho) \le (1-\rho)D, \tag{6.48}$$

in the sense of positive definite matrices. Also by assumption, $\alpha\rho \le \rho_j \le \rho$. Hence

$$\int_0^T \int_{\mathbf{T}^d} \frac{1-\rho}{\rho}|\nabla\rho_j|^2 dxdt \le c^{-1} \int_0^T \int_{\mathbf{T}^d} \frac{\nabla\rho_j \cdot S(\rho)\nabla\rho_j}{\rho_j} dxdt < \infty. \tag{6.49}$$

By the previous lemma, $\frac{1-\rho_\epsilon}{\rho_\epsilon}|\nabla\rho_{j,\epsilon}|^2$ is uniformly integrable. Since

$$\frac{\nabla\rho_{j,\epsilon} \cdot S(\rho_\epsilon)\nabla\rho_{j,\epsilon}}{\rho_{j,\epsilon}} \le \alpha^{-1}\frac{1-\rho_\epsilon}{\rho_\epsilon}|\nabla\rho_{j,\epsilon}|^2 \tag{6.50}$$

it is also uniformly integrable. Therefore, as $\epsilon \to 0$,

$$\int_0^T \int_{\mathbf{T}^d} \frac{\nabla\rho_{j,\epsilon} \cdot S(\rho_\epsilon)\nabla\rho_{j,\epsilon}}{\rho_{j,\epsilon}} dxdt \to \int_0^T \int_{\mathbf{T}^d} \frac{\nabla\rho_j \cdot S(\rho)\nabla\rho_j}{\rho_j} dxdt. \tag{6.51}$$

For the second term, by convexity on the one hand, and lower semicontinuity on the other,

$$\int_0^T \int_{\mathbf{T}^d} \frac{\nabla\rho_\epsilon \cdot D\nabla\rho_\epsilon}{\rho_\epsilon(1-\rho_\epsilon)} dxdt \to \int_0^T \int_{\mathbf{T}^d} \frac{\nabla\rho \cdot D\nabla\rho}{\rho(1-\rho)} dxdt. \tag{6.52}$$

The third term is pointwise bounded by the first. The uniform integrability of

$$\frac{\nabla\rho_\epsilon \cdot S(\rho_\epsilon)\nabla\rho_\epsilon}{\rho_\epsilon} \tag{6.53}$$

follows.

Lemma 6.8. *Suppose that $I_{dyn}(\tilde{\rho}(\cdot,\cdot)) < \infty$ and, for some $\alpha > 0$, $\rho_j(t,x) \geq \alpha\rho(t,x)$ for all $j = 1, ..., m$, $t \in [0,T]$ and $x \in \mathbf{T}^d$. Let*

$$\rho_{j,\epsilon}(t,x) = \int_{\mathbf{T}^d} \rho_j(t,x+y)\phi_\epsilon(y)dy. \tag{6.54}$$

Then, as $\epsilon \to 0$,

$$I_{dyn}(\tilde{\rho}_\epsilon(\cdot,\cdot)) \to I_{dyn}(\tilde{\rho}(\cdot,\cdot)). \tag{6.55}$$

Proof. The finiteness of $I_{dyn}(\tilde{\rho}(\cdot,\cdot))$ implies in particular that $\frac{\partial\tilde{\rho}}{\partial t}$ can be written as $\nabla\cdot P$ for some P with $\int_0^T\int_{\mathbf{T}^d}\langle P; A^{-1}P\rangle dxdt < \infty$. Expanding the $H_{-1,A}$ norm,

$$I_{dyn}(\rho(\cdot,\cdot)) = \frac{1}{2}\int_0^T\int_{\mathbf{T}^d}\left\langle P; A^{-1}P\right\rangle dxdt - \frac{1}{2}\int_0^T\int_{\mathbf{T}^d}\langle P;\chi\nabla\rho\rangle\, dxdt + \frac{1}{8}I_0(\tilde{\rho}). \tag{6.56}$$

By assumption, $I_0(\tilde{\rho}(\cdot,\cdot)) < \infty$. Therefore each of the three terms is finite on its own. Note also that

$$\langle P;\chi\nabla\tilde{\rho}\rangle \leq \left\langle P; A^{-1}P\right\rangle + \frac{1}{4}\langle\chi\nabla\tilde{\rho}; A\chi\nabla\tilde{\rho}\rangle \tag{6.57}$$

so it suffices to show that

$$\int_0^T\int_{\mathbf{T}^d}\left\langle P; A^{-1}P\right\rangle dxdt < \infty \tag{6.58}$$

and

$$\int_0^T\int_{\mathbf{T}^d}\langle\chi\nabla\tilde{\rho}; A\chi\nabla\tilde{\rho}\rangle < \infty \tag{6.59}$$

imply the uniform integrability of

$$\left\langle P_\epsilon; A^{-1}(\tilde{\rho}_\epsilon)P_\epsilon\right\rangle \tag{6.60}$$

and

$$\langle\chi(\tilde{\rho}_\epsilon)\nabla\tilde{\rho}_\epsilon; A(\tilde{\rho}_\epsilon)\chi(\tilde{\rho}_\epsilon)\nabla\tilde{\rho}_\epsilon\rangle\,. \tag{6.61}$$

The second of these is the content of the previous lemma. To prove the first, we write it more explicitly as

$$\left\langle P; A^{-1}P\right\rangle = \sum_{j=1}^n \frac{P_j\cdot S^{-1}(\rho)P_j}{\rho_j} + \frac{\left(\sum_{j=1}^n P_j\right)\cdot\left(S^{-1}(\rho) - (D(1-\rho))^{-1}\right)\left(\sum_{j=1}^n P_j\right)}{\sum_{j=1}^n\rho_j}. \tag{6.62}$$

Now

$$\int_0^T \int_{\mathbf{T}^d} \frac{|P_j|^2}{\rho(1-\rho)} dxdt \leq \int_0^T \int_{\mathbf{T}^d} \frac{P_j \cdot S^{-1}(\rho) P_j}{\rho_j} dxdt < \infty \tag{6.63}$$

so

$$\frac{P_{j,\epsilon} \cdot S{-}1(\rho_\epsilon) P_{j,\epsilon}}{\rho_{j,\epsilon}} \leq (c\alpha)^{-1} \frac{|P_{j,\epsilon}|^2}{\rho_\epsilon(1-\rho_\epsilon)} \leq (c\alpha)^{-1} \left(\frac{|P_j|^2}{\rho(1-\rho)} \right)_\epsilon \tag{6.64}$$

is uniformly integrable. Also

$$\frac{\left(\sum_{j=1}^n P_{j,\epsilon}\right) \cdot S^{-1}(\rho_\epsilon) \left(\sum_{j=1}^n P_{j,\epsilon}\right)}{\sum_{j=1}^n \rho_{j,\epsilon}} \leq \sum_{j=1}^n \frac{P_{j,\epsilon} \cdot S^{-1}(\rho_\epsilon) P_{j,\epsilon}}{\rho_{j,\epsilon}} \tag{6.65}$$

and

$$\frac{\left(\sum_{j=1}^n P_{j,\epsilon}\right) \cdot (D(1-\rho_\epsilon))^{-1} \left(\sum_{j=1}^n P_j\right)}{\sum_{j=1}^n \rho_{j,\epsilon}} \leq const. \sum_{j=1}^n \frac{|P_{j,\epsilon}|^2}{\rho_\epsilon(1-\rho_\epsilon)} \tag{6.66}$$

so these terms are also uniformly integrable.

We now return to complete the proof of theorem 6.4. When we start with $\tilde{\rho}$ and convolute it with a time dependent mollifier $\phi_{\frac{\epsilon(\cdot)}{k}}$ to obtain $\tilde{\rho}_k$ in the calculation of $I_{dyn}(\tilde{\rho}_k)$ there are two types of terms. One of them for each time t is obtained by differentiating with respect to t in the $\tilde{\rho}$ and this leads to exactly the type of term handled by lemma 6.8. The other term obtained by differentiating with respect to the t in $\epsilon(t)$ leads to a term that is of the form

$$\frac{1}{k^2} \int_0^T [\epsilon'(t)]^2 \|\Delta\tilde{\rho}\|_{-1,A(\tilde{\rho}_k)} dt. \tag{6.67}$$

The changeover in $\epsilon(\cdot)$ from 0 to 1 occurs inside the interval $[\delta_1, \delta_2]$ where $\nabla\tilde{\rho}$ is bounded with mean 0. Moreover ρ_j and $1-\rho$ have a uniform lower bound and this implies the same bounds for $\tilde{\rho}_k$ as well, leading to a uniform bound for $\|\Delta\tilde{\rho}\|_{-1,A(\tilde{\rho}_k)}$. This completes the proof of Theorem 6.4.

7. Contraction Principle

Let us consider $\tilde{\rho}(\cdot\,,\cdot) : [0\,,T] \times \mathbf{T}^d \to \mathcal{C}$ that satisfies

$$\hat{I}_j(\tilde{\rho}(\cdot\,,\cdot)) = \int_0^T \int_{\mathbf{T}^d} (1-\rho)\frac{|\nabla \rho_j|^2}{\rho_j}\, dxdt < \infty \tag{7.1}$$

for $j = 1, \cdots, n$ and

$$\hat{I}(\rho(\cdot,\cdot)) = \int_0^T \int_{\mathbf{T}^d} \frac{|\nabla \rho|^2}{\rho(1-\rho)}\, dx\, dt < \infty \tag{7.2}$$

where $\rho = \sum_j \rho_j$.

We have the n color rate function $I_{dyn}(\tilde{\rho}(\cdot\,,\cdot))$ defined for such maps by

$$\begin{aligned} I_{dyn}(\tilde{\rho}(\cdot\,,\cdot)) &= \frac{1}{2}\int_0^T \|\frac{\partial \tilde{\rho}}{\partial t} - \frac{1}{2}\nabla A\chi(\tilde{\rho})\nabla\tilde{\rho}\|^2_{-1,A(\tilde{\rho})}\, dt \\ &= \sup_{\Phi}\Big[\int_0^T\int_{\mathbf{T}^d}\langle \frac{\partial\tilde{\rho}}{\partial t} - \mathcal{A}^*_\rho\tilde{\rho}, \Phi\rangle dxdt - \frac{1}{2}\int_0^T\int_{\mathbf{T}^d}\langle\nabla\Phi, A(\tilde{\rho})\nabla\Phi\rangle dxdt\Big] \end{aligned} \tag{7.3}$$

where the supremum is taken over all Φ that map $[0,T]\times\mathbf{T}^d$ smoothly into R^n and $\mathcal{A}^*_\rho$ is the forward or adjoint operator

$$\mathcal{A}^*_\rho = \frac{1}{2}\nabla\cdot S(\rho)\nabla + \frac{1}{2}\nabla\cdot(D - S(\rho))\frac{\nabla\rho}{\rho} \tag{7.4}$$

corresponding to the diffusion generator

$$\mathcal{A}_\rho f = \frac{1}{2}\nabla\cdot S(\rho)\nabla f + \frac{1}{2}(S(\rho) - D)\frac{\nabla\rho}{\rho}\cdot\nabla f \tag{7.5}$$

If $\tilde{\rho}$ does not satisfy either (7.1) or (7.2) we define

$$I_{dyn}(\tilde{\rho}(\cdot\,,\cdot)) = +\infty$$

The variational formula (7.3) has to be carefully interpreted. There is no apriori assumption of regularity on $\tilde{\rho}(\cdot\,,\cdot)$. Since Φ is smooth we can attempt an integration by parts and the best that we can do is

$$\begin{aligned} \int_0^T\int_{\mathbf{T}^d}(\frac{\partial\tilde{\rho}_i}{\partial t} - \mathcal{A}^*_\rho\tilde{\rho}_i)\Phi_i dxdt = &\int_{\mathbf{T}^d}\Phi_i(T,x)\rho_i(T,x)\,dx - \int_{\mathbf{T}^d}\Phi_i(0,x)\rho_i(0,x)\,dx \\ &- \int_0^T\int_{\mathbf{T}^d}\frac{\partial\Phi_i}{\partial t}(t,x)\rho_i(t,x)\,dx\,dt \\ &+ \frac{1}{2}\int_0^T\int_{\mathbf{T}^d}\nabla\rho_i\cdot S(\rho)\nabla\Phi_i\,dx\,dt \\ &+ \frac{1}{2}\int_0^T\int_{\mathbf{T}^d}(D - S(\rho))\frac{\nabla\rho}{\rho}\cdot\nabla\Phi_i(t,x)\rho_i(t,x)\,dx\,dt. \end{aligned}$$

We therefore need to have some information on $\nabla \rho_i$. It turns out that (7.1) and (7.2) are sufficient. If we have an upper bound C on $I_{dyn}(\tilde{\rho})$ then we can get bounds on

$$\hat{I}_j(\tilde{\rho}(\cdot\,,\cdot)) = \int_0^T \int_{\mathbf{T}^d} (1-\rho)\frac{|\nabla \rho_j|^2}{\rho_j}\, dxdt \le C' \tag{7.6}$$

for $j = 1\cdots, n$, and

$$\hat{I}(\rho(\cdot,\cdot)) = \int_0^T \int_{\mathbf{T}^d} \frac{|\nabla \rho|^2}{\rho(1-\rho)}\, dx\, dt \le C' \tag{7.7}$$

with C' depending only on C.

We start with a dual variational formula. Let $\mathcal{G}_{\tilde{\rho}} = \{(b, b_1, \dots, b_n)\}$ that satisfy

$$\frac{\partial \rho_j}{\partial t} - \mathcal{A}^*_{\rho}\, \rho_j \; + \nabla\cdot \rho_j (1-\rho) Db + \nabla \cdot \rho_j\, b_j = 0\ ;\ j = 1, \dots, n \tag{7.8}$$

$$\mathcal{E}[b] = \frac{1}{2}\int_0^T \int_{\mathbf{T}^d} \rho(1-\rho) < b\,,\, Db > \; dx\, dt < \infty. \tag{7.9}$$

$$\mathcal{E}_n[b_1 \cdots, b_n] = \sum_{j=1}^n \frac{1}{2}\int_0^T \int_{\mathbf{T}^d} < b_j\,, S^{-1}(\rho) b_j > \; \rho_j\, dxdt < \infty \tag{7.10}$$

$$\sum_{j=1}^n \rho_j\, b_j \equiv 0 \tag{7.11}$$

We note that $(b, b_1 \cdots, b_n) \in \mathcal{G}_{\tilde{\rho}}$ implies in particular that

$$\frac{\partial \rho}{\partial t} - \frac{1}{2}\nabla D \nabla \rho = -\nabla \cdot \rho(1-\rho) Db \tag{7.12}$$

or $b \in \mathcal{B}_\rho$ where we define $\mathcal{B}_\rho$ as those b that satisfy (7.9) and (7.12).

Theorem 7.1. *For any* $\tilde{\rho} : [0, T] \times \mathbf{T}^d \to \mathcal{C}$

(7.13)

$$\begin{aligned} I_{dyn}(\tilde{\rho}) &= \sup_{\Phi}\left[\int_0^T \int_{\mathbf{T}^d} \langle \frac{\partial \tilde{\rho}}{\partial t} - \mathcal{A}^*_\rho \tilde{\rho}, \Phi\rangle dxdt - \frac{1}{2}\int_0^T \int_{\mathbf{T}^d} \langle \nabla\Phi ; A(\tilde{\rho})\nabla\Phi\rangle dxdt \right] \\ &= \inf_{(b, b_1, \cdots, b_n)\in \mathcal{G}_{\tilde{\rho}}} \frac{1}{2}\int_0^T \int_{\mathbf{T}^d} \left\{ \sum_{j=1}^n \rho_j < b_j, S^{-1}(\rho)\, b_j > + \rho(1-\rho)\; < b\,, Db > \right\} dx\, dt \end{aligned}$$

Proof. If $\tilde{\rho} \in \mathcal{E}_n^0$, we can solve

$$\frac{\partial \tilde{\rho}}{\partial t} - \mathcal{A}_\rho^* \tilde{\rho} = \nabla A(\tilde{\rho}) \nabla H$$

with a function $H \in L_2([0,T]; H_1)$. The two sides can then be shown to be equal by using the fact that the matrix $A(\rho)$ can be explicitly inverted and its inverse is seen to correspond to the quadratic form

$$\langle \mathbf{f}; A^{-1}(\rho)\mathbf{f} \rangle = \sum_{j=1}^{n} \frac{f_j \cdot S^{-1}(\rho) f_j}{\rho_j} - \frac{f \cdot S^{-1}(\rho) f}{\rho} + \frac{f \cdot D^{-1} f}{\rho(1-\rho)}$$

where $f = \sum_j f_j$. The clue to the explicit inversion is the fact that $A(\rho)$ is a rank one perturbation of $S(\rho) \otimes \Lambda$ where $\Lambda = \mathrm{diag}\{\rho_1, \cdots, \rho_n\}$.

It is not hard to show that the left-hand side of (7.13) is less than its right-hand side. For the other direction, let us first note that the quantity defined by the formula on the left-hand of (7.13) is lower semicontinuous as a functional of $\tilde{\rho}$. We now show the right hand side is lower semicontinuous as well. Let us take a sequence $\tilde{\rho}^m$ converging weakly to $\tilde{\rho}$ with uniform bound ℓ for the right-hand side. As a result, we have a uniform bound ℓ for the left-hand side which in turn imply

$$\int_0^T \int_{\mathbf{T}^d} \frac{|\nabla \rho^m|^2}{\rho^m(1-\rho^m)} \, dx \, dt \le C,$$

$$\int_0^T \|\frac{\partial \rho}{\partial t}\|_{-1}^2 + \|\frac{\partial \rho_i^m}{\partial t}\|_{-1}^2 dt \le C,$$

$$\int_0^T \int_{\mathbf{T}^d} (1-\rho^m) \frac{|\nabla \rho_i^m|^2}{\rho_i^m} \, dx \, dt \le C.$$

for some constant $C < \infty$. We then use Lemma 4.2 to deduce the convergence in $L^1([0\,,T] \times \mathbf{T}^d)$ of ρ^m to ρ as well as that of $(1-\rho^m)\rho_i^m$ to $(1-\rho)\rho_i$

Since $S(\rho) \le C(1-\rho)$ we obtain a bound on

$$\int_0^T \int_{\mathbf{T}^d} \frac{\rho_i^m}{(1-\rho^m)} \, |b_i^m|^2 \, dx \, dt$$

and therefore we can take a weak limit c_i in $L_2[[0\,,T] \times \mathbf{T}^d]$ of

$$c_i^m = \sqrt{\frac{\rho_i^m}{(1-\rho^m)}} \, b_i^m$$

and write

$$b_i = \sqrt{\frac{(1-\rho)}{\rho_i}} c_i$$

on the set where $\rho_i > 0$ and 0 elsewhere. Since we can write

$$\rho_i^m\, b_i^m = \sqrt{\rho_i^m (1-\rho^m)}\, c_i^m$$

we have the weak convergence in $L_1[[0\,,T]\times \mathbf{T}^d\,]$ of $\rho_i^m\, b_i^m$ to $\rho_i\, b_i$. It is now easy to produce a b such that $(b, b_1, \cdots, b_n) \in \mathcal{G}_{\tilde\rho}$ and

$$\frac{1}{2}\sum_j \int_0^T \int_{\mathbf{T}^d} \rho_j\, b_j \cdot S^{-1}(\rho)\, b_j\, dx\, dt + \frac{1}{2}\int_0^T \int_{\mathbf{T}^d} \rho(1-\rho)\, b\cdot Db\, dx\, dt \le \ell$$

thereby proving the lower semicontinuity of right-hand side.

By Theorem 6.1 we can approximate arbitrary $\tilde\rho$ nicely by $\tilde\rho^m \in \mathcal{E}_n^0$ so that the right hand side converges to the corresponding value of the right hand side. By lower semicontinuity of right hand side, we must have that the right hand side is less than or equal to the left hand side in the limit. The reverse inequality is trivial and we are done.

The following theorem summarizes results that can be found in [QV] on the question of existence and uniqueness of the diffusion processes that we need.

Theorem 7.2. *For each $b \in \mathcal{B}_\rho$ we have a unique diffusion process P^b that has $\rho(t,\cdot)$ as its marginal density at time t and has as its generator*

$$\mathcal{A}_{\rho,b} = \frac{1}{2}\nabla\cdot S(\rho(t,x))\nabla + \frac{1}{2}(S(\rho(t,x)) - D)\frac{(\nabla\rho)(t,x)}{\rho(t,x)}\cdot\nabla + (1-\rho)Db\cdot\nabla. \tag{7.14}$$

The family $\{P^b : b\in\mathcal{B}_\rho\}$ has the following additional properties:

By the orthogonality relations, the minimizer b_0 of $\mathcal{E}[b]$ in the class $\mathcal{B}_\rho$, is of gradient type. We have

$$\mathcal{E}[b] = \mathcal{E}[b_0] + \mathcal{E}[b-b_0]$$

and of course $\nabla\cdot\rho(1-\rho)D(b-b_0) = 0$ for any other $b\in\mathcal{B}_\rho$.

The relative entropy $H(P^b, P^{b_0})$ satisfies

$$\begin{aligned} H(P^b; P^{b_0}) &= \frac{1}{2}\int_0^T\int_{\mathbf{T}^d} (S^{-1}(\rho)(1-\rho)D(b-b_0)\,\cdot\,(1-\rho)D(b-b_0))\rho\, dx\, dt \\ &\le C\mathcal{E}[b-b_0] \end{aligned}$$

for some constant C, because of the lower bound

$$S(\rho) \ge c(1-\rho)D$$

which is valid with some positive constant c.

Any weak solution of the forward equation

$$\frac{\partial R}{\partial t} = \mathcal{A}^*_{\rho,b} R$$

satisfying the bound

$$0 \le R(t,x) \le C\rho(t,x) \tag{7.15}$$

for all $(t,x) \in [0,T] \times \mathbf{T}^d$ *with some constant* C*, also satisfies the energy estimate*

$$\int_0^T \int_{\mathbf{T}^d} (1-\rho)\frac{|\nabla R|^2}{R}\, dx\, dt \le C'$$

where $C' < \infty$ *depends only on the constant* C *in equation (7.15), the bound on the gradient estimate on* ρ *specified in equation (7.2) and a bound on the quantity* $\mathcal{E}[b]$ *defined in equation (7.9).*

In addition let $0 \le R(t\,,x) \le 1$ *satisfy*

$$\int_0^T \int_{\mathbf{T}^d} (1-\rho)\frac{|\nabla R|^2}{R} < \infty \tag{7.16}$$

and be a weak solution of

$$\frac{\partial R}{\partial t} = \mathcal{A}^*_{\rho\,,b\,,\tilde{b}} R.$$

Here $\mathcal{A}^*_{\rho\,,b\,,\tilde{b}}$ *is the adjoint of*

$$\mathcal{A}_{\rho\,,b\,,\tilde{b}} = \mathcal{A}_{\rho\,,b} + \tilde{b}\cdot\nabla$$

with a $\tilde{b}$ *satisfying the bound*

$$\frac{1}{2}\int_0^T \int_{\mathbf{T}^d} (\tilde{b}\,\cdot\, S^{-1}(\rho)\tilde{b})\, R\, dxdt < \infty$$

Then there is a unique diffusion corresponding to $\mathcal{A}_{\rho\,,b\,,\tilde{b}}$ *that has* $R(\cdot\,,\cdot)$ *for its marginals. If we denote this diffusion by* $P^{b\,,\tilde{b}}$*, it is absolutely continuous with respect to* $P^b_{R_0}$ *the diffusion corresponding to* $\mathcal{A}_{\rho\,,b}$*, but with initial density* $R(0\,,\cdot)$*.*

We will use the next theorem in the following two contexts.

1. We will have ρ satisfying (7.2) and $\mathcal{L}_t$ will be the operator given in (7.4), with $b \in \mathcal{B}_\rho$, written in nondivergence form. The process Q will have ρ as one dimensional marginals and will be absolutely continuous with respect to P^b with finite relative entropy $H(Q,P^b)$ for some $b \in \mathcal{B}_\rho$. We will find a unique $b_Q \in \mathcal{B}_\rho$ such that

$$H(Q\,,P^{b_Q}) \le H(Q\,,P^b) < \infty$$

and the process P^{b_Q} satisfies

$$E^Q\left[\int_0^T \phi(t\,,x(t))\,\cdot\, dx(t)\right] = E^{P^{b_Q}}\left[\int_0^T \phi(t\,,x(t))\,\cdot\, dx(t)\right] \tag{7.17}$$

for all smooth ϕ.

2. In addition to ρ satisfying (7.2) and $b \in \mathcal{B}_\rho$, we will have a function $c : [0\,,T] \times \mathbf{T}^d \to R^d$ satisfying

$$\int_0^T \int_{\mathbf{T}^d} (c \cdot S^{-1}(\rho) c)\, dx dt < \infty.$$

The process $P^{b\,,c}$ will correspond to some initial distribution $R(0\,,\cdot)$ satisfying $R(0\,,\cdot) \le \rho(0\,,\cdot)$ and generator $\mathcal{L}_t$, the nondivergence form version of

$$\mathcal{A}_{\rho\,,b\,,c} = \frac{1}{2} \nabla \cdot S(\rho) \nabla + (S(\rho) - D) \frac{\nabla \rho}{\rho} \cdot \nabla + (1 - \rho) D b \cdot \nabla + c \cdot \nabla$$

The process Q will have marginals $R(\cdot\,,\cdot)$ coinciding at time 0 with the initial density $R(0\,,\cdot)$ of $P^{b\,,c}$. Moreover $R(\cdot\,,\cdot)$ will satisfy (7.16). The process Q will also be absolutely continuous with respect to $P^{b\,,c}$ with finite relative entropy $H(Q\,, P^{b\,,c})$. We want to replace c by c_Q, so that

$$E^Q \left[\int_0^T \phi(t\,, x(t)) \cdot dx(t) \right] = E^{P^{b\,,c_Q}} \left[\int_0^T \phi(t\,, x(t)) \cdot dx(t) \right] \tag{7.18}$$

with $H(Q\,, P^{b\,,c_Q}) \le H(Q\,, P^{b\,,c}) < \infty$ and

$$\frac{1}{2} \int_0^T \int_{\mathbf{T}^d} ((c - c_Q) \cdot S^{-1}(\rho)(c - c_Q)) R(t\,, x)\, dx dt < \infty$$

Let P be a diffusion process on $[0\,,T]$ with marginals $p(t\,,\cdot)$ that corresponds to the generator

$$\mathcal{L}_t = \frac{1}{2} \sum_{i,j=1}^d a_{i,j}(t, x) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{j=1}^d b_j(t\,, x) \frac{\partial}{\partial x_j}.$$

We will assume that

$$\int_0^T \int_{\mathbf{T}^d} [\sum_{i,j=1}^d |a_{i,j}(t\,, x)| + \sum_{j=1}^d |b_j(t, x)|\,]\, p(t\,, x)\, dx dt < \infty$$

so that

$$y(t) = \int_0^t dx(s) - \int_0^t b(s\,, x(s))\, ds$$

is well defined as a square integrable martingale with quadratic variation

$$\int_0^t a_{i,j}(s\,, x(s)) ds.$$

Let Q be the law of another process that has marginals $q(t\,,\cdot)$. Suppose that Q is absolutely continuous with respect to P and the relative entropy $H(Q\,, P)$ is finite. Assume further that

$$\int_0^T \int_{\mathbf{T}^d} [\sum_{i,j=1}^d |a_{i,j}(t\,, x)| + \sum_{j=1}^d |b_j(t, x)|\,]\, q(t\,, x)\, dx dt < \infty.$$

We then have,

Theorem 7.3. *There exist coefficients $\tilde{b}(t\,,x)$ and a diffusion process $\widehat{P}$ with generator*

$$\tilde{\mathcal{L}}_t = \frac{1}{2}\sum_{i,j=1}^{d} a_{i,j}(t,x)\frac{\partial^2}{\partial x_i \partial x_j} + \sum_{j=1}^{d}[b_j + \tilde{b}_j](t\,,x)\frac{\partial}{\partial x_j}$$

such that

$$E^P\Big[\int_0^T f(t\,,x(t))\,dt\, + \phi(t\,,x(t))\,\cdot\,dx(t)\Big]$$
$$= E^{\widehat{P}}\Big[\int_0^T f(t\,,x(t))\,dt\, + \phi(t\,,x(t))\,\cdot\,dx(t)\Big]$$

for all smooth functions f and ϕ. In particular $\widehat{P}$ has the same marginals $p(t\,,\cdot)$ as P. We have the estimate

$$H(\widehat{P}\,,Q) = \frac{1}{2}\int_0^T\int_{\mathbf{T}^d}(\tilde{b}\,\cdot\,a^{-1}\,\tilde{b}\,)\,p(t\,,x)\,dxdt \leq H(P\,,Q) \tag{7.19}$$

Proof: The proof is fairly routine. We note that

$$E^P\Big[\exp\Big\{\int_0^T \phi(t\,,x(t))\,\cdot\,dy(t)\,-\frac{1}{2}\int_0^T(\phi\,\cdot\,a\,\phi)(t\,,x(t))\,dt\,\Big\}\Big] = 1$$

and by Jensen's inequality

$$E^Q\Big[\int_0^T \phi(t\,,x(t))\,\cdot\,dy(t)\,-\frac{1}{2}\int_0^T(\phi\,\cdot\,a\,\phi)(t\,,x(t))\,dt\Big] \leq H(Q\,,P)$$

providing us with an estimate

$$\Big|E^Q\Big[\int_0^T \phi(t\,,x(t))\,\cdot\,dy(t)\Big]\Big| \leq H(Q\,,P) + \frac{1}{2}\int_0^T\int_{\mathbf{T}^d}(\phi\,\cdot\,a\,\phi)(t\,,x)\,q(t\,,x)\,dxdt$$

By the Riesz representation theorem there is a $\tilde{b}$ satisfying (7.19) for which

$$E^Q\left[\int_0^T \phi(t\,,x(t))\,\cdot dy(t)\right] = \int_0^T\int_{\mathbf{T}^d}(\phi\,\cdot\,\tilde{b})(t\,,x)\,q(t\,,x)\,dxdt$$

In particular

$$E^Q\left[\int_0^T \phi(t\,,x(t))\,\cdot dx(t)\right] = \int_0^T\int_{\mathbf{T}^d}(\phi\,\cdot\,(b+\tilde{b}))(t\,,x)\,q(t\,,x)\,dxdt \tag{7.20}$$

Ito's formula is valid almost surely with respect to P and therefore almost surely with respect to Q,

$$f(t\,,x(t)) - f(0\,,x(0)) = \int_0^t [\frac{\partial f}{\partial s} + \frac{1}{2}\sum_{i,j=1}^n a_{i,j}\frac{\partial^2 f}{\partial x_i \partial x_j}](s\,,x(s))\,ds + \int_0^t \nabla f(s\,,x(s))\,\cdot\,dx(s)$$

It follows, by taking expectations with respect to Q and using (7.20), that $q(\cdot\,,\cdot)$ is a weak solution of

$$\frac{\partial q}{\partial t} = \tilde{\mathcal{L}}^* q$$

where $\tilde{\mathcal{L}}^*$ is the adjoint or forward operator corresponding to the generator

$$\tilde{\mathcal{L}}_t = \frac{1}{2}\sum_{i,j=1}^n a_{i,j}(t,x)\frac{\partial^2}{\partial x_i \partial x_j} + \sum_{j=1}^n [b_j + \tilde{b}_j](t\,,x)\frac{\partial}{\partial x_j}.$$

At this point we can assert the existence of a unique diffusion with marginals q and generator $\tilde{\mathcal{L}}_t$. We denote this by $\widehat{Q}$ and clearly $\widehat{Q}$ works.

Let a measure Q on $C[[0,T];\mathbf{T}^d]$ with total mass $\bar{\rho}$ be given. We define the following rate function $\mathcal{I}(Q)$. If the marginals of Q have densities $\rho(t,x)$ with $0 \le \rho(t,x) \le 1$ satisfying (7.2) and if Q is absolutely continuous with respect to P^b for some $b \in \mathcal{B}_\rho$ with finite relative entropy $H(Q\,,P^b)$, we define

$$\mathcal{I}(Q) = H(Q,P^{b_Q}) + \mathcal{E}[b_Q]$$

where b_Q and P^{b_Q} are as in (7.17). Otherwise $\mathcal{I}(Q) = +\infty$.

It will be convenient to introduce a more general version of the same rate function. Given n measures $Q_1,\ldots,Q_n$ and their sum $Q = \sum_{j=1}^n Q_j$, assume that their marginal densities $\rho_1,\ldots,\rho_n$ and ρ satisfy (7.1) and (7.2). Let us construct $b_Q \in \mathcal{B}_\rho$ and P^{b_Q} as in (7.17). Let $P_j^{b_Q}$ be the diffusion corresponding to the generator $\mathcal{A}_{\rho\,,b_Q}$ with initial density that given by Q_j at time 0. We define

$$\mathcal{I}_n(Q_1,\ldots,Q_n) = \sum_{j=1}^n H(Q_j\,,P_j^{b_Q}) + \mathcal{E}[b]$$

when it is well defined. Otherwise $\mathcal{I}_n(Q_1,\ldots,Q_n) = +\infty$. Note that for $n = 1$ it reduces to $\mathcal{I}(Q)$.

We now establish some properties of the functional $\mathcal{I}$.

Theorem 7.4. *The functional $\mathcal{I}(Q)$ is lower semicontinuous and has compact level sets.*

Proof: Let us suppose that we have a sequence Q_k such that $I(Q_k) \le \ell$. We denote by $\rho_k(t,x)$ the marginal densities of Q_k. Let us denote b_{Q_k} by b_k and P^{b_k} by P_k. From the definition of $I(Q)$ we have the following relations

$$\frac{\partial \rho_k}{\partial t} = \frac{1}{2}\nabla\cdot D\nabla\rho_k - \nabla\cdot\,\rho_k(1-\rho_k)Db_k \tag{7.21}$$

and

$$\mathcal{E}[b_k] = \frac{1}{2}\int_0^T \int_{\mathbf{T}^d} \rho_k(1-\rho_k)(b_k \cdot Db_k)\,dx\,dt$$

and

$$H(Q_k\,, P_k) + \mathcal{E}[b_k] \le \ell. \tag{7.22}$$

We write the generator of P_k in nondivergence form

$$\mathcal{A}_k = \frac{1}{2}S(\rho_k(t,x))\nabla\cdot\nabla + c_k(t,x)\cdot\nabla$$

with

$$c_k = \frac{1}{2}S'(\rho_k)\nabla\rho_k + \frac{1}{2}[S(\rho_k) - D]\frac{\nabla\rho_k}{\rho_k} + (1-\rho_k)Db_k$$

We shall complete the proof of the theorem in several steps.
Step1: The bound (7.22) yields the following estimate on $\nabla\rho_k(t,x)$:

$$\sup_k \int_0^T \int_{\mathbf{T}^d} \frac{|\nabla\rho_k(t,x)|^2}{\rho_k(t,x)(1-\rho_k(t,x))}\,dxdt \le C_\ell \tag{7.23}$$

with a constant C_ℓ that depends only on ℓ. This in turn provides the estimate,

$$\sup_k \int_0^T \int_{\mathbf{T}^d} |c_k(t,x)|^2\rho_k(t.x)\,dt\,dx < \infty \tag{7.24}$$

For each k, P_k is a diffusion process on $C[[0,T];\mathbf{T}^d]$ with the generator given by (7.22). In nondivergence form, its second order diffusion coefficients are uniformly bounded by the bound on $S(\rho)$ and and its first order term satisfies (7.24), $\rho_k(t,x)$ being the marginal density of P_k at time t. This establishes the compactness of the sequence P_k.

Since we have a uniform bound on the relative entropy $H(Q_k\,, P_k)$ the compactness of Q_k follows. Using the bound (7.21) and the weak compactness of Q_k and P_k, we can take a subsequence and assume that weak limits b of b_k and Q and P of Q_k and P_k exist. (The precise definition of b will be given after (7.27) in Step 2.) Since P_k and Q_k have the same marginal densities ρ_k, the same is true of their weak limits P and Q and we denote their their common marginal by $\rho(t,x)$. From lower semicontinuity it will follow that

$$H(Q\,,P) + \frac{1}{2}\int_0^T \int_{\mathbf{T}^d} (b(t,x) \cdot Db(t,x))\rho(t,x)(1-\rho(t,x))\,dt\,dx \le \ell$$

The proof of the theorem will be completed by showing that $b \in \mathcal{B}_\rho\,$, $P = P^b$ and $b_Q = b$ which will establish that $I(Q) \le \ell$.
Step 2: In order to simplify notation we will ignore the need for a subsequence and assume that all weak limits exist. From (7.23) we get the estimate

$$\sup_k \int_0^T \int_{\mathbf{T}^d} |\nabla\rho_k(t,x)|^2\,dxdt \le C_\ell \tag{7.25}$$

¿From (7.21-23), we get the estimate

$$\sup_k \int_0^T \| \frac{\partial \rho_k}{\partial t} \|_{-1}^2 \, dt < \infty \tag{7.26}$$

On account of Lemma 4.2, the estimates (7.25) and (7.26) are sufficient to obtain strong convergence of ρ_k to ρ in $L^1([0,T] \times \mathbf{T}^d)$. It is not hard to show that the sequence $\sqrt{\rho_k(1-\rho_k)}$ also converges to $\sqrt{\rho(1-\rho)}$ in L^1. We can use this to take some sort of weak limit of b_k. First, by the uniform L^1 bound on the $b_k^2 \rho_k (1-\rho_k)$, a subsequnce of $b_k \sqrt{\rho_k(1-\rho_k)}$ converges weakly in L^2 to a function h. We claim that indeed $h = b\sqrt{\rho(1-\rho)}$ for some function b. To see this, observe that by the strong convergence of $\sqrt{\rho_k(1-\rho_k)}$, we also that the weak convergence of $b_k \rho_k (1-\rho_k)$ to $u = h\sqrt{\rho(1-\rho)}$ and evidently u vanishes whenever $\rho = 0$ or 1. Hence b can be defined to be $u[\rho(1-\rho)]^{-1}$.

Since the convergence of ρ_k to ρ is strong in L^1 there is no problem in passing to the limit in

$$\frac{\partial \rho_k}{\partial t} - \frac{1}{2} \nabla D \nabla \rho_k = -\nabla \cdot \rho_k (1-\rho_k) D b_k$$

and showing that $b \in \mathcal{B}_\rho$. Moreover, if R_k is a solution of

$$\frac{\partial R_k}{\partial t} = \mathcal{A}_k^* R_k \, , \tag{7.27}$$

with $0 \leq R_k \leq C\rho_k$, similar estimates i.e. (7.6) and (7.16) are valid. These imply

$$\sup_k \int_0^T \int_{\mathbf{T}^d} |\nabla \left(R_k \sqrt{\frac{1-\rho_k}{\rho_k}} \right) |^2 \, dt \, dx < \infty. \tag{7.28}$$

We can use the equation (7.27) to get an estimate on $\frac{\partial R_k}{\partial t}$.

$$\sup_k \int_0^T \| \frac{\partial R_k}{\partial t} \|_{-1}^2 \, dt < \infty. \tag{7.29}$$

Estimates (7.6),(7.28) and (7.29) imply

$$\lim_{k \to \infty} \int_0^T \int_{\mathbf{T}^d} \left| \sqrt{\frac{1-\rho_k(t,x)}{\rho_k(t,x)}} R_k(t,x) - \sqrt{\frac{1-\rho(t,x)}{\rho(t,x)}} R(t,x) \right| \, dt \, dx = 0.$$

We can now pass to the limit in equation (7.27) and see that

$$\frac{\partial R}{\partial t} = \mathcal{A}_{\rho,b}^* R \tag{7.30}$$

From the uniqueness of solutions to (7.30), we conclude that $P = P^b$.

Step 3: The last step is the identification $b = b_Q$. This requires establishing the relation

$$E^Q\left[\int_0^T \phi(t,x(t)) \cdot dx(t)\right] = E^P\left[\int_0^T \phi(t,x(t)) \cdot dx(t)\right]$$

by passing to the limit as $k \to \infty$ in both sides of

$$E^{Q_k}\left[\int_0^T \phi(t,x(t)) \cdot dx(t)\right] = E^{P_k}\left[\int_0^T \phi(t,x(t)) \cdot dx(t)\right].$$

The right hand side is an easy calculation because P_k is a diffusion and we know the first order term of the generator $\mathcal{A}_k$ in nondivergence form. The left hand side is somewhat more complicated because, although we know that $Q_k \Rightarrow Q$, the functional

$$F(\omega) = \int_0^T \phi(t,x(t)) \cdot dx(t)$$

is neither bounded nor continuous. We shall first establish the uniform integarability of $F(\cdot)$ with respect to $\{Q_k\}$. We write $F = G_k + (F - G_k)$, where

$$G_k(\omega) = \int_0^T \phi(t,x(t))\, cdot\, c_k(t,x(t))\, dt$$

and

$$c_k = \frac{1}{2}S'(\rho_k)\nabla\rho_k + \frac{1}{2}[S(\rho_k) - D]\frac{(\nabla\rho_k)}{\rho_k} + (1-\rho_k)Db_k$$

It is easy to bound

$$E^{Q_k}\left[\left|G_k(\omega)\right|^2\right] \le TC(\phi)\int_0^T \int_{\mathbf{T}^d} |c_k|^2 \rho_k \, dt\, dx$$

by a quantity independent of k. The functional $F - G_k$ is a stochastic integral with respect to a continuous martingale of uniformly bounded quadratic variation with respect to P_k having a uniformly bounded integrand. It has finite exponential moments with any constant relative to P_k. Since we have uniform bound on the relative entropy $H(Q_k, P_k)$, we obtain the uniform integrability of $(F - G_k)$ with respect to Q_k. Finally for $h > 0$ we will construct continuous functionals F_h such that

$$\lim_{h\to 0}\sup_k E^{Q_k}\left[\left|F_h(\omega) - F(\omega)\right|\right] = 0$$

This will complete the proof.

Let us do the usual approximation by simple functions by defining

$$F_h(\omega) = \int_0^T \phi(\tau_h(t), x(\tau_h(t))) \cdot dx(t)$$

where

$$\tau_h(t) = rh \quad \text{if} \quad rh \le t < (r+1)h \quad \text{for} \quad \text{integral} \quad r \ge 0.$$

If we use the earlier decomposition

$$dx(t) = c_k(t, x(t))dt + [dx(t) - c_k(t, x(t))dt] = dz_k(t) + dy_k(t)$$

and split our approximations as

$$F_h(\omega) = \int_0^T \phi\left(\tau_h(t),\, x(\tau_h(t))\right) \cdot dz_k(t) + \int_0^T \phi\left(\tau_h(t), x(\tau_h(t))\right) \cdot dy_k(t)$$

and correspondingly

$$F(\omega) = \int_0^T \phi(t\,, x(t)) \cdot dz_k(t) + \int_0^T \phi(t\,, x(t)) \cdot dy_k(t)$$

it is routine to estimate the difference between the two first terms and control it by the tightness of the family Q_k and the bound

$$\sup_k \int_0^T \int_{\mathbf{T}^d} |c_k(t,x)|^2 \rho_k(t,x)\, dt\, dx < \infty$$

Standard methods in the theory of stochastic integration will allow us to estimate the difference between the two second terms with respect to P_k to conclude that for every $\lambda > 0$,

$$\lim_{h\to 0} \sup_k E^{P_k}\left[\exp\left\{\lambda \left|\int_0^T [\phi\left(\tau_h(t), x(\tau_h(t))\right) - \phi(t, x(t))] \cdot dy_k(t)\right|\right\}\right] = 1.$$

By the entropy bounds this implies

$$\lim_{h\to 0} \sup_k E^{Q_k}\left[\left|\int_0^T [\phi(\tau_h(t), x(\tau_h(t))) - \phi(t, x(t))] \cdot dy_k(t)\right|\right] = 0$$

and we are done.

Now we turn to the relation of the rate function $\mathcal{I}_n(Q_1, \ldots, Q_n)$ to the n-color rate function $I_{dyn}(\tilde{\rho}(\cdot\,,\cdot))$ that was defined earlier. Let $\tilde{\rho}(\cdot\,,\cdot) : [0\,,T] \times \mathbf{T}^d \to \mathcal{C}$ be given, satisfying (7.1) and (7.2). For any given $\rho : [0\,,T] \times \mathbf{T}^d \to [0\,,1]$ we define $\mathcal{M}(\rho)$ to be the set of measures Q on $C[[0\,,T]\,;\mathbf{T}^d]$ that have $\rho(t\,,\cdot)$ as marginals at time t.

Theorem 7.5. *The n color rate function $I_{dyn}(\tilde{\rho}(\cdot\,,\cdot))$ is given by the contraction formula*

$$I_{dyn}(\tilde{\rho}(\cdot\,,\cdot)) = \inf_{\substack{Q_1 \in \mathcal{M}(\rho_1) \\ \cdots \\ Q_n \in \mathcal{M}(\rho_n)}} \mathcal{I}_n(Q_1, \ldots, Q_n)$$

Proof: We have the variational formula of Theorem 7.1 for $I_{dyn}(\tilde{\rho})$,

$$I_{dyn}(\tilde{\rho}) = \inf_{(b,b_1,\ldots,b_n)\in\mathcal{G}_{\tilde{\rho}}} \frac{1}{2}\int_0^T \int_{\mathbf{T}^d} \left\{ \sum_{j=1}^n \rho_j \langle b_j, S^{-1}(\rho)\, b_j\rangle + \rho(1-\rho)\,(b\,\cdot\, Db)\right\} dx\, dt$$

Because $Q_i \in \mathcal{M}(\rho_i)$ for $i = 1,\ldots,n$, we conclude that $\sum_{j=1}^n Q_j = Q \in \mathcal{M}(\rho)$. Recall that $\rho = \sum_{j=1}^n \rho_j$. Let b_Q be as in Theorem 7.3. For any diffusion process P we denote by $\mathcal{M}(P)$ the family of processes Q that satisfy $H(Q,P) < \infty$, have the same one dimensional marginals as P and satisfy in addition

$$E^Q\left[\int_0^T \phi(t,x(t))\,\cdot\, dx(t)\right] = E^P\left[\int_0^T \phi(t,x(t))\,\cdot\, dx(t)\right]$$

for all smooth test functions ϕ. We have the diffusions $P^{b_Q}, P_1^{b_Q},\ldots,P_n^{b_Q}$ with generator

$$\mathcal{A}_{\rho,b_Q} = \frac{1}{2}\nabla\cdot S(\rho(t,x))\nabla + \frac{1}{2}(S(\rho(t,x)) - D)(\frac{\nabla\rho}{\rho})(t,x)\cdot\nabla + (1-\rho(t,x))Db_Q(t,x)\cdot\nabla$$

and initial densities $\rho(0\,,\cdot), \rho_1(0\,,\cdot),\ldots,\rho_n(0\,,\cdot)$ respectively. By definition

$$\inf_{\substack{Q_1\in\mathcal{M}(\rho_1)\\ \cdots \\ Q_n\in\mathcal{M}(\rho_n)}} \mathcal{I}_n(Q_1,\ldots,Q_n) = \inf_{\substack{Q_1\in\mathcal{M}(\rho_1)\\ \cdots \\ Q_n\in\mathcal{M}(\rho_n)}} \left\{\sum_{j=1}^n H(Q_j\,,P_j^{b_Q}) + \mathcal{E}[b_Q]\right\}$$

Having already chosen b_Q, we can use Theorem 7.3 to find $(b_1,\ldots,b_n)$ so that the processes $P_j^{b_Q\,,b_j}$ with generators

$$\mathcal{A}_{\rho,b_Q,b_j} = \mathcal{A}_{\rho,b_Q} + b_j(x,t)\cdot\nabla$$

satisfy

$$\frac{\partial\rho_j}{\partial t} = \mathcal{A}^*_{\rho,b_Q,b_j}\,\rho_j$$

and $Q_j \in \mathcal{M}(P_j^{b_Q\,,b_j})$. The latter property and Girsanov Formula imply

$$E^{Q_j}\left[\log\frac{dP_j^{b_Q,b_j}}{dP_j^{b_Q}}\right] = H(P_j^{b_Q,b_j}\,,P_j^{b_Q}).$$

Hence

$$\begin{aligned} H(Q_j\,,P_j^{b_Q}) &= H(Q_j\,,P_j^{b_Q,b_j}) + H(P_j^{b_Q,b_j}\,,P_j^{b_Q}) \\ &\geq H(P_j^{b_Q,b_j}\,,P_j^{b_Q}) \\ &= \frac{1}{2}\int_0^T\int_{\mathbf{T}^d}(b_j\cdot S^{-1}(\rho)b_j)\rho_j(t\,,x)\,dxdt \end{aligned}$$

From $Q_j \in \mathcal{M}(P_j^{b_Q, b_j})$ for $1 \leq j \leq n$ and $Q \in \mathcal{M}(P^{b_Q})$ it follows that $\sum_{j=1}^n \rho_j b_j \equiv 0$. If we replace Q_j by $P_j^{b_Q, b_j}$ in the variational formula we get

$$
\begin{aligned}
&\inf_{\substack{Q_1 \in \mathcal{M}(\rho_1) \\ \cdots \\ Q_n \in \mathcal{M}(\rho_n)}} \left\{ \sum_{j=1}^n H(Q_j\,, P_j^{b_Q}) + \mathcal{E}[b_Q] \right\} \\
&\qquad \geq \inf_{(b,b_1,\ldots,b_n) \in \mathcal{G}_{\tilde\rho}} \frac{1}{2} \int_0^T \int_{\mathbf{T}^d} \left\{ \sum_{j=1}^n \rho_j (b_j \cdot S^{-1}(\rho)\, b_j) + \rho(1-\rho)\,(b \cdot Db) \right\} dx\, dt
\end{aligned}
$$

Since any choice of $(b, b_1 \ldots, b_n)$ leads to a choice of $\{P_j^b\}$ and $\{Q_j\}$, the reverse inequality is immediate. This proves the theorem.

A special situation is when $\rho_j(0\,,\cdot)$ are supported on disjoint subsets of $\mathbf{T}^d$. In this case $Q_1, \ldots, Q_n$ are mutually orthogonal. Then it is easily seen that $\mathcal{I}_n(Q_1, \ldots, Q_n) = \mathcal{I}(Q)$ where $Q = \sum_{j=1}^n Q_j$. The variational formula becomes

$$
I_{dyn}(\tilde\rho) = \inf_{\substack{Q_1 \in \mathcal{M}(\rho_1) \\ \cdots \\ Q_n \in \mathcal{M}(\rho_n)}} \mathcal{I}(Q)
$$

8. Superexponential tightness.

One of the key steps in any large deviation estimate is the superexponetial estimate. If Q_n is a family of measures on a Polish space X, the natural definition of superexponential tightness is that for any $\ell < \infty$ there is a compact set C_ℓ in X such that

$$
Q_n[X - C_\ell] \leq e^{-\ell n} \quad \text{for all } n \tag{8.1}
$$

The construction of such a C_ℓ is usually awkward, involving lots of intersections and unions. Ideally our measures Q_n will be supported on a fixed compact set C. The next best thing is that they will be supported on some neighborhood C^{ϵ_n} of a compact set with $\epsilon_n \to 0$ as $n \to \infty$. Failing that, the least one should demand is that for each $\epsilon > 0$ and $\ell < \infty$ there exists a compact set $C_{\ell,\epsilon}$ such that

$$
\limsup_{n \to \infty} \frac{1}{n} \log Q_n\big[X - C_{\ell,\epsilon}^{\epsilon}\big] \leq -\ell \tag{8.2}
$$

It is a general fact that (8.2) implies what looks like the stronger definition (8.1). Here is a proof. Pick a sequence $\epsilon_k \downarrow 0$. For each ℓ and k we can clearly find a compact set C_{ℓ,ϵ_k} such that

$$
Q_n\big[X - C_{\ell,\epsilon_k}^{\epsilon_k}\big] \leq e^{-(\ell+1)n} \quad \text{for } n \geq n_0(\ell, k)
$$

We can assume without loss of generality that $n_0(\ell, k) \geq k$ so that

$$
Q_n\big[X - C_{\ell,\epsilon_k}^{\epsilon_k}\big] \leq e^{-k}\, e^{-\ell n} \quad \text{for } n \geq n_0(\ell, k)
$$

For each ℓ, k and $n = 1, 2, \cdots, n_0(\ell, k)$ by Prohorov's theorem we can find a compact set $D_{\ell,k}$ such that

$$Q_n\big[X - D_{\ell,k}\big] \le e^{-k}\, e^{-\ell n}$$

If we now define $B_{\ell,k} = D_{\ell,k} \cup C_{\ell,\epsilon_k}$, clearly $B_{\ell,k}$ is compact and

$$Q_n\big[X - B_{\ell,k}^{\epsilon_k}\big] \le e^{-k}\, e^{-\ell n} \quad \text{for all} \quad k \quad \text{and} \quad n\ .$$

If we now define K_ℓ to be the closure of $\cap_k B_{\ell,k}^{\epsilon_k}$, as the closure of a totally bounded set it is compact and

$$Q\big[X - K_\ell\big] \le \sum_{k \ge 1} e^{-k}\, e^{-\ell n} \le e^{-\ell n}$$

which is (8.1).

In our context our basic Polish space is the space of all probability measures on the Skorohod space $\Omega = D[[0,T]; \mathbf{T}^d]$ although the jumps once scaled are only of size $\frac{1}{N}$. The large deviations are going to happen only at measures supported on $C[[0,T]; \mathbf{T}^d]$. The set

$$A_{\delta,\epsilon} = \left\{ x(\cdot) : \sup_{\substack{|s-t| \le \delta \\ 0 \le s,t \le T}} |x(t) - x(s)| \le \epsilon \right\}$$

is contained $C_\ell^{2\epsilon}$ where C_ℓ is the space of functions that are Lipschitz with a Lipschitz constant $\ell = 2\frac{\epsilon}{\delta}$. Here by A^ϵ we mean the space of functions within a distance ϵ from A. The set $\{Q : Q\big[\Omega - A_\delta, \epsilon\big] \le \epsilon\}$ is within an ϵ neighborhood of a compact set in the Prohorov metric. So the superexponential estimate reduces to the following theorem.

Theorem 8.1. *Let P_N be the measure corresponding to the L interacting trajectories starting from an arbitrary initial configuration. Then for every $\epsilon > 0$,*

$$\limsup_{\delta \to 0} \limsup_{N \to \infty} \frac{1}{N^d} \log P_N\big[card\{j : \sup_{\substack{|s-t| \le \delta \\ 0 \le s,t \le T}} |x_j(s) - x_j(t)| \ge \epsilon\} \ge \epsilon N\big] = -\infty.$$

We will break the proof into several steps.

Although each $x_i(t)$ takes values on the torus $\mathbf{T}^d$ it is convenient to think of it as the projection onto the torus of an $\mathbf{R}^d$–valued process. Since the jumps of $x_i(\cdot)$ are all of size $\frac{C}{N}$, for sufficiently large N, each path $x_i(\cdot)$ can be lifted to a unique path $y_i(\cdot)$ with values in $\mathbf{R}^d$, at least if we pick the starting point $y_i(0)$ in some canonical fashion. To be meticulously correct we need to consider the two sets of processes together because the $y_i(\cdot)$ are governed by the configuration $\eta(\cdot)$ of the $x_i(\cdot)$ on $\mathbf{T}^d$ and $x_i(\cdot)$ shadow $y_i(\cdot)$. The $y_i(\cdot)$ are semimartingales and have the decomposition

$$y_i(t) = A_i(t) + M_i(t)$$

as the sum of a continuous function of bounded variation and a martingale. The function $A_i(t)$ is seen to be

$$A_i(t) = N \int_0^t \sum_z z\, p(z)(1 - \eta(x_i(t) + z, t)). \tag{8.3}$$

Each term in the summation corresponing to jumps of size $\frac{z}{N}$, to the sites $x_i(t)+z$ on the torus that are currently empty, that could potentially take place with rates $N^2p(z)$ in the speeded up time scale. The superexponential estimate that we are seeking for $\{x_i(\cdot)\}$ naturally breaks up into two parts, for $\{A_i(\cdot)\}$ and for $\{M_i(\cdot)\}$: 1. For every $\epsilon>0$,

$$\limsup_{\delta\to 0}\limsup_{N\to\infty}\frac{1}{N^d}\log P_N\big[card\{j:\sup_{\substack{|s-t|\le\delta\\ 0\le s,t\le T}}|A_j(s)-A_j(t)|\ge\epsilon\}\ge\epsilon N\big]=-\infty, \tag{8.4}$$

and 2. For every $\epsilon>0$,

$$\limsup_{\delta\to 0}\limsup_{N\to\infty}\frac{1}{N^d}\log P_N\big[card\{j:\sup_{\substack{|s-t|\le\delta\\ 0\le s,t\le T}}|M_j(s)-M_j(t)|\ge\epsilon\}\ge\epsilon N\big]=-\infty. \tag{8.5}$$

If we define for any trajectory $f(\cdot)$

$$S(f)=\sup_{0<\delta\le T}\sup_{\substack{|t-s|\le\delta\\ 0\le s,t\le T}}\frac{|f(t)-f(s)|}{\sqrt{\delta}\,\log\frac{1}{\delta}}$$

the following inequality was proved in [R: Lemma 2.4 inequality (2.11)]

$$\lim_{\ell\to\infty}\limsup_{N\to\infty}\frac{1}{N^d}\log P_N\Big[\sum_j S(A_j(\cdot))\ge\ell N^d\Big]=-\infty. \tag{8.6}$$

Equation (8.4) is an easy consequence of (8.6), because

$$|f(s)-f(t)|\le S(f)\sqrt{\delta}\,\log\frac{1}{\delta}$$

implies that

$$card\{j:\sup_{\substack{|s-t|\le\delta\\ 0\le s,t\le T}}|A_j(s)-A_j(t)|\ge\epsilon\}\le card\{j:S(A_j(\cdot))\ge\theta(\delta,\epsilon)\}$$

where

$$\theta(\delta,\epsilon)=\frac{\epsilon}{\sqrt{\delta}\,\log\frac{1}{\delta}}\to\infty\text{as}\quad\delta\to 0$$

The rest of this section will be devoted to the proof of (8.5). As a jump Markov process our process of L interacting particles comes with certain basic counting processes

$$F_N(i,z;t)=\quad\text{number of jumps of size }\frac{z}{N}\text{ of }x_i(\cdot)\text{ during }0\le s\le t$$

that have a jump rate

$$\lambda_N(i,z;t)=N^2p(z)(1-\eta(x_i(t)+z,t))$$

and our martingales $M_i(t)$ have the representation

$$\begin{aligned} M_i(t) &= \frac{1}{N}\sum_z z\left[F_N(i,z;t) - \int_0^t \lambda_N(i,z;s)ds\right] \\ &= \sum_z z\, M_N(i,z;t) \end{aligned}$$

as a finite sum of basic martingales that are mutually orthogonal for different z and i. It is important initially to estimate for each z,

$$P_N\big[H \geq \epsilon N\big]$$

where

$$H = card\{j : \sup_{0\leq s\leq \delta} |M_N(j,z;s) - M_N(j,z;0)| \geq \epsilon\}$$

If it were not for exclusion, $M_N(j,z;t)$ will be independent centered Poisson processes and the corresponding probability will be given by a simple binomial probability

$$\bar{P}[H \geq \epsilon N] = \sum_{r\geq \epsilon N} \binom{L}{r} [\beta_N(\delta,\epsilon)]^r [1-\beta_N(\delta,\epsilon)]^{L-r}$$

with

$$\beta_N(t,\epsilon) = \Pr\left[\sup_{0\leq s\leq t} |\frac{1}{N} Z_N(s) - Np| \geq \epsilon\right]$$

Here $Z_N(\cdot)$ is a standard Poisson process with intensity $N^2 p$ and $p = p(z)$. Our key step is the following theorem.

Theorem 8.2. *For every N, every initial configuation, and every ℓ, t and $\delta > 0$ we have*

$$P_N\left[H \geq \ell\right] \leq \bar{P}\left[H \geq \ell\right].$$

Proof. The idea is that for the actual exclusion process the rates are sometimes less than $N^2 p(z)$ i.e. 0 when blocked by another particle. We use ideas from control theory or dynamic programming to show that the Poisson case is the worst case. Without loss of generality, we assume that $N^2 p(z) = 1$. The proof of the theorem will be broken up into several lemmas.

Let $Z(\cdot)$ be the standard Poisson process with intensity 1. Let $\alpha(s) = Z(s) - s$ be the centered version. It has the infinitesimal generator

$$(\Omega f)(\alpha) = f(\alpha+1) - f(\alpha) - \frac{\partial f}{\partial \alpha}(\alpha)$$

Given $\alpha, t \in \mathbf{R}$ we can define a version that starts from α at time t by defining $\alpha(s) = \alpha + Z(s-t) - (s-t)$ for $s \geq t$. Given $m > 0$ we consider the stopping time τ_m which is the exit time of the process from the finite interval $(-m,m)$. For any $h(t)$ that is a smooth nonincreasing function on $\mathbf{R}$, with $0 \le h(\cdot) \le 1$ we define

$$q(\alpha,t) = \mathbf{E}_{\alpha,t}\big[h(\tau_m)\big].$$

Lemma 8.3. *The function q defined on $(-m,m)\times\mathbf{R}$ is continuously differentiable on $[-m,m)\times\mathbf{R}$ and satisfies*

$$\begin{cases} 0 \le q \le 1 \\ \frac{\partial q}{\partial t} \le 0 \\ \frac{\partial q}{\partial t}(\alpha,t) + (\Omega q)(\alpha,t) = 0 \end{cases} \tag{8.7}$$

with "boundary" conditions

$$\begin{cases} q(-m,t) = h(t) \quad \text{for} \quad t \in \mathbf{R} \\ q(\alpha,t) = h(t) \quad \text{for} \quad \alpha \in [m,m+1) \quad \text{and} \quad t \in \mathbf{R} \end{cases}.$$

Proof. By solving the Dirichlet problem

$$\frac{\partial q}{\partial t} + \Omega q = 0 \tag{8.8}$$

for the space time generator on $(-m,m)\times\mathbf{R}$ with boundary data $h(t)$ specified on the boundary $\{\{-m\}\cup[m,m+1)\}\times\mathbf{R}$, we see that q solves the equation (8.8). Since we can differentiate the equation with respect to t and obtain the same equation with boundary data $h'(t)$ it is clear by the maximum principle that we can get a bound on $\frac{\partial q}{\partial t}$ in terms of the bound on $h'(t)$. From the equation we have an apriori bound on $\frac{\partial q}{\partial t} - \frac{\partial q}{\partial \alpha}$. This yields bounds on both partial derivatives of q. Finally, because of the maximum principle, $h'(t) \le 0$ implies that $\frac{\partial q}{\partial t} \le 0$.

Given k numbers $x_1,\cdots,x_k$ with $0 \le x_j \le 1$, for each $1 \le \ell \le k$ we define

$$v_\ell^k(x_1,\cdots,x_k) = \sum_{\mathbf{A}} \prod_{j\in\mathbf{A}} x_j \prod_{j\notin\mathbf{A}} (1-x_j)$$

where the summation extends over all subsets $\mathbf{A} \subset \{1,\cdots,k\}$ of cardinality at least ℓ. If we interpret v_ℓ^k as the probability of getting at least ℓ heads in k independent coin tosses, x_j being the probability of head for the j-th toss, it is clear that v_ℓ^k is a multilinear function of $x_1,\cdots,x_k$ which is nondecreasing in each x_j. Let us note that if $x_r = 1$, then the summation extends only over those $\mathbf{A}$ that contain r and the function coincides on $x_r = 1$ with the function $v_{\ell-1}^{k-1}$ of the remaining $\ell - 1$ variables.

Lemma 8.4. *Suppose $\{\alpha_j(\cdot) : 1 \le j \le L\}$ are L independent centered Poisson processes as in the previous lemma, having the generator*

$$\tilde{\Omega}_L u = \sum_j \Omega_j u$$

where for each j, Ω_j is the generator Ω acting on the variable α_j that was defined in the previous lemma. If we define

$$u_\ell^L(\alpha_1,\cdots,\alpha_L,t) = v_\ell^L(q(\alpha_1,t),\cdots,q(\alpha_L,t))$$

then u_ℓ^L is sufficiently regular and satisfies:

$$\frac{\partial u_\ell^L}{\partial t} + \tilde{\Omega}_L u_\ell^L = 0 \quad \text{in} \quad (-m,m)^L \times \mathbf{R}$$

and

$$\Omega_j u_\ell^L \geq 0$$

for every j.

Proof. Follows immediately from lemma 8.3 and the fact that that v_ℓ^L is multilinear in its variables and nondecreasing with respect to each one.

At this point we can take limits and replace the smooth function $h(\cdot)$ by the function $h_T(t)$ which is 1 if $t \leq T$ and 0 otherwise. If we do that, in the limit,

$$q(\alpha,t) = \bar{P}_{\alpha,t}[\tau_m \leq T]$$

in $(-m\,m)$ and equals 1 outside. Recall that $\bar{P}_{\alpha,t}$ denotes the probability associated with the centered Poisson process $\alpha(s), s \geq t$ with $\alpha(t) = \alpha$. If we write $\bar{P}$ for the law of the process $(\alpha_1(\cdot),\ldots,\alpha_L(\cdot))$ generated by $\tilde{\Omega}$ with $\alpha_i(0) = \alpha_i$, then

$$u_\ell^L(\alpha_1,\ldots,\alpha_L,0) = \bar{P}(H \geq \ell)\ .$$

We see from (8.8) that the functions v_ℓ^L reduce to $v_{\ell-k}^{L-k}$ if k of the α_j's are outside $(-m,\, m)$ and the rest remain inside.

We now turn to the proof of Theorem 8.2. If we take L mutually orthogonal counting processes $\{c_j(t)\}$, i.e. no two of which will jump at the same time, having intensities $\{\lambda_j(t,\omega)\}$ satisfying $0 \leq \lambda_j(t,\omega) \leq 1$ for all j and if $\alpha_j(t) = c_j(t) - \int_0^t \lambda_j(s,\omega)\,ds$ are their centered versions, we can stop these processes one after another when they get out of $(-m,\, m)$ until ℓ of them escape before time T. Of course this may not happen. From Lemma 8.4 and the comments immediately following it, we see that $U = u_\ell^L(\alpha_1(t),\cdots,\alpha_L(t),t)$ remains a supermartingale until ℓ processes have escaped or until time T whichever is sooner. In particular, the probability that at least ℓ particles exit from $(-m,\, m)$ before time T for any L arbitrary centered pairwise orthogonal counting processes with intensities bounded by one 1, is dominated by the corresponding probability for L independent centered canonical Poisson processes with intensity 1. This concludes the proof of Theorem 8.2.

We now take up the proof of our main result, Theorem 8.1. For each $\epsilon > 0$ and $\delta > 0$ we want to count the number of particles that "misbehave": i.e.

$$\sup_{\substack{|s-t|\leq\delta \\ 0\leq s,t\leq T}} |M_i(s) - M_i(t)| \geq \epsilon$$

Let us divide the interval $[0,T]$ into k intervals $\{I_1,\cdots,I_k\}$ of length $\delta=\frac{T}{k}$ with $I_r=[\frac{(r-1)T}{k},\frac{rT}{k}]$. If for any trajectory $M(\cdot)$ and any $\epsilon>0$

$$\sup_{t\in I_r}|M(t)-M(\frac{(r-1)T}{k})|\le\frac{\epsilon}{3} \tag{8.9}$$

for every $1\le r\le k$, then

$$\sup_{\substack{|s-t|\le\delta(k)\\0\le s,t\le T}}|M(s)-M(t)|\le\epsilon$$

where $\delta(k)=\frac{T}{k}$. Let K_r be the number of particles that violate (8.9). The total number of particles K that misbehave can clearly be estimated by

$$K\le\sum_r K_r$$

If we denote by $\pi(N,k,\frac{\epsilon}{3})$ the probability that a normalized centered Poisson process

$$\alpha_N(t)=\frac{1}{N}[Z(N^2t)-N^2]$$

violates (8.9) in an interval I_r of length $\frac{T}{k}$, by Theorem 8.2, the processes M_i's are uniformy conditionally stochastically dominated by L independent binomials with probability $\pi(N,k,\frac{\epsilon}{3})$. Let γ be the cardinality of the set of possible jump sizes i.e the cardinality of the set $\{z:p(z)>0\}$ and $\Gamma=\sum_{z:p(z)>0}|z|$. ¿From the representation

$$M_i(t)=\sum_z zM_N(i,z;t)\ ,$$

we can bound the oscillation of $M_i(\cdot)$ by

$$|M_i(t)-M_i(s)|\le\Gamma\sup_z|M_N(i,z;t)-M_N(i,z;s)|.$$

With all of this we have the following inequality:

$$P_N\big[K\ge\ell\big]\le F(Lk\gamma,\pi(N,\frac{T}{k},\frac{\epsilon}{3\Gamma}),\ell)$$

where $F(n,\pi,\ell)$ is the probability of at least ℓ successes in n binomial trials with probability π of success in a single trial. We conclude now that with $\ell=N^d\epsilon$

$$\limsup_{N\to\infty}\frac{1}{N^d}\log P_N\big[K\ge N^d\epsilon\big]\le-k\gamma\bar\rho\sup_{p\ge\theta(k)}[p\log\frac{p}{q}+(1-p)\log\frac{1-p}{1-q}] \tag{8.10}$$

Here

$$\bar\rho=\lim_{N\to\infty}\frac{L}{N^d}$$

is the overall density of particles,

$$q = q(k) = \lim_{N\to\infty} \pi(N, \frac{T}{k}, \frac{\epsilon}{3\Gamma}) = \Pr\big[\sup_{0\le t\le \frac{T}{k}} |\beta(t)-\beta(0)| \ge \frac{\epsilon}{3\Gamma}\big]$$

β being the standard Brownian motion, and

$$\theta(k) = \lim_{N\to\infty} \frac{\ell}{Lk\gamma} = \frac{\epsilon}{k\bar\rho\gamma}$$

In this estimate $\bar\rho$, γ and ϵ are fixed and letting $\delta \to 0$ amounts to letting $k \to \infty$. Then $\log q(k) \simeq -c_1(\epsilon)k$ and $\theta(k) \simeq \frac{c_2(\epsilon)}{k}$. The right hand side of () is then clearly behaving like

$$-\gamma\bar\rho k\theta(k)\log\frac{\theta(k)}{q(k)} \simeq -c_3(\epsilon)k$$

and can be made large with k.

9. Proof of the main Theorem.

We are now ready to prove our main theorem on the large deviations of the empirical process. Let us recall the notation. $D[[0,T];\mathbf{T}^d]$ is the space of trajectories on $[0,T]$ with values in $\mathbf{T}^d$ with discontinuities, if any, of only the first kind. The space $\mathcal{M}(D[[0,T];\mathbf{T}^d])$ is the space of nonnegative measures on $D[[0,T];\mathbf{T}^d]$. Given the L rescaled trajectories $\{y_1(\cdot),\ldots,y_L(\cdot)\}$ of our interacting particles, we define the empirical process as

$$R_{N,\omega} = \frac{1}{N^d}\sum_{i=1}^{L}\delta_{y_i(\cdot)}$$

with values in the space $\mathcal{M}(D[[0,T];\mathbf{T}^d])$. We will denote by P_N^ξ the measure induced on this space from the underlying particle system with an arbitrary initial configuration $\{y_1,\ldots,y_L\}$ of L particles in the lattice Z_N^d imbedded in $\mathbf{T}^d$. If the initial configuration is random with a distribution μ_N it is assumed to satisfy a large deviation principle on the space of macroscopic density profiles $q(\cdot)$ with a rate function $I_{init}(q(\cdot))$. We want to show that a large deviation principle holds for the distributions

$$P_N = \int P_N^\xi\,\mu_N(d\xi)$$

of $R_{N,\omega}$, and that the rate function for this large deviation principle is the one given in section 7. For any measure $Q \in \mathcal{M}(D[[0,T];\mathbf{T}^d])$, $\mathcal{I}(Q)$ is infinite unless Q is supported on the set of continuous trajectories $(C[[0,T];\mathbf{T}^d])$. Moreover if we denote by $q(t,\cdot)$ the marginal density of Q at time t, we must have

(9.1) $$\int_0^T\int_{\mathbf{T}^d}\frac{(\nabla q\cdot D\nabla q)}{q(1-q)}\,dtd\theta < \infty$$

For any $q = q(\cdot,\cdot)$, satisfying (9.1), $\mathcal{B}_q$ is defined as $b(\cdot,\cdot)$ that satisfy

$$\mathcal{E}(b) = \frac{1}{2}\int_{\mathbf{T}^d}\int_0^T (Db\cdot b)\, q(1-q)\, d\theta dt < \infty \tag{9.2}$$

and solve in weak sense the equation

$$\frac{\partial q}{\partial t} - \frac{1}{2}\nabla\cdot D\nabla q + \nabla\cdot Dbq(1-q) = 0 \tag{9.3}$$

We know that for $b \in \mathcal{B}_q$, we have a unique time inhomogeneous diffusion process with marginals equal to q and generator

$$\frac{1}{2}\nabla\cdot S(q(t,\theta))\nabla + \frac{1}{2}(S(q(t,\theta)) - D)\frac{\nabla q(t,\theta)}{q(t,\theta)}\cdot\nabla + Db(t,\theta)(1-q(t,\theta))\cdot\nabla \tag{9.4}$$

We denote this diffusion processe by P^b. As a Markov Process, we can start this diffusion with any initial distribution ρ that is dominated by $q(0\,,\cdot)$ and we will denote these processes by P^b_ρ. Let us note that these are not necessarily probability measures and will instead have their total masses equal to that of the initial distribution.

The rate function $\mathcal{I}(Q) = \infty$ unless $Q << P^b$ with finite relative entropy $H(Q\,,P^b)$ for some $b \in \mathcal{B}_Q$. In that case, according to Theorem 7.2, there is a unique $b \in \mathcal{B}_Q$, which we denote by b_Q, that will satisfy

$$E^Q[\int_0^T \phi(t,x(t))\cdot dx(t)\,] = E^{P^{b_Q}}[\int_0^T \phi(t,x(t))\cdot dx(t)\,] \tag{9.5}$$

for every smooth function $\phi : [0,T]\times\mathbf{T}^d \to R^d$. Moreover $H(Q, P^{b_Q}) < \infty$ and

$$\mathcal{I}(Q) = I_{[0,T]}(Q) = I_{init}(q(0,\cdot)) + \mathcal{E}[b_Q] + H(Q\,,P^{b_Q}). \tag{9.6}$$

If the initial configuration of the particle system were nonrandom with a macroscopic profile of $\rho_0(\cdot)$, then the large deviation rate function of the empirical process would be given by

$$\mathcal{I}(Q) = \mathcal{I}_{[0,T]}(Q\,;\rho_0) = \mathcal{E}[b_Q] + H(Q, P^{b_Q}) \tag{9.7}$$

provided Q is such that $q(0\,,\cdot) = \rho_0(\cdot)$, and $\mathcal{I}(Q) = +\infty$ otherwise. The transition from this deterministic version to the more general version with random initial data is routine. The proof of our main theorem will depend on proving large deviations for certain finite dimensional projections.

Let us take a partition $\mathcal{P}$ of the torus $\mathbf{T}^d$ into r sets $B_1, B_2, \ldots, B_r$ having boundaries with zero Lebesgue measure. Let $\mathcal{A}^k$ denote the set of indices $\{a_1,\cdots,a_k\}$ where each

$a_i \in \{1, \ldots, r\}$. For a choice $\mathbf{a} \in \mathcal{A}^k$, and $k+1$ time points $0 = t_1 < t_2 < \cdots < t_{k+1} \le T$, we let

$$E_{\mathbf{a}} = \big\{y(\cdot) : y(t_i) \in B_{a_i} \quad \text{for} \quad i = 1, \cdots, k\big\}$$

and define $n = r^k$ empirical measures $\nu^k_{\mathbf{a}} \in \mathcal{M}(\mathbf{T}^d)$ on $\mathbf{T}^d$ that correspond to each of the r^k indices $\mathbf{a} \in \mathcal{A}^k$,

$$\nu^k_{\mathbf{a}} = \frac{1}{N^d} \sum_{i : y_i(\cdot) \in E_{\mathbf{a}}} \delta_{y_i(t_{k+1})}$$

We have simplified the past history by partitioning the space of trajectories on $[0, t_k]$ into r^k groups and we have painted the particles in each group a different color. We are then looking at the empirical distributions at time $t = t_{k+1}$ of particles with r^k different colors. As a first step, we fix r, the sets $B_1, \ldots, B_r$, k, and the time points $0 = t_1 < t_2 < \cdots < t_k < t_{k+1} \le T$ and we are interested in establishing a large deviation theorem for the collection $\{\nu^k_{\mathbf{a}}, \, \mathbf{a} \in \mathcal{A}^k\}$ of r^k measures on $\mathbf{T}^d$.

Theorem 9.1. *The large deviation princilple is valid for $\{\nu^k_{\mathbf{a}}\}$ on the space $[\mathcal{M}(\mathbf{T}^d)]^{r^k}$ with a rate function*

$$I^{(k)}(\{p_{\mathbf{a}}\}) = \inf_{Q \in \mathcal{M}_k(\{p_{\mathbf{a}}\})} \mathcal{I}(Q) \tag{9.8}$$

where $\mathcal{M}_k(\{p_{\mathbf{a}}\})$ is the set of measures Q on $C[[0,T]; \mathbf{T}^d]$ with the property that for all $\mathbf{a} \in \mathcal{A}^k$, and for all Borel sets $A \subset \mathbf{T}^d$,

$$Q[\cap_{1 \le j \le k}\{y(t_j) \in B_{a_j}\} \cap \{y(t_{k+1}) \in A\}] = Q[E_{\mathbf{a}} \cap \{y(t_{k+1}) \in A\}] = \int_A p_{\mathbf{a}}(\theta) d\theta.$$

Remark: Given any Q, $\mathbf{a} \in \mathcal{A}^k$ and $F \in \mathcal{F}^k = \sigma\{y(s) : t_k \le s \le T\}$ we define

$$Q_{\mathbf{a}}[F] = Q[\cap_{1 \le j \le k}\{y(t_j) \in B_{a_j}\} \cap F]$$

then $Q \in \mathcal{M}_k(\{p_{\mathbf{a}}\})$ if and only if for each $\mathbf{a} \in \mathcal{A}^k$, the marginal at time t_{k+1} of $Q_{\mathbf{a}}$ has density $p_{\mathbf{a}}$.

Remark: Although we are mainly interested in the large deviation rates for the quantities $\{\nu^k_{\mathbf{a}}(B_a)\}$, in order to make the induction in k work it is important to carry along complete information about the empirical distributions $\{\nu^k_{\mathbf{a}}\}$ of particles of different colors at time t_{k+1}.

Proof: The proof of the theorem will proceed by induction on k. For $k = 1$, there is nothing to prove since it all happens at time 0 and I_{init} controls everything. For $k = 2$ we can color the particles with r colors depending on which B_a they belong to in the partition at time $t_1 = 0$. Since B_a are continuity sets for the Lebesgue measure, the large deviation principle can be deduced from Theorem 4.13, with a rate function that can be obtained by the contraction principle of Theorem 7.5.

$$\begin{aligned} I^{(1)}(\{p_a\}) &= \inf_{\tilde\rho(\cdot\,,\cdot) \in \mathbf{C}} \left[I_{init}(\tilde\rho(0\,,\cdot)) + I_{dyn}(\tilde\rho(\cdot\,,\cdot))\right] \\ &= \inf_{Q \in \mathcal{M}_1(\{p_a\})} \left[I_{init}((q(0,\cdot)) + \mathcal{I}_{[t_1,t_2]}(Q)\right] \end{aligned} \tag{9.9}$$

where $\mathbf{C}$ is the set of densities

$$\mathbf{C} = \left\{ \tilde{\rho}(\cdot\,,\cdot) : \rho_i(t_1\,,\theta) = 0 \text{ if } \theta \notin B_i\,,\ \text{and } \rho_i(t_2\,,\cdot) = p_i(\cdot) \text{ for } 1 \le i \le r \right\}$$

If the initial configuration is deterministic and corresponds to the densities $\{\rho_a\}$ for the different colors $a \in \mathcal{A}$, with the support of ρ_a contained B_a for each $a \in \mathcal{A}$, then the rate function becomes

$$\begin{aligned} I^{(1)}(\{p_a\}\,;\{\rho_a\}) &= \inf_{\tilde{\rho}(\cdot\,,\cdot)\in\mathbf{C}} I_{dyn}(\tilde{\rho}(\cdot\,,\cdot)) \\ &= \inf_{Q\in\mathcal{M}_1(\{p_a\}\,;\{\rho_a\})} \mathcal{I}_{[t_1,t_2]}(Q) \end{aligned} \tag{9.9}$$

with

$$\mathcal{M}_1(\{p_a\}\,;\{\rho_a\}) = \mathcal{M}_1(\{p_a\}) \bigcap \left\{ Q : q(0\,,\cdot) = \sum_a \rho_a(\cdot) \right\}$$

Let us assume that the result is valid for $k+1$ time points $0 = t_0 < t_1 < \cdots < t_k < t_{k+1} \le T$ and try to establish the induction. We add an extra time point t_{k+2} at the end so that we now have $0 = t_1 < t_2 < \cdots < t_{k+1} < t_{k+2} \le T$.

The basic ingredient in the induction step is the following elementary property concerning relative entropy. Suppose we consider the three σ-fields $\mathcal{F}_t = \sigma\{x(s) : 0 \le s \le t\}$, $\mathcal{F}^t = \sigma\{x(s) : t \le s \le T\}$ and $\mathcal{F}_t^t = \sigma\{x(t)\}$, in $\Omega = C[[0\,,T]; \mathbf{T}^d]$. Let Q be a measure on $(\Omega\,,\mathcal{F}_T)$. Let $\{P_{t,x}\}$ be some inhomogeneous diffusion in the time interval $[0\,,T]$. The measure Q is not necessarily a probability measure, but has a marginal $q(0\,,dx)$ at time 0. The diffusion initialized with the measure $q(0\,,\cdot)$ at $t = 0$ will be denoted by $P_{0\,,q}$. Let us suppose that $H(Q\,,P_{0\,,q}) < \infty$. The regular conditional probabilities of Q given the σ-fields $\mathcal{F}_t$ and $\mathcal{F}_t^t$ are denoted respectively by $Q_\omega^{[0\,,t]}$ and Q_ω^t. These are of course probability measures. Let us replace the measure Q by $\bar{Q}$ defined by, $\bar{Q} = Q$ on $\mathcal{F}_t$ and the regular conditional probability $\bar{Q}_\omega^{[0\,,t]}$ of $\bar{Q}$ given $\mathcal{F}_t$ is taken to be Q_ω^t on $\mathcal{F}^t$. In other words we force the Markov property at time t. Then $\bar{Q} = Q$ on both the σ-fields $\mathcal{F}_t$ and $\mathcal{F}^t$.

$$H(\bar{Q}\,,P_{0\,,q}) \le H(Q\,,P_{0\,,q})$$

We now proceed with the induction argument. Suppose Q is given as a measure on $(\Omega\,,\mathcal{F}_T)$. Let $E_{\mathbf{a},a}$ be the partition of Ω indexed by $(\mathbf{a},a) \in \mathcal{A}^{k+1}$ according to the colors at times $0 = t_1 <, \cdots, < t_{k+1} \le T$ and $t_{k+2} \in (t_{k+1}\,,T]$ is the next inspection time. We denote by

$$Q_{\mathbf{a},a}(\cdot) = Q(E_{\mathbf{a},a} \cap \cdot)$$

the restriction of Q to the various subsets of the partition. If we replace each $Q_{\mathbf{a},a}$ by the corresponding $\bar{Q}_{\mathbf{a},a}$ at time t_{k+1} and resum to get

$$\bar{Q} = \sum_{\mathbf{a},a} \bar{Q}_{\mathbf{a},a}$$

we can make the following observations.

1) For each $(\mathbf{a}, a) \in \mathcal{A}^{k+1}$, $Q_{\mathbf{a},a}$ and $\bar{Q}_{\mathbf{a},a}$ agree on both $\mathcal{F}_{t_{k+1}}$ and $\mathcal{F}^{t_{k+1}}$. In particular $Q \in \mathcal{M}_{k+1}(\{p_{\mathbf{a},a}\})$ if and only if $\bar{Q} \in \mathcal{M}_{k+1}(\{p_{\mathbf{a},a}\})$.

2) The two measures Q and $\bar{Q}$ have the same marginals for every t.

3) The quantities b_Q and $b_{\bar{Q}}$ are the same. They are detrmined by the expected values of stochastic integrals and so do not depend on the fact that the conditionals across t_{k+1} have been changed. In particular $P^{b_Q} = P^{b_{\bar{Q}}}$.

4) Therefore in the definition of $\mathcal{I}(Q)$ the two terms $\mathcal{E}(b_Q)$ and $\mathcal{E}(b_{\bar{Q}})$ are equal.

5) The entropy terms satisfy

$$H(\bar{Q}\,, P^{b_Q}) \leq H(Q\,, P^{b_{\bar{Q}}})$$

6) In the infimum

$$\inf_{Q \in \mathcal{M}_{k+1}(\{p_{\mathbf{a},a}\})} \mathcal{I}(Q)$$

we can replace Q by $\bar{Q}$.

Let us now compute $\mathcal{I}(\bar{Q})$. The term $\mathcal{E}(b)$ is additive over disjoint time intervals. By the Markov property that was forced at time t_{k+1} on $\bar{Q}$,

$$\begin{aligned} H_{[0\,,t_{k+2}]}(\bar{Q}\,, P^{b_Q}) &= H_{[0\,,t_{k+1}]}(Q\,, P^{b_Q}) + \sum_{\mathbf{a},a} H_{[t_{k+1}\,,t_{k+2}]}(\hat{Q}_{\mathbf{a},a}\,, P^{b_Q}_{\mathbf{a},a}) \\ &= H_{[0\,,t_{k+1}]}(Q\,, P^{b_Q}) + \sum_{\mathbf{a}} H_{[t_{k+1}\,,t_{k+2}]}(\hat{Q}_{\mathbf{a}}\,, P^{b_Q}_{\mathbf{a}}) \end{aligned}$$

where

$$\hat{Q}_{\mathbf{a}} = \sum_a \hat{Q}_{\mathbf{a},a} = \int_{\mathbf{T}^d} Q^{t_{k+1}}_{\mathbf{a},x} p_{\mathbf{a}}(x) dx$$

are just the restrictions of $Q_{\mathbf{a}}$ to $\mathcal{F}^{t_{k+1}}$ represented as integrals of the conditionals given the initial position at time t_{k+1} with respect to the marginals at time t_{k+1} and

$$P^{b_Q}_{\mathbf{a}} = \int_{\mathbf{T}^d} P^{b_Q}_{t_{k+1},x} p_{\mathbf{a}}(x) dx$$

is the diffusion with initial density $p_{\mathbf{a}}(\cdot)$ at time t_{k+1}. We can now calculate

$$\inf_{Q \in \mathcal{M}_{k+1}(\{p_{\mathbf{a},a}\})} \mathcal{I}(Q) = \inf_{Q \in \mathcal{M}_{k+1}(\{p_{\mathbf{a},a}\})} \mathcal{I}_{[0\,,t_{k+2}]}(Q)$$

$$\geq \inf_{\substack{\{p_{\mathbf{a}}\} \in \mathcal{D} \\ Q \in \mathcal{M}_k(\{p_{\mathbf{a}}\}) \\ \{\bar{Q}_{\mathbf{a}}\} \in \mathcal{M}}} \left\{ \mathcal{I}_{[0\,,t_{k+1}]}(Q) + \sum_{\mathbf{a}} H_{[t_{k+1},t_{k+2}]}(\bar{Q}_{\mathbf{a}}\,, P^{b_{\bar{Q}}}_{\mathbf{a}}) + \mathcal{E}_{[t_{k+1},t_{k+2}]}(b_{\bar{Q}}) \right\}$$

where

$$\mathcal{D} = \left\{ \{p_{\mathbf{a}}\} : \int_{B_a} p_{\mathbf{a}}(x)\, dx = \bar{Q}_{\mathbf{a}}(B_a) \ \text{ for all } \ \mathbf{a} \in \mathcal{A}^k \right\}$$

and

$$\mathcal{M} = \left\{ \{\bar{Q}_{\mathbf{a}}\} : \bar{Q}_{\mathbf{a}} \left[\{y(t_{k+1}) \in B_a\} \cap \{y(t_{k+2}) \in A\} \right] = \int_A p_{\mathbf{a},a}(x)\, dx \right\}$$

for all $\mathbf{a}\,, a$ and measurable $A \subset \mathbf{T}^d$. The inequality is actually an equality because, starting from Q on $\mathcal{F}_{t_{k+1}}$ and $\bar{Q}_{\mathbf{a}}$ that are compatible, i.e. occur in the infimum, we can patch together a measure Q on $\mathcal{F}_{t_{k+2}}$ that figures in the original infimum.

Now it is easy to see that the induction works. If at time t_{k+1} we had r^k colors and they had density profiles $\{p_{\mathbf{a}}\}$ and we refined the colors at that time to have r^{k+1} colors, we will have the profiles $\{\rho_{\mathbf{a},a}\}$ where $\rho_{\mathbf{a},a}(x) = p_{\mathbf{a}}(x)\chi_{B_a}(x)$. The rate function for them to have profiles $p_{\mathbf{a},a}$ at time t_{k+2} is given by

$$\inf_{\{\bar{Q}_{\mathbf{a}}\} \in \mathcal{M}} \left\{ \sum_{\mathbf{a}} H_{[t_{k+1}, t_{k+2}]}(\bar{Q}_{\mathbf{a}}\,, P_{\mathbf{a}}^{b_{\bar{Q}}}) + \mathcal{E}_{[t_{k+1}, t_{k+2}]}(b_{\bar{Q}}) \right\}$$

according to the induction step with $k = 2$. To remove the condition of fixing the profiles at time t_{k+1}, we have to add the large deviation rate for $\{p_{\mathbf{a}}\}$, which is equal to

$$\inf_{Q \in \mathcal{M}_k(\{p_{\mathbf{a}}\})} \mathcal{I}_{[0\,, t_{k+1}]}(Q)$$

by the induction hypothesis and we are done.

We can now prove our main theorem that the large deviation principle holds for $R_{N,\omega}$ with the rate function $I(Q)$. In Theorem 9.1 keeping track of the distribution at the last time t_k was needed only for for the purpose of carrying out the induction.

By contraction we have the following large deviation principle.

Corollary 9.2. *The large deviation principle is valid for*

$$\{\pi_{\mathbf{a}}\} = \frac{1}{N^d} \{ \text{ number of particles } \ i : y(t_i) \in B_{a_i} \ \text{ for } \ 1 \le i \le k \}$$

on the space $\mathbf{R}^{r^k}$ *with a rate function*

$$I^{(k)}[\pi] = \inf_{Q \in \mathcal{M}_k(\{\pi_{\mathbf{a}}\})} I(Q)$$

where $\mathcal{M}_k(\{\pi_{\mathbf{a}}\})$ *is the set of measures* Q *on* $C[[0,T]; \mathbf{T}^d]$ *with the property*

$$Q[\cap_{1 \le j \le k} \{y(t_j) \in B_{a_j}\}] = Q[E_{\mathbf{a}}] = \pi_{\mathbf{a}}$$

for all $\mathbf{a}$.

The proof of the main theorem is now routine by refining the partition and making the sampling times dense in $[0, T]$. We have superexponential tightness and compactness of the level sets of the rate function $I(Q)$. Because the rate function $I(Q)$ is finite only for measures supported on $C[[0,T];\mathbf{T}^d]$ and the sets in our partition are continuity sets for the marginals of Q with $I(Q) < \infty$ there are no technical difficulties in carrying out the limiting procedure and establishing our main theorem.

Acknowledgement. J. Q. and F. R. would like to thank S. R. S. Varadhan, H. -T. Yau and the Courant Institute for their warm hospitality while a part of this work was done. F. R. was partially supported by a Sloan Foundation Fellowship.

References

[KOV] C. Kipnis, S. Olla, and S. R. S. Varadhan, "Hydrodynamics and large deviations for simple exclusion processes", Comm. Pure Appl. Math. **42**, 115–137 (1989).

[KV] C. Kipnis, S.R.S. Varadhan, "Central limit theorem for additive functionals of reversible Markov processes and applications to simple exclusion", Comm. Math. Phys. 106, 1-19.

[Q1] J. Quastel, "Diffusion of color in the simple exclusion process", Comm. Pure Appl. Math. **45**, 623–679 (1992).

[Q2] J. Quastel, " Large deviations from a hydrodynamic scaling limit for a nongradient system", Ann. Prob. **23** No. 2, (1995).

[QV] J. Quastel, and S. R. S. Varadhan, "Diffusion semigroup and diffusion processes corresponding to degenerate divergence form operators", to appear in Comm. Pure Appl. Math.

[R] F. Rezakhanlou, "Propagation of chaos for symmetric simple exclusion", Comm. Pure Appl. Math. **117**, 943–957 (1994).

[V1] S. R. S. Varadhan, "Nonlinear diffusion limit for a system with nearest neighbor interaction II", in Proc. Taniguchi Symp., Kyoto (1990).

[V2] S. R. S. Varadhan, "Regularity of self–diffusion coefficient",