The simplest example is a system of noninteracting particles undergoing independent motions. For instance we could have on $\mathbf{T}^3$, $L \simeq \bar{\rho} N^3$ particles all behaving like independent Browninan Particles. If the initial configuration of the L particles is such that the empirical distribution

$$\nu_0(dx) = \frac{1}{N^3}\sum_i \delta_{x_i}$$

has a deterministic limit $\rho_0(x)dx$, then the empirical distribution

$$\nu_t(dx) = \frac{1}{N^3}\sum_i \delta_{x_i(t)}$$

of the configuration at time t, has a deterministic limit $\rho(t\,,x)dx$ as $N \to \infty$ and $\rho(t\,,x)$ can be obtained from $\rho_0(x)$ by solving the heat equation

$$\frac{\partial \rho}{\partial t} = \frac{1}{2}\Delta \rho$$

with the initial condition $\rho(0\,,x) = \rho_0(x)$. The proof is an elementary law of large numbers argument involving a calculation of two moments. Let $f(x)$ be a continuous function on $\mathbf{T}$ and let us calculate for

$$U = \frac{1}{N^3}\sum_i f(x_i(t))$$

the first two moments given the initial configuration $(x_1, \cdots, x_L)$

$$E(U) = \frac{1}{N^3}\sum \int_{\mathbf{T}^3} f(y)p(t\,,x_i\,,y)dy$$

and an elementary calculation reveals that the conditional expectation converges to the following constant.

$$\int_{\mathbf{T}^3}\int_{\mathbf{T}^3} f(y)p(t\,,x\,,y)\rho_0(x)dydx = \int_{\mathbf{T}^3} f(y)\rho(t\,,y)dy$$

The independence clearly provides a uniform upper bound of order N^{-3} for the conditional variance that clearly goes to 0. Of course on $\mathbf{T}^3$ we could have had a process obtained by rescaling a random walk on a large torus of size N. Then the hydrodynamic scaling limit would be a consequence of central limit theorem for the scaling limit of a single particle and the law of large numbers resulting from the averaging over a large number of independently moving particles. The situation could be different if the particles interacted with each other.

The next class of examples are called simple exclusion processes. They make sense on any finite or countable set X and for us X will be either the integer lattice $\mathbf{Z}^d$ in d-dimensions or $\mathbf{Z}^d_N$ obtained from it as a quotient by considering each coordinate modulo N. At any given time a subset of these lattice sites will be occupied by partcles, with atmost one particle at each site. In other words some sites are empty while others are occupied with one particle. The particles move randomly. Each particle waits for an exponential random time and then tries to jump from the current site x to a new site y. The new site y is picked randomly according to a probability distribution $\pi(x\,,y)$. In particular $\sum_y \pi(x\,,y) = 1$ for every x. Of course a jump to y is not always possible. If the site is empty the jump is possible and is carried out. If the site already has a particle the jump cannot be carried out and the particle forgets about it and waits for another chance, i.e. waits for a new exponential waiting time.

If we normalize so that all waiting times have mean 1, the generator of the process can be written down as

$$(\mathcal{A}\, f)(\eta) = \sum_{x,y} \eta(x)(1-\eta(y))\pi(x\,,y)[f(\eta^{x,y}\,) - f(\eta)]$$

where η represents the configuration with $\eta(x) = 1$ if there is a particle at x and $\eta(x) = 0$ otherwise. For each configuration η and a pair of sites x, y the new configuration $\eta^{x,y}$ is defined by

$$\eta^{x,y}(z) = \begin{cases} \eta(y) & if\ Z = x \\ \eta(x)\ if\ z = y \\ \eta(z)\ if\ z \neq x\,, y \end{cases}$$

We will be concerned mainly with the situation where the set X is $\mathbf{Z}^d$ or $\mathbf{Z}^d_N$, viewed naturally as an Abelian group with $\pi(x\,,y)$ being translation invariant and given by $\pi(x\,,y) = p(y-x)$ for some peobability distribution p. It is convenient to assume that p has finite support. There are various possibilities.

$$p \text{ is symmetric i.e. } p(z) = p(-z)$$

or more generally

$$p \text{ has mean zero i.e. } \sum_z zp(z) = 0$$

and finally

$$\sum_z z\,p(z) = m \neq 0$$

We shall first concentrate on the symmetric case. Let us look at the function

$$V_J(\eta) = \sum J(x)\eta(x)$$

and compute

$$\begin{aligned}(\mathcal{A}V_J)(\eta) &= \sum_{x,y} \eta(x)(1-\eta(y))p(y-x)(J(y)-J(x)) \\ &= \sum_{x,y} \eta(x)p(y-x)(J(y)-J(x)) \\ &= \sum_{x,y} \eta(x)[(\mathbf{P}-I)J](x) \\ &= V_{(\mathbf{P}-I)J}(\eta)\end{aligned}$$

The space of linear functionals is left invariant by the generator. It is not difficult to see that

$$E_\eta\big[V_J(\eta(t))\big] = V_{J(t)}(\eta)$$

where

$$J(t) = \exp[t(\mathbf{P} - I)]J$$

is the solution of

$$\frac{d}{dt}J(t\,,x) = (\mathbf{P} - I)J(t\,,x)$$

It is almost as if the interaction has no effect and in fact in the calculation of expectations of 'one particle' functions it clearly does not. Let us start with a configuration on $\mathbf{Z}_N^d$ and scale space by N and time by N^2. The generator becomes $N^2\mathcal{A}$ and the particles can be visualized as moving in a lattice imbedded in the unit torus $\mathbf{T}^d$, with spacing $\frac{1}{N}$, and becoming dense as $N \to \infty$.

Let J be a smooth function on $\mathbf{T}^d$. We consider the functional

$$\xi(t) = \frac{1}{N^d}\sum_x J(\frac{x}{N})\eta_t(x)$$

and we can write

$$\xi(t) - \xi(0) = \int_0^t V_N(\eta(s))ds + M_N(t)$$

where

$$V_N(\eta) = (N^2\mathcal{A}V_J)(\eta) = V_{J_N}(\eta)$$

with

$$\begin{aligned}(J_N)(\theta) &= N^2\sum\big[J(\theta + \frac{z}{N}) - J(\theta)\big]p(z)\\ &\simeq \frac{1}{2}(\Delta_C J)(\theta)\end{aligned}$$

for $\theta \in \mathbf{T}^d$. Here Δ_C refers to the Laplacian

$$\sum_{i,j} C_{i,j}\frac{\partial^2}{\partial x_i \partial x_j}$$

with the covariance matrix C given by

$$C_{i,j} = \sum_z z_i z_j p(z)$$

$M_N(t)$ is a martingale and a very elementary calculation yields

$$E\Big\{[M_N(t)]^2\Big\} \le C\,t\,N^{-d}$$

essentially completing the proof in this case. Technically the empirical distribution $\nu_N(t)$ is viewed as a measure on $\mathbf{T}^d$ and $\nu_N(\cdot)$ is viewed as a stochastic process with values in the space $\mathcal{M}(\mathbf{T}^d)$ of nonnegative measures on $\mathbf{T}^d$. In the limit it lives on the set of weak solutions of the heat equation

$$\frac{\partial \rho}{\partial t} = \frac{1}{2}\Delta_C \, \rho$$

and the uniqueness of such weak soultions for given initial density establishes the validity of the hydrodynamic limit.

Let us make the problem slightly more complicated by adding a small bias. Let $q(z)$ be an odd function with $q(-z) = -q(z)$ and we will modify the problem by making p depend on N in the form

$$p_N(z) = p(z) + \frac{1}{N} q(z)$$

Assuming that q is nonzero only when p is so, p_N will be an admissible transition probability for large enough N. A calculation yields that in the slightly modified model referred to as weakly asymmetric simple exclusion model V_N is given by

$$V_N(\eta) \simeq V_{J_N}(\eta) + \frac{1}{N^d}\sum_x \eta(x)(1-\eta_x) < m \, , \nabla J(x) >$$

with

$$m = \sum_z z \, q(z)$$

If one thinks of $\rho(t\,,\theta)$ as the density of particles at the (macroscopic) time t and space θ the first term clearly wants to have the limit

$$\int_{\mathbf{T}^d} \frac{1}{2}(\Delta_C J)(\theta)\rho(t\,,\theta) d\theta$$

It is not so clear what to do with the second term. The 'invariant' measures in this model are the Bernoulli measures with various densities ρ and the 'averaged' version of the second term should be

$$\int_{\mathbf{T}^d} < m \, , (\nabla J)(\theta) > \rho(t\,,\theta)(1-\rho(t\,,\theta) d\theta$$

Replacing the linear heat equation by the nonlinear equation

$$\frac{\partial \rho}{\partial t} = \frac{1}{2}\Delta_C \rho - \nabla \cdot m\rho(1-\rho)$$

This requires justification that will be the content of our next lecture.

Let us now turn to the case where p has mean zero but is not symmetric. In this case

$$V_N(\eta) = N^{2-d}\sum_{x,y} \eta(x)(1-\eta(y))p(y-x)[J(\frac{y}{N}) - J(\frac{x}{N})]$$

and we get stuck at this point. If p is symmetric, as we saw, we gain a factor of N^{-2}. Otherwise the gain is only a factor of N^{-1} which is not enough. We seem to end up with

$$N^{-d}\sum_x\sum_y \eta(x) < \frac{1}{2}[(\nabla J)(\frac{x}{N}) + (\nabla J)(\frac{y}{N})]\,, N(1-\eta(y))(y-x)p(y-x) >$$
$$= \frac{1}{2N^d}\sum_x (\nabla J)(\frac{x}{N}) N\mathbf{\Psi}_x$$

where

$$\begin{aligned}\mathbf{\Psi}_x &= [\eta(x)\sum_z (1-\eta(x+z))z\,p(z) + (1-\eta(x))\sum_z \eta(x-z)z\,p(z)]\\ &= [-\eta(x)\sum_z \eta(x+z)z\,p(z) + (1-\eta(x))\sum_z \eta(x-z)z\,p(z)]\\ &= [\sum_z \eta(x-z)z\,p(z) - \eta(x)\sum_z (\eta(x+z)+\eta(x-z))z\,p(z)]\\ &= \tau_x\mathbf{\Psi}_0\end{aligned}$$

with τ_x being the shift by x. The second sum is zero in the symmetric case and $\mathbf{\Psi}_0$ can then be written as a 'gradient'

$$\mathbf{\Psi}_0 = \sum_j \tau_{e_j}\xi_j - \xi_j$$

where τ_{e_j} are shifts in the coordinate directions. This allows us to do summation by parts and gain a factor of N^{-1}. When this is not the case, we have a 'nongradient' model and the hydrodynamic limit can no longer be established by simple averaging.

The important ingredient in the analysis of gradient models is the ability to do averaging and replace quantities by their expected values calculated under various equlibrium distributions. Suppose $\mu_N$is a probability measure on the space Ω_N of configurations η on the periodic lattice $\mathbf{Z}_N^d$. We wish to think of μ_N as being a Bernoulli measure with some density ρ. The density ρ is not quite a constant, but a slowly varying function on function on $\mathbf{Z}_N^d$, and in fact a function of the macroscopic variable $\frac{x}{N}$. Let $g = g(\eta)$ be a local function depending on the configuration in some finite box around 0. By $Ave_\ell\, g$ we will denote the averaging process over the block(cube) $B_\ell = \{z : |z_j| \le \ell;\ 1 \le j \le d\}$ of side $2\ell + 1$

$$(Ave_\ell g)(\eta) = \frac{1}{(2\ell+1)^d}\sum_{z\in B_\ell} g(\tau_z\eta)$$

If $g(\eta) = \eta(0)$ this produces the empirical density $\bar\eta_{\ell,0}$. For any local g we can calculate its expected value with respect to the Bernoulli measure with density ρ and get

$$\hat g(\rho) = E^{P_\rho}\big[g(\eta)\big]$$

Averaging means, replacing $(Ave_\ell g)(\eta)$ by $\hat g(\bar\eta_{\ell,0})$ or more precisely showing that the error

$$\delta_{N,\ell} = E^{\mu_N}\left[\frac{1}{N^d}\sum_x |(Ave_\ell g)(\tau_x\eta) - \hat g(\bar\eta_{\ell,x})|\right]$$

goes to zero as $N \to \infty$ so long as $\ell \to \infty$ and $\frac{\ell}{N} \to 0$. What property of the sequence μ_N will allow us to make such a conclusion? **Dirichlet Form.** Given the density $\{f(\eta)\}$ with respect to the uniform distribution of a probability distribution μ on $\mathbf{Z}_N^d$, the Dirichlet form $\mathcal{D}_N(\mu)$ is the quantity

$$\mathcal{D}_N(\mu) = \frac{1}{2^{N^d}} \sum_{\eta} \sum_{\substack{x,y \in \mathbf{Z}_N^d \\ |x-y|=1}} [\sqrt{f(\eta^{x,y})} - \sqrt{f(\eta)}|^2$$

Remark: One should think of the inner summation as being over all the nearest neighbor bonds. The outer summation with the factor $\frac{1}{2^{N^d}}$ is integration with respect to the uniform distribution on Ω_N. Since f is a nonnegative L_1 function, its square root will be naturally an L_2 function and we are really dealing with the Dirichlet form of $\sqrt{f}$. **Theorem (Averaging principle).** If μ_N on Ω_N is such that

$$\mathcal{D}_N(\mu_N) \leq C N^{d-2}$$

for some constant C independent of N, then for every local g

$$\lim_{\epsilon \to 0} \limsup_{N \to \infty} \delta_{N,\epsilon N} = 0$$

Proof: The proof is carried out in two steps. First we show that

$$\lim_{\ell \to \infty} \limsup_{N \to \infty} \delta_{N,\ell} = 0$$

and then

$$\lim_{\substack{\epsilon \to 0 \\ \ell \to \infty}} E^{\mu_N} \left[\frac{1}{N^d} \sum_{x \in \mathbf{Z}_N^d} |\bar{\eta}_{\ell,x} - \bar{\eta}_{N\epsilon,x}| \right] = 0$$

Since $\hat{g}(\rho)$ is a polynomial in ρ, the two steps together will suffice to prove the theorem.

Step 1: The details of the proof can be found in [KOV] we will only sketch it here. Suppose B_ℓ is a block of size $2\ell + 1$ and $q_\ell(\eta)$ is an assignment of probabilities for the $2^{(2\ell+1)^d}$ possible configurations. We view the block non-periodically and we have $(2\ell)^d$ interior nearest neighbor bonds. The Dirichlet form is given by

$$\mathcal{D}_\ell(q_\ell) = \sum_{\substack{x,y \in B_\ell \\ |x-y|=1}} |\sqrt{q_\ell(\eta^{x,y})} - \sqrt{q_\ell(\eta)}|^2$$

Given a probability distribution μ_N on Ω_N, we denote its marginal on a block of size $2\ell+1$ centered at x by $\nu_{\ell,N,x}$ and by $\bar{\nu}_{\ell,N}$ we denote the average

$$\bar{\nu}_{\ell,N} = \frac{1}{N^d} \sum_{x \in \mathbf{Z}_N^d} \nu_{\ell,N,x}$$

over all the sites x. If $q_{\ell,N}(\eta)$ are the individual probabilities a simple consequence of the convexity of $\mathcal{D}(\nu)$ in ν yields

$$\mathcal{D}_\ell(q_{\ell,N}) \le (\frac{2\ell}{N})^d \mathcal{D}_N(\mu_N) \le (\frac{2\ell}{N})^d C N^{d-2} = \frac{C(2\ell)^d}{N^2}$$

It follows that any possible limit q_ℓ of $q_{\ell,N}$ as $N \to \infty$ has the property $\mathcal{D}_\ell(q_\ell) = 0$. Any probability distribution with $\mathcal{D}_\ell(q_\ell) = 0$ is invariant under permutations and the conditional distribution of occupied sites, given the total number of particles k, is uniform over all possible subsets of cardinality k. If we denote this uniform distribution by $\lambda_{\ell,k}$, then

$$\nu_l = \sum_k \lambda_{\ell,k} \pi_{\nu_\ell}(k)$$

Since $\lambda_{\ell,k}$ converges to the Bernoulli with density ρ, as $\frac{k}{(2\ell+1)^d} \to \rho$ (uniformly in ρ) it follows that

$$\lim_{\ell\to\infty} \sup_{\nu_\ell} E^{\nu_\ell}[|Ave_\ell\, g(\eta) - \hat{g}(Ave_\ell\, \eta)|] = 0$$

This completes step 1.

Proof of Step 2: For step 2, we need only establish that the local density $\bar{\eta}_{\ell,x}$ does not fluctuate over small macroscopic length scales. This will follow from the estimate

$$\limsup_{\substack{\ell\to\infty\\ \epsilon\to 0}} \limsup_{N\to\infty} \sup_{|y|\le N\epsilon} E^{\mu_N}\left[\frac{1}{N^d}\sum_{x\in\mathbf{Z}_N^d} |\bar{\eta}_{\ell,x} - \bar{\eta}_{\ell,x+y}|\right] = 0$$

If we take two blocks of size $2\ell+1$ centeerd at x and $x+y$ there is no bond connecting x and $x+y$ which are far from each other. We can introduce a Dirichlet form for this bond.

$$\mathcal{D}_{x,x+y}(q) = \sum_\eta |\sqrt{q(\eta^{x,x+y})} - \sqrt{q(\eta)}|^2$$

and this can be estimated, by writing the difference as a telescopic sum over exchanges involving nearest neighbors and applying Schwartz's inequlaity, interms of Dirichlet forms for nearest neighbor bonds. If we denote by $\nu_{\ell,N,x,x+y}$ the joint marginal over two widely seperated blocks of size $2\ell+1$ centered at x and $x+y$, and by

$$\bar{\nu}_{\ell,N,y} = \frac{1}{N^d}\sum_{x\in\mathbf{Z}_N^d} \nu_{\ell,N,x,x+y}$$

the average over all rigid spatial translations of the pair of blocks

$$\mathcal{D}_{\ell,0,y}(\bar{\nu}_{\ell,N,y}) \le C|y|^2 N^{-2} \le C\epsilon^2$$

It follows, by pure thought, (once $\epsilon \to 0$, the two averages are essentially averages over two halves of a combined system of $2(2\ell+1)^d$ sites which is in equilibrium), that

$$\limsup_{\ell\to\infty} \limsup_{\epsilon\to 0} \limsup_{N\to\infty} \sup_{|y|\le N\epsilon} E^{\mu_N}\left[\frac{1}{N^d}\sum_{x\in\mathbf{Z}_N^d} |\bar{\eta}_{\ell,x} - \bar{\eta}_{\ell,x+y}\,|\right] = 0$$

Because $\bar{\eta}_{N\epsilon,x}$ is not very different from the average $Ave_{|y|\leq N\epsilon}\,\bar{\eta}_{\ell,x+y}$ step 2 is essentially complete.

Super Exponential Estimates: Let us denote by P_N the measure corresponding to the process on $D[[0,T];\Omega_N]$ starting from some initial distribution. We want to replace additive functionals of the form

$$\int_0^T \frac{1}{N^d}\sum_x J(\frac{x}{N})g((\tau_x\eta)(s))ds$$

by

$$\int_0^T \frac{1}{N^d}\sum_x J(\frac{x}{N})\hat{g}(\bar{\eta}_{N\epsilon,x})(s))ds$$

and for the difference

$$F_N = \int_0^T \frac{1}{N^d}\sum_x J(\frac{x}{N})g((\tau_x\eta)(s))ds - \int_0^T \frac{1}{N^d}\sum_x J(\frac{x}{N})\hat{g}(\bar{\eta}_{N\epsilon,x})(s))ds$$

we obtain the following superexponential bound.

Theorem: For any $\delta > 0$

$$\limsup_{\epsilon\to 0}\limsup_{N\to\infty}\frac{1}{N^d}\log P_N[|F_N|\geq\delta\,] = -\infty$$

Proof: It suffices to prove that for any $\lambda > 0$

$$\limsup_{\epsilon\to 0}\limsup_{N\to\infty}\frac{1}{N^d}E^{P_N}\Big[\exp\big[\lambda N^d F_N\big]\Big] = 0$$

By Hölder inequality we can reduce it to equilibrium and, it is sufficient to prove

$$\limsup_{\epsilon\to 0}\limsup_{N\to\infty}\frac{1}{N^d}E^{\bar{P}_N}\Big[\exp\big[\lambda N^d F_N\big]\Big] = 0$$

where $\bar{P}_N$ is initializes to start from Bernoulli with density $\frac{1}{2}$. By Feynamn-Kac it is reduced to an estimation of an eigen value. More precisely we only need to show that for every $\lambda > 0$

$$\limsup_{\epsilon\to 0}\limsup_{N\to\infty}\frac{1}{N^d}\Big[\sup_\nu E^\nu\big[N^d\lambda F_N\big] - cN^2\mathcal{D}_N(\nu)\Big] = 0$$

or

$$\limsup_{\epsilon\to 0}\limsup_{N\to\infty}\Big[\sup_\nu E^\nu\big[\lambda F_N\big] - \frac{c}{N^{d-2}}\mathcal{D}_N(\nu)\Big] = 0$$

The apperance of the constant c is because the actual Dirichlet form involves $p(\cdot)$ and if we assume irreducibility we can get a bound. The factor N^2 is due to the speed up of time. If F_N is bounded by A, we will only see ν with

$$\mathcal{D}_N(\nu)\leq\frac{\lambda A}{c}N^{d-2} = CN^{d-2}$$

We are done due to the earlier result.

Application 1: Weak asymmetry. If Q_N is the process with weak asymmetry then

$$R_N = \frac{dQ_N}{dP_N}$$

can be explicitly calculated by Girsanov type formula and

$$E^{P_N}\left[R_N^p\right] \leq \exp\left[CN^d\right]$$

Therefore Q_N inherits the super exponential bounds. This establishes the hydrodynamic limit for weakly asymmetric perturbations.

In fact if we perturb $p(\cdot)$ by a skew symmetric $\frac{q}{N}$ with q of the form $q(t, \frac{x}{N}, \cdot)$ so that the generator looks like

$$\begin{aligned} N^2 \sum_{x,y} \eta(x)(1-\eta(y))p(y-x)[f(\eta^{x,y}) - f(\eta)] \\ + N \sum_{x,y} \eta(x)(1-\eta(y))q(t, \frac{x}{N}, y-x)[f(\eta^{x,y}) - f(\eta)] \end{aligned}$$

and the means are calculated as

$$b(t, \theta) = \sum_z z\, q(t, \theta, z),$$

then the hydrodynamic limit is

$$\frac{\partial \rho}{\partial t} = \frac{1}{2} \nabla \cdot C \nabla \rho - \nabla \cdot b\rho(1-\rho)$$

Application 2: Large Deviations. The Entropy cost of the perturbation is calculated easily by the use of Girsanov's formula

$$H(Q_N\,, P_N) = E^{Q_N}\left[\int_0^T \sum_{x,y} \left[\eta_t(x)(1-\eta_t(y))c_N(t, \frac{x}{N}, y-x)\right] dt\right]$$

where c is the relative entropy of one Poisson distribution to another

$$\begin{aligned} C_N(t, \theta, z) &= (N^2 p(z) + Nq(t, \theta, z)) \log \frac{N^2 p(z) + Nq(t, \theta, z)}{N^2 p(z)} \\ &\quad - (N^2 p(z) + Nq(t, \theta, z)) + N^2 p(z) \\ &\simeq \frac{q^2(t, \theta, z)}{2p(z)} \end{aligned}$$

It is clearly in our interest to minimize

$$\frac{1}{2} \sum_z \frac{q^2(t, \theta, z)}{p(z)}$$

subject to

$$\sum_z z\, q(t, \theta, z) = b$$

and the minimum is seen to be

$$\frac{1}{2} < b\,, C^{-1}\, b >$$

so that the minimal entropy cost H_N for the large deviation of the empirical density from the solution of the normal hydrodynamic limiting equation

$$\frac{\partial \rho}{\partial t} = \frac{1}{2}\nabla \cdot C\nabla \rho \tag{3.1}$$

to the solution of

$$\frac{\partial \rho}{\partial t} = \frac{1}{2}\nabla \cdot C\nabla \rho - \nabla \cdot b\rho(1-\rho) \tag{3.2}$$

is

$$H_N \simeq N^d \mathcal{E}(b)$$

where

$$\mathcal{E}(b) = \frac{1}{2} N^d \int_0^T \int_{\mathbf{T}^d} \rho(t\,,\theta)(1-\rho(t\,,\theta)) < b(t,\theta)\,, C^{-1} b(t,\,\theta) > dt d\theta \tag{3.3}$$

We have now a large deviation lower bound for a ρ not satisfying (3.1). Find the class $\mathcal{B}(\rho)$ of b's that satisfy (3.2) and optimize $\mathcal{E}(b)$ given in (3.3) over $b \in \mathcal{B}(\rho)$. In other words we have established the large deviation lower bound with the rate function

$$I(\rho) = \inf_{b \in \mathcal{B}(\rho)} \mathcal{E}(b) \tag{3.4}$$

We now work on the upper bound. We start with a class of exponential martingales. For any smooth test function $J(t,\theta)$

$$E^{P_N}\Big[\exp\Big[\sum_x J(T,\frac{x}{N})\,\eta_T(x) - \sum_x J(0,\frac{x}{N})\,\eta_0(x) - \int_0^T \sum_x J_t(t,\frac{x}{N})\eta_t(x)dt$$
$$- N^2 \int_0^T \sum_{x,y} p(y-x)\eta_t(x)(1-\eta_t(y))\big[\exp[J(t,\frac{y}{N}) - J(t,\frac{x}{N})] - 1\big]dt\Big]\Big] = 1$$

The quantity

$$N^2 \sum_{x,y} p(y-x)\eta_t(x)(1-\eta_t(y))\big[\exp[J(t,\frac{y}{N}) - J(t,\frac{x}{N})] - 1\big]$$

simplifies to

$$\frac{1}{2}\sum_x (\nabla \cdot C\nabla J)(t,\frac{x}{N})\eta_t(x) + \frac{1}{2}\sum_{x,y} p(y-x)\eta_t(x)(1-\eta_t(y))\,|< y-x\,, (\nabla J)(t,\frac{x}{N}) >|^2$$

Because of super exponential bounds we can effectively average the second term, even for large deviation purposes and interms of the density ρ, the expression looks like

$$N^d\Big[\frac{1}{2}\int_0^T \int_{\mathbf{T}^d} \nabla \cdot C\nabla J(t,\theta)\rho(t,\theta)dt d\theta$$
$$+ \frac{1}{2}\int_0^T \int_{\mathbf{T}^d} < \nabla J(t,\theta), C\nabla J(t,\theta) > \rho(t,\theta)(1-\rho(t,\theta))dt d\theta\Big]$$

Formally we are looking at

$$E^{P_N}\left[\exp\left[N^d F_N\right]\right] = 1$$

where

$$\begin{aligned}
F_N &= F_{N,J}(\rho(\cdot,\cdot)) \\
&\simeq \int_{\mathbf{Z}^d} J(T,\theta)\rho(T,\theta)d\theta - \int_{\mathbf{Z}^d} J(0,\theta)\rho(0,\theta)d\theta - \int_0^T \int_{\mathbf{Z}^d} J_t(t,\theta)\rho(t,\theta)dtd\theta \\
&- \frac{1}{2}\int_0^T \int_{\mathbf{T}^d} \nabla\cdot C\nabla J(t,\theta)\rho(t,\theta)dtd\theta \\
&- \frac{1}{2}\int_0^T \int_{\mathbf{T}^d} < \nabla J(t,\theta), C\nabla J(t,\theta) > \rho(t,\theta)(1-\rho(t,\theta))dtd\theta
\end{aligned}$$

By standard large deviation theory one gets an upper bound with a rate function

$$\sup_J F_{N,J}(\rho(\cdot,\cdot))$$

that is easily seen to equal (3.4).