QUANTITATIVE HOMOGENIZATION AND HYDRODYNAMIC LIMIT OF NON-GRADIENT EXCLUSION PROCESS

TADAHISA FUNAKI, CHENLIN GU, HAN WANG

ABSTRACT. For the non-gradient exclusion process, we prove its approximation rate of diffusion matrix/conductivity by local functions. The proof follows the quantitative homogenization theory developed by Armstrong, Kuusi, Mourrat and Smart, while the new challenge here is the hard core constraint of particle number on every site. Therefore, a coarse-grained method is proposed to lift the configuration to a larger space without exclusion, and a gradient coupling between two systems is applied to capture the spatial cancellation. Moreover, the approximation rate of conductivity is uniform with respect to the density via the regularity of the local corrector. As an application, we integrate this result in the work by Funaki, Uchiyama and Yau [*IMA Vol. Math. Appl.*, 77 (1996), pp. 1–40.] and yield a quantitative hydrodynamic limit. In particular, our new approach avoids to show the characterization of closed forms. We also discuss the possible extensions in the presence of disorder on the bonds.

MSC 2010: 82C22, 35B27, 60K35.

KEYWORDS: interacting particle system, non-gradient exclusion process, Kawasaki dynamics, diffusion matrix, quantitative homogenization, hydrodynamic limit.

CONTENTS

1. INTRODUCTION

The diffusion matrix plays an important role in the study of the large-scale behaviors of interacting particle systems. Among these systems, some are classified as *gradient model* if the current of the conserved quantity can be written as a sum of the difference between a local function and its spatial shift, and the others are called *non-gradient model*. Unlike the gradient model, the non-gradient model usually requires more techniques to derive the hydrodynamic limit, because the scaling yields a diverging factor and, moreover, its diffusion matrix does not have an explicit expression and one needs to add a correction. This idea goes back to the seminal work [74] of Varadhan, who studied the Ginzburg–Landau model. Later, the hydrodynamic limit was proved in several classical particle systems of non-gradient type: *the generalized symmetric exclusion process (GSEP)* [52] by Kipnis, Landim and Olla; *the lattice gas* [34], also known as *the non-gradient Kawasaki dynamics*, by Funaki, Uchiyama and Yau, and the general lattice gas with mixing condition [75] by Varadhan and Yau; *the multi-type simple symmetric exclusion process (multi-type SSEP)* [69] by Quastel, etc. The equilibrium fluctuation in non-gradient models was studied later in [31, 57, 19], and the regularity of the diffusion matrix was discussed in a series of work [54, 65, 66, 67, 16]. One can also refer [72, 51] for the basic background and [71] for the relation between the gradient condition and the Green–Kubo formula.

It is natural to ask the convergence rate, and the quantitative hydrodynamic limit has received attentions recently. The related results can be found for the Ginzburg–Landau model in [26, 27], which was developed on [45]. Very recently, [60] proposed a *consistence-stability approach* to obtain the quantitative hydrodynamic limit in Wasserstein-1 distance for several models including the zero-range process, the Ginzburg–Landau model, and the simple exclusion process (see [59, Chapter 5.5]). However, all these results are for the gradient model, and there is no results for the non-gradient model in the literature to the best of our knowledge. This is because, as mentioned in the last paragraph, a diverging factor appears and the diffusion matrix is more complicated in the non-gradient model. Such obstacle was already observed in the proof in [34, 31], where several key error terms are finally reduced to the approximation of the diffusion matrix or *the conductivity*; see [34, (2.5) and Section 5]. These two fundamental quantities are defined using variational formulas and are related by the Einstein relation. Therefore, both of them can be approximated qualitatively by a sequence of local functions, but the convergence rate is unknown.

In this paper, we answer the question above by a concrete construction of desired local functions. As Varadhan observed the link between the interacting systems and homogenization, in the sense of averaging and gradient replacement which kills the diverging factor, in the earlier work [74], the new improvement comes from the recent progress in the quantitative homogenization theory; see [15, 14, 9, 10, 6, 5, 11, 7] based on the renormalization approach, and [64, 62, 39, 40, 37, 41, 38] based on another approach using spectral inequalities. As an example, for the $\nabla\phi$ interface model studied in [33], a quantitative hydrodynamic limit is obtained in [4] using the renormalization approach. Also inspired by the renormalization approach, [35] studies a similar diffusion matrix problem in continuous configuration space, and a quantitative equilibrium fluctuation is obtained recently in [47] under the same setting. These work pave the way for the quantitative homogenization theory in interacting particle systems, but the continuous configuration model there relaxes the hard core constraint by allowing arbitrarily large number of particles in the unit volume. As a consequence, to apply the existing results to a lattice particle model of exclusion rule still meets technical challenges in math. The present paper aims to resolve these difficulties and is the first example to establish quantitative homogenization theory on the non-gradient exclusion process. Our main result not only generalizes [35] to the non-gradient exclusion process, but also improves in the sense that we construct *one* local corrector to realize uniform convergence of conductivity for *every* particle density. Moreover, this density-uniform homogenization can be integrated into the relative entropy method in the classical work [34], and then establishes a quantitative hydrodynamic limit; namely, our result provides the convergence rate in the hydrodynamic limit for non-gradient exclusion process. We emphasize that our method is new and, in particular, avoids to prove the characterization of closed forms which is usually required to show the hydrodynamic limit for non-gradient models.

Our proof is also robust. Viewing the recent interests on the exclusion process in random environment (see [70, 49, 50, 42, 30, 43, 28, 29]), we give a quick generalization for our case when the external disorder is posed on the bonds. We believe the results in this paper can be extended to other models including GSEP and multi-type SSEP in the future work.

1.1. **Main results.** In this part, we state our main results.

We recall quickly the necessary notations of the exclusion process and the results in the previous work. Let $\mathbb{Z}^d$ be the Euclidean lattice and we use $\mathcal{X} := \{0,1\}^{\mathbb{Z}^d}$ to stand for the space of the configuration of particles under exclusion rule. The element of $\mathcal{X}$ will be denoted by $\eta = \{\eta_x : x \in \mathbb{Z}^d\}$. Here $\eta_x = 0$ means the site x is vacant and $\eta_x = 1$ means the site is occupied by one particle. We denote by $y \sim x$ for $x, y \in \mathbb{Z}^d$ if $|x-y| = 1$. Then $\{x,y\}$ is called an (undirected) bond. For every $\Lambda \subseteq \mathbb{Z}^d$, we denote by Λ^* the bond in Λ that

$$\Lambda^* := \{\{x,y\} : x, y \in \Lambda, x \sim y\}. \tag{1.1}$$

For $x, y \in \mathbb{Z}^d$, the exchange operator $\eta^{x,y}$ is defined as

$$(\eta^{x,y})_z := \begin{cases} \eta_z, & z \neq x, y; \\ \eta_y, & z = x; \\ \eta_x, & z = y. \end{cases} \tag{1.2}$$

Especially, when $b = \{x,y\}$ is a bond, we also write η^b instead of $\eta^{x,y}$, and define the Kawasaki operator $\pi_b \equiv \pi_{x,y}$

$$\pi_b f(\eta) := f(\eta^b) - f(\eta). \tag{1.3}$$

For any $x \in \mathbb{Z}^d$, the translation operator τ_x is defined as

$$(\tau_x \eta)_y := \eta_{x+y}, \tag{1.4}$$

and for function f on $\mathcal{X}$, we also define $\tau_x f$ as

$$(\tau_x f)(\eta) = f(\tau_x \eta). \tag{1.5}$$

The *non-gradient exclusion process* on $\mathbb{Z}^d$ is defined by the generator below

$$\mathcal{L} := \sum_{b \in (\mathbb{Z}^d)^*} c_b(\eta) \pi_b = \frac{1}{2} \sum_{x,y \in \mathbb{Z}^d : |x-y|=1} c_{x,y}(\eta) \pi_{x,y}, \tag{1.6}$$

where the family of functions

$$\{c_b(\eta) \equiv c_{x,y}(\eta) = c_{y,x}(\eta); b = \{x,y\} \in (\mathbb{Z}^d)^*\}, \tag{1.7}$$

determine the jump rate of particles on the nearest bonds. This model is also called *the speed-change Kawasaki dynamics* or *the lattice gas* in the literature, and we will also use these names alternatively from time to time in the paper.

We suppose the following conditions for the jump rate throughout the paper without specific explanation.

Hypothesis 1.1. The following conditions are supposed for $\{c_b\}_{b \in (\mathbb{Z}^d)^*}$.

(1) Non-degenerate and local: $c_{x,y}(\eta)$ depends only on $\{\eta_z : |z-x| \leqslant \mathbf{r}\}$ for some integer $\mathbf{r} > 0$, and is bounded on two sides $1 \leqslant c_{x,y}(\eta) \leqslant \lambda$.

(2) Spatially homogeneous: for all $\{x,y\} \in (\mathbb{Z}^d)^*$, $c_{x,y} = \tau_x c_{0,y-x}$.
(3) Detailed balance under Bernoulli measures: $c_{x,y}(\eta)$ is independent of $\{\eta_x, \eta_y\}$.

This model is known of non-gradient type, i.e. we cannot find functions $\{h_{i,j}\}_{1\leqslant i,j\leqslant d}$ such that $c_{0,e_i}(\eta)(\eta_{e_i} - \eta_0) = \sum_{j=1}^d \left((\tau_{e_j} h_{i,j})(\eta) - h_{i,j}(\eta)\right)$ for general $\{c_b\}_{b\in(\mathbb{Z}^d)^*}$, with $\{e_i\}_{1\leqslant i\leqslant d}$ the canonical basis of $\mathbb{Z}^d$.

The hydrodynamic limit of this speed-change Kawasaki dynamics on torus is proved in [34]. Let $\mathbb{T}_N^d := (\mathbb{Z}/N\mathbb{Z})^d$ be the lattice torus of scale N, where we can define all the notations by replacing $\mathbb{Z}^d$ with $\mathbb{T}_N^d$. We denote by $\mathcal{X}_N := \{0,1\}^{\mathbb{T}_N^d}$ the configuration space on $\mathbb{T}_N^d$, and define $\eta^N(t) = \{\eta_x^N(t), x \in \mathbb{T}_N^d\}$ as the $\mathcal{X}_N$-valued Markov jump process on torus governed by the generator $\mathcal{L}_N := N^2\mathcal{L}$, the counterpart of (1.6) on $\mathbb{T}_N^d$. The macroscopic empirical measure of $\eta^N(t)$ is defined as

$$\rho^N(t, \mathrm{d}v) := N^{-d} \sum_{x\in\mathbb{T}_N^d} \eta_x^N(t)\delta_{x/N}(\mathrm{d}v), \qquad v \in \mathbb{T}^d, \tag{1.8}$$

and the limit is the solution of a nonlinear diffusion equation

$$\partial_t \rho(t,v) = \nabla \cdot \left(D(\rho(t,v))\nabla\rho(t,v)\right), \qquad (t,v) \in \mathbb{R}_+ \times \mathbb{T}^d. \tag{1.9}$$

Here $\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$ is the continuous torus and $D : (0,1) \to \mathbb{R}^{d\times d}$ is *the diffusion matrix* defined by *the Einstein relation*

$$D(\rho) := \frac{\mathbf{c}(\rho)}{2\chi(\rho)}, \tag{1.10}$$

where $\chi(\rho)$ is *the compressibility*

$$\chi(\rho) := \rho(1-\rho), \tag{1.11}$$

and $\mathbf{c}(\rho)$ is *the effective conductivity* defined as follows. We construct at first a quadratic form with respect to the function $F \in \mathcal{F}_0^d$

$$\xi \cdot \mathbf{c}(\rho; F)\xi = \frac{1}{2}\sum_{|x|=1} \left\langle c_{0,x}\left(\xi \cdot \left\{x(\eta_x - \eta_0) - \pi_{0,x}\Big(\sum_{y\in\mathbb{Z}^d} \tau_y F\Big)\right\}\right)^2\right\rangle_\rho, \tag{1.12}$$

where $\mathcal{F}_0$ is the local function space on $\mathcal{X}$ and $\mathcal{F}_0^d := (\mathcal{F}_0)^d$, and $\langle\cdot\rangle_\rho$ stands for the expectation under Bernoulli product measure of density $\rho \in [0,1]$. Then $\mathbf{c}(\rho)$ is the minimization of $\mathbf{c}(\rho; F)$

$$\xi \cdot \mathbf{c}(\rho)\xi := \inf_{F\in\mathcal{F}_0^d} \xi \cdot \mathbf{c}(\rho; F)\xi. \tag{1.13}$$

Under the assumption that (1.9) has a smooth initial density $\rho_0 = \rho_0(v)$ and $\eta^N(0)$ is close to the local equilibrium with density profile $\rho_0(v)$ in the sense of relative entropy $h_N(0) = o(1)$ (see (7.13) for the definition), [34, Theorem 1.1] proves that for every $\phi \in C^\infty(\mathbb{T}^d)$ and $\varepsilon > 0$

$$P\left[\left|\int_{\mathbb{T}^d} \phi(v)\rho^N(t,\mathrm{d}v) - \int_{\mathbb{T}^d} \phi(v)\rho(t,\mathrm{d}v)\right| > \varepsilon\right] \xrightarrow{N\to\infty} 0. \tag{1.14}$$

Here P stands for the probability space of the process $(\eta^N(t))_{t\in\mathbb{R}_+}$, and $\rho(t,\mathrm{d}v) := \rho(t,v)\mathrm{d}v$.

The proof of hydrodynamic limit in [34] relies on *the relative entropy method.* One key step is to prove that, for every $\beta > 0$ and small $\delta > 0$, we have the following estimate for the normalized relative entropy (with some other terms omitting details on the right-hand side)

$$h_N(t) \leqslant h_N(0) + \frac{1}{\delta}\int_0^t h_N(s)\,\mathrm{d}s + C(\beta+1)\sup_{\rho\in[0,1]}|R(\rho;F)| + \frac{C}{\beta} + Q_N. \tag{1.15}$$

Here the quantity $R(\rho;F)$ comes from the conductivity (1.12) and (1.13), which is defined as

$$R(\rho;F) := \mathbf{c}(\rho;F) - \mathbf{c}(\rho), \tag{1.16}$$

see (7.16) for the other error terms in Q_N. To conclude (1.14), we take a large β such that the right-hand side of (1.15) is small, and then apply Gronwall's inequality and the entropy inequality. Therefore, we also need the decay from the term $R(\rho;F)$, which was proved in [34, Lemma 2.1] that

$$\inf_{F\in\mathcal{F}_0^d}\sup_{\rho\in[0,1]}|R(\rho;F)| = 0. \tag{1.17}$$

The object of this paper is to give a convergence rate of (1.14). As explained briefly above, this is finally reduced to $R(\rho;F)$ and we need to study (1.17) more precisely. It can be considered as a quantitative homogenization of the fundamental quantity $\mathbf{c}(\rho)$, which is our main result stated as follows. In the statement, $\Lambda_L := (-\frac{L}{2},\frac{L}{2})^d\cap\mathbb{Z}^d$ stands for a hypercube of side length around $L\in\mathbb{R}_+$, and $\mathcal{F}_0^d(\Lambda_L)$ is the subset of $\mathcal{F}_0^d$ which contains $\sigma(\{\eta_x\}_{x\in\Lambda_L})$-measurable local functions.

Theorem 1.2. *Under Hypothesis 1.1, there exists an exponent $\gamma(d,\lambda,\mathbf{r})>0$ and a positive constant $C(d,\lambda,\mathbf{r})<\infty$, such that*

$$\inf_{F_L\in\mathcal{F}_0^d(\Lambda_L)}\sup_{\rho\in[0,1]}|R(\rho;F_L)| \leqslant CL^{-\gamma}. \tag{1.18}$$

We will also give the concrete construction of the local function achieving the estimate above in our proof; see Section 6.4 for details. Then as expected, we can insert the estimate in (1.15), and obtain a quantitative hydrodynamic limit after careful investigation.

Theorem 1.3. *Let $\rho(t,v)$ be the solution of the hydrodynamic equation (1.9) for $t\in[0,T]$ with a smooth initial value ρ_0 such that $0<\rho_0(v)<1$. Assume that f_0 and ψ_0 defined in (7.14) satisfy the entropy condition $h_N(0)\leqslant CN^{-\alpha}$ for some $C,\alpha>0$. Then, for every $\varepsilon>0$ and $\phi\in C^\infty(\mathbb{T}^d)$, there exist $\kappa>0$ and $C=C(\varepsilon,\phi)>0$ such that*

$$P\left[\left|\int_{\mathbb{T}^d}\phi(v)\rho^N(t,\mathrm{d}v) - \int_{\mathbb{T}^d}\phi(v)\rho(t,\mathrm{d}v)\right| > \varepsilon\right] \leqslant CN^{-\kappa} \tag{1.19}$$

holds for all $t\in[0,T]$.

Our proof relies on the homogenization theory, so let us also state a variant of (1.18), which is an intermediate step but is related to the CLT variance estimate in [34, Section 5]. Consider a formal sum $\ell_\xi = \sum_{x\in\mathbb{Z}^d}(\xi\cdot x)\eta_x$, then we notice that the term $\xi\cdot x(\eta_x-\eta_0)$ in (1.12) is

$$\xi\cdot x(\eta_x-\eta_0) = -\pi_{0,x}\ell_\xi.$$

Thus (1.13) is the minimization of the Dirichlet energy of a linear statistic plus some correction term. Inspired from the ergodic theory, a natural finite-volume approximation in $\Lambda\subseteq\mathbb{Z}^d$ of (1.13) should be

$$\frac{1}{2}\xi\cdot\overline{\mathbf{c}}(\rho,\Lambda)\xi := \inf_{v\in\ell_\xi+\mathcal{F}_0(\Lambda^-)}\frac{1}{|\Lambda|}\left\langle v(-\mathcal{L}_\Lambda v)\right\rangle_\rho. \tag{1.20}$$

Here Λ^- is the interior of Λ and $\overline{\Lambda^*}$ is the set of bonds issued from Λ (see (1.34), (1.35) for details). $\mathcal{F}_0(\Lambda^-)$ is the set of $\sigma(\{\eta_x\}_{x\in\Lambda^-})$-measurable local functions, and $\mathcal{L}_\Lambda$ is the generator on Λ

$$\mathcal{L}_\Lambda v := \sum_{b\in\overline{\Lambda^*}} c_b \pi_b v. \tag{1.21}$$

A well-known integration by part formula under $\langle\cdot\rangle_\rho$ also tells us

$$\langle v(-\mathcal{L}_\Lambda v)\rangle_\rho = \frac{1}{2}\sum_{b\in\overline{\Lambda^*}} \left\langle c_b(\pi_b v)^2\right\rangle_\rho, \tag{1.22}$$

so (1.20) can be interpreted by the minimization of the Dirichlet energy contributed by every particle. We expect that $\overline{\mathbf{c}}(\rho,\Lambda)$ converges to $\mathbf{c}(\rho)$ when $\Lambda \nearrow \mathbb{Z}^d$.

A similar definition like (1.20) can be also posed under the canonical ensemble

$$\frac{1}{2}\xi\cdot\hat{\mathbf{c}}(\Lambda,N)\xi := \inf_{v\in\ell_\xi+\mathcal{F}_0(\Lambda^-)} \frac{1}{|\Lambda|}\langle v(-\mathcal{L}_\Lambda v)\rangle_{\Lambda,N}, \tag{1.23}$$

where $\langle\cdot\rangle_{\Lambda,N}$ is the expectation under the uniform measure of N particles in Λ. Notice that the quantity $\hat{\mathbf{c}}(\Lambda,N)$ still depends on the configuration outside Λ. On the other hand, because the jump rate c is of finite range $\mathbf{r}$ by Hypothesis 1.1, the influence from the boundary layer vanishes when $\Lambda \nearrow \mathbb{Z}^d$. Thus $\hat{\mathbf{c}}(\Lambda,N)$ should be close to $\mathbf{c}(N/|\Lambda|)$ in large scale. We prove the convergence of these two quantities in the following theorem.

Theorem 1.4. *Under Hypothesis 1.1, there exists a constant $C(d,\lambda,\mathbf{r})<\infty$ and two exponents $\gamma_1(d,\lambda,\mathbf{r}),\gamma_2(d,\lambda,\mathbf{r})>0$ such that for every $L,M\in\mathbb{N}_+$,*

$$|\overline{\mathbf{c}}(\rho,\Lambda_L)-\mathbf{c}(\rho)| \leqslant CL^{-\gamma_1}, \tag{1.24}$$

and

$$|\hat{\mathbf{c}}(\Lambda_L,M)-\mathbf{c}(M/|\Lambda_L|)| \leqslant CL^{-\gamma_2}. \tag{1.25}$$

Recalling the Einstein relation (1.10), the results above also imply the convergence rate of the diffusion matrix $D(\rho)$, for $\rho\in(0,1)$, by local functions or finite-volume approximation.

We also obtain an estimate similar to (1.24) when the disorder is posed on bonds. To lighten the notation, we leave the related discussion in Section 8.

Remark 1.5. The choice of notation $\overline{\Lambda^*}$ here is just for the technical convenience and the consistence. Lemma A.1 ensures the stability of Theorem 1.4 in the general domain, so we can replace $\overline{\Lambda^*}$ in (1.21) by the canonical notation Λ^* defined in (1.1) and the statement still holds.

1.2. **Strategy of the proof.** The idea of the proof in this paper is inspired by recent developments in the quantitative homogenization, and in particular on the renormalization approach developed in [15, 14, 9, 10, 6, 5]; see monographs [11, 7] and [63] for a gentle introduction. This renormalization approach has shown its robustness in a number of other settings including the parabolic equations [1], finite-difference equations on percolation clusters [3, 21, 23], differential forms [22], the "$\nabla\phi$" interface model [20, 12, 13, 4], and the Villain model [24]. Recently [35, 36, 47] also generalizes the theory to an interacting particle system in continuous space *without* exclusion, thus let us discuss the novelty and contribution in this paper.

1.2.1. *Renormalization with coarse-grained lifting.* We employ the renormalization approach to prove (1.24) of Theorem 1.4, which serves as the cornerstone for the other results. The ingredients of the renormalization approach can be roughly divided into two parts:

(1) Find the subadditive quantity for the desired limit, and use the gap between the subadditive quantity and its dual quantity to control the convergence rate.
(2) Establish various analytic tools, where the two key estimates are
  - the Caccioppoli inequality;
  - the multiscale Poincaré inequality.

The first part is more conceptual, and the quantity $\overline{\mathbf{c}}(\rho, \Lambda)$ defined in (1.20) is a good candidate satisfying the subadditivity in our setting, i.e.

$$\Lambda = \bigsqcup_{n=1}^{N} \Lambda^{(i)}, \qquad \overline{\mathbf{c}}(\rho, \Lambda) \leqslant \sum_{i=1}^{N} \frac{|\Lambda^{(i)}|}{|\Lambda|} \overline{\mathbf{c}}(\rho, \Lambda^{(i)}). \tag{1.26}$$

The main issue comes from the second part. The two key inequalities are well-developed for the elliptic equation on $\mathbb{R}^d$, but can become challenging in other settings. We should also highlight that, the two inequalities are more than the technical estimates, but the essentials of quantitative homogenization, because they characterize the elliptic conditions in the large scale; see the recent work for the homogenization in high contrast [2], [8]. In Kawasaki dynamics, these inequalities are not accessible directly, and we need to relax them respectively to *the modified Caccioppoli inequality* and *the weighted multiscale Poincaré inequality*, which are explained in the following paragraphs more carefully.

The classical Caccioppoli inequality describes the inner regularity of the elliptic equations, but seems missing in the particle systems due to the influence of particles near the boundary. Therefore, the modified Caccioppoli inequality is developed in the previous work [35, Proposition 3.9], which differs from the classical one, but captures the same spirit. In the present work, its counterpart in Kawasaki dynamics is also recovered in Proposition 2.6. The proof requires more work due to the microscopic behavior; see Lemma 2.1 and Lemma 2.7. As new inputs, the Glauber derivative and *the reverse Efron–Stein inequality* are also involved.

The multiscale Poincaré inequality (see [11, Proposition 1.12 and Corollary 1.14] for example) improves the estimate of the classical Poincaré inequality when the function has the spatial cancellation property, which is the case in homogenization. It meets obstacles to derive the counterpart for Kawasaki dynamics, because the generator is not smooth enough to ensure the H^2 estimate needed in the proof. Actually, even the generator of the simple symmetric exclusion shows interaction in higher order derivative; see Remark 2.10 for details. For this reason, we believe the multiscale Poincaré inequality should live in the homogenized particle system $\widetilde{\mathcal{X}} = \mathbb{N}^{\mathbb{Z}^d}$ where $\mathbb{N} = \{0, 1, 2, 3, \cdots\}$, i.e. the independent particles, and we prove it in Section 2.2.

Then a crucial problem is how to apply an inequality on $\widetilde{\mathcal{X}} = \mathbb{N}^{\mathbb{Z}^d}$ to the functions on $\mathcal{X} = \{0,1\}^{\mathbb{Z}^d}$. Similar problem on the percolation setting was also posed, and a possible solution is the coarse-grained strategy; see [3, 21, 46, 23]. We hope to implement this idea in the exclusion processes: for every function $u : \mathcal{X} \to \mathbb{R}$, we aim to find a coarsened function $[u] : \widetilde{\mathcal{X}} \to \mathbb{R}$ on the larger space such that for every $\eta \in \mathcal{X}$ as the grain of $\widetilde{\eta} \in \widetilde{\mathcal{X}}$, it satisfies

$$\widetilde{\eta} \simeq \eta \Longrightarrow [u](\widetilde{\eta}) \simeq u(\eta). \tag{1.27}$$

A naive candidate of grain η is the one close to $\widetilde{\eta}$ under some distance. However, different from the Bernoulli percolation setting, in particle systems the space of grain $\mathcal{X}$ is a very sparse subset of $\widetilde{\mathcal{X}}$, so we face *the curse of dimension* which may extremely enlarge the error in (1.27). This is also the key difficulty compared to the previous work [35]. Our solution turns out not only a coarse-grained method of functions, but also a *lift* from $\mathcal{X}$ to $\widetilde{\mathcal{X}}$, i.e. we can represent the function on Kawasaki dynamics using a coupled independent particles. More precisely, for every $\widetilde{\eta} \in \widetilde{\mathcal{X}}$, we set its grain $[\widetilde{\eta}] \in \mathcal{X}$ as

$$\forall x \in \mathbb{Z}^d, \qquad [\widetilde{\eta}]_x := \mathbf{1}_{\{\widetilde{\eta}_x \geqslant 1\}},$$

and the coarsened function for $u : \mathcal{X} \to \mathbb{R}$ as

$$[u] : \widetilde{\mathcal{X}} \to \mathbb{R}, \qquad [u](\widetilde{\eta}) := u([\widetilde{\eta}]).$$

This *coarse-grained lifting* is introduced in Section 3. Based on this technique, we also obtain a *gradient coupling* between two systems (see Propositions 3.2 and 3.6), and a weighted multiscale Poincaré inequality on Kawasaki dynamics. They are the main tools to evaluate the flatness of the functions in Proposition 5.3.

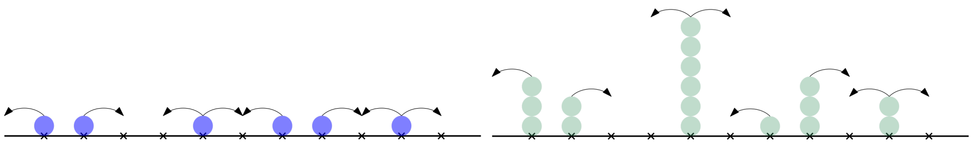

FIGURE 1. An illustration for the coarse-grained lifting between the Kawasaki dynamics and independent particles.

1.2.2. *Regularity and uniform convergence.* Usually the convergence rate depends on the particle density, so let us explain why a uniform convergence is valid. A first qualitative argument is that, our finite-volume approximation decreases to the limit, and the limit function $\rho \mapsto \mathbf{c}(\rho)$ is continuous thanks to [65], so Dini's theorem applies and the convergence is uniform. At the quantitative level, we highlight that, the only step where the density involves in analysis is the weighted multiscale Poincaré inequality, where some large factors can be added for the low density cases. Meanwhile, the modified Caccioppoli inequality (2.10) uses the elliptic regularity and the variance decay estimate (5.20) uses the spatial independence, so they are free from particle density. Here we notice that $\mathbf{c}$ has a trivial bound by $\chi(\rho) = \rho(1-\rho)$, and this can help us at two endpoints.

We still need some more ingredients to pass the results from Theorem 1.4 to Theorem 1.2. The uniform estimates in (1.24) can be seen as a weak version

$$\sup_{\rho \in [0,1]} \inf_{F_{\rho,L} \in \mathcal{F}_0^d(\Lambda_L)} R(\rho; F_{\rho,L}) \leqslant CL^{-\gamma_1},$$

while quantity in (1.18) is a strong version, and usually we have

$$\sup_{\rho \in [0,1]} \inf_{F_{\rho,L} \in \mathcal{F}_0^d(\Lambda_L)} R(\rho; F_{\rho,L}) \leqslant \inf_{F_L \in \mathcal{F}_0^d(\Lambda_L)} \sup_{\rho \in [0,1]} R(\rho; F_L).$$

We do not know whether there exists any duality property in the function $R(\rho; F)$, thus we make the proof by a direct construction. The minimizer F in this variational problem is actually *the correctors* in homogenization theory. In the renormalization step, we already get a candidate $\phi_{\rho,\Lambda,\xi}$ for the problem (1.20), but it has dependence on the density. Our main task is to remove this dependence, so we propose a modified version of the local corrector

$$\Lambda = \bigsqcup_{n=1}^{N} \Lambda^{(i)}, \qquad \widetilde{\phi}_{\hat{\rho},\Lambda,\xi} = \sum_{i=1}^{N} \phi_{\hat{\rho},\Lambda_i,\xi}, \tag{1.28}$$

where $\hat{\rho}$ is an empirical density on the domain Λ instead of a fixed designed density. Insert this function in the problem (1.20), we obtain a uniform convergence under grand canonical ensemble by the following reasons.

(1) The homogenization appears in the large scale, so (1.26) is nearly an equality and $\sum_{i=1}^{N} \phi_{\rho,\Lambda_i,\xi}$ nearly equals $\phi_{\rho,\Lambda,\xi}$.
(2) Given an empirical density $\hat{\rho}$, each local corrector lives as if under the grand canonical ensemble thanks to the local equivalence of ensembles. This is also the trick in the proof of (1.25).
(3) The empirical density $\hat{\rho}$ may also fluctuate when applying the Kawasaki operator, but this can be handled. On the one hand, we have the regularity of the mapping $\rho \mapsto \phi_{\rho,\Lambda_i,\xi}$, and each fluctuation of density is just $1/|\Lambda|$. On the other hand, such fluctuation only happens on the boundary layer of Λ, whose order is dominated by the volume order in (1.20).

Similar argument actually has already appeared in the proof of (1.17) in [34, Lemma 2.1]. Besides the quantitative homogenization in (1), we also need to calculate carefully the errors in (2) and (3), i.e. that from local equivalence of ensembles and the regularity of density. They are discussed in detail in Sections 6.2 and 6.1, and then we justify the density-free corrector (1.28) in Section 6.3.

1.2.3. *Hydrodynamic Limit.* Overall, our method well fits proving the hydrodynamic limit even with a quantitative convergence rate for the non-gradient Kawasaki dynamics. The main difficulty to study non-gradient models lies, in general, in the fact that the microscopic current does not have a gradient form and this yields a diverging factor under the scaling. To overcome this difficulty, we need to show that, under a large space-time domain, one can replace such a term by a well-behaving function of gradient form asymptotically. This is called *the gradient replacement* (see [34, Theorem 3.2 and Lemma 3.4]). For this, we usually need to show Varadhan's lemma (see [74, Theorem 5.2] and [34, Theorem 4.1]) which gives the characterization of closed forms defined on a configuration space. We observe a connection between the gradient replacement and the dual quantity employed in the renormalization approach (see Section 7.1), thus our method provides another new route for the non-gradient hydrodynamic limit avoiding Varadhan's lemma.

1.3. Organization of paper.

The rest of paper is organized as follows. We finish the introduction with a resume of notations, especially those about function spaces. In Section 2, we present the necessary tools including the modified Caccioppoli inequality for the Kawasaki dynamics. Then we introduce the coarse-grained lifting technique in Section 3, and use it to derive the weighted multiscale Poincaré inequality. Sections 4 and 5 are devoted to the convergence rate, where we make use of the renormalization approach. Afterwards, we study the regularity of the local corrector and remove the density dependence in Section 6. The quantitative hydrodynamic limit is proved in Section 7 and the extension for the disordered cases is discussed in Section 8. See the outline in Figure 2 for details.

1.4. Notations.

We resume the notations used throughout the paper.

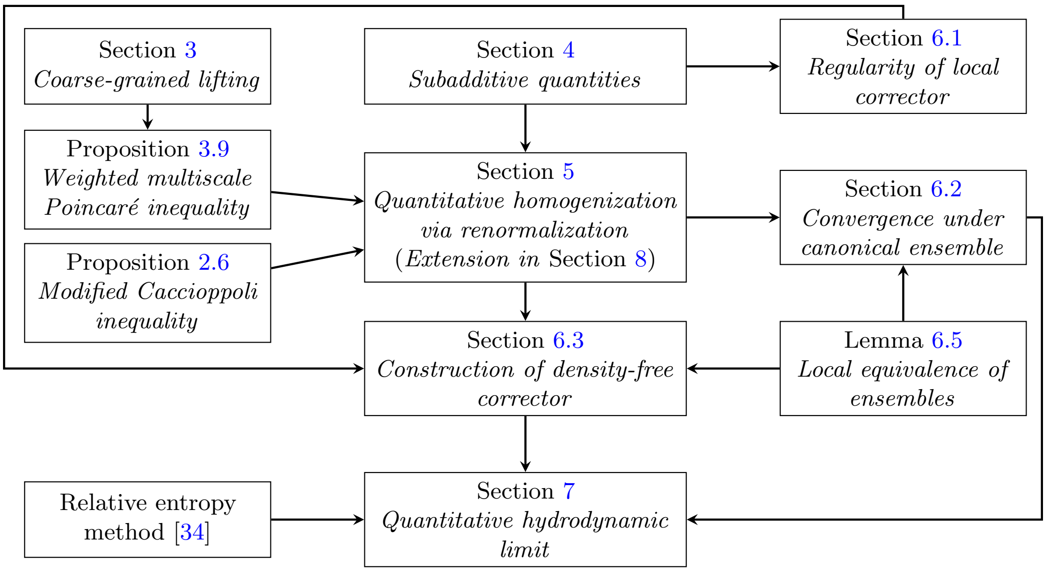

FIGURE 2. The outline of proof.

1.4.1. *Geometry.* We use $|\cdot|$ to stand for the usual ℓ^2-norm for the finite dimensional vector or matrices. Meanwhile, for any $x, y \in \mathbb{Z}^d$, we define

$$\operatorname{dist}(x,y) := \max\{|x_1 - y_1|, |x_2 - y_2|, \cdots, |x_d - y_d|\}. \tag{1.29}$$

This also generalizes to $\operatorname{dist}(x,\Lambda) := \sup_{y\in\Lambda} \operatorname{dist}(x,y)$ for every $\Lambda \subseteq \mathbb{Z}^d$.

We denote by $\Lambda_L := (\frac{L}{2}, \frac{L}{2})^d \cap \mathbb{Z}^d$ the hypercube of side length around L, where $L \in \mathbb{R}_+$ is not necessarily an integer for the flexibility. For every $m \in \mathbb{N} = \{0, 1, 2, \cdots\}$, we also denote by $\square_m := \left(-\frac{3^m}{2}, \frac{3^m}{2}\right)^d \cap \mathbb{Z}^d$ the hypercube of side length 3^m. For any $n, m \in \mathbb{N}$ such that $n < m$, we denote by $\mathcal{Z}_{m,n} := 3^n\mathbb{Z}^d \cap \square_m$ and $\mathcal{Z}_n := 3^n\mathbb{Z}^d$. Then we have the following partition

$$\square_m = \bigsqcup_{z\in\mathcal{Z}_{m,n}} (z + \square_n), \tag{1.30}$$

which provides convenience to implement the renormalization.

For any finite set $\Lambda \subseteq \mathbb{Z}^d$, we denote by $|\Lambda|$ the number of vertices

$$|\Lambda| := \#\{x : x \in \Lambda\}, \tag{1.31}$$

and define the diameter as

$$\operatorname{diam}(\Lambda) := \max\{|x - y| : x, y \in \Lambda\}. \tag{1.32}$$

We also define $\partial\Lambda$ the boundary set of Λ that

$$\partial\Lambda := \{x \in \Lambda : \exists y \notin \Lambda, y \sim x\}, \tag{1.33}$$

and denote by Λ^- the interior of Λ

$$\Lambda^- := \Lambda \setminus \partial\Lambda. \tag{1.34}$$

Recall that the set of bonds of Λ is defined as Λ^* in (1.1). We define its enlarged version

$$\overline{\Lambda^*} := \{\{x,y\} : x \in \Lambda, y = x + e_i, i = 1, 2, \cdots, d\}, \tag{1.35}$$

where $e_i \in \mathbb{Z}^d$ is the i-th directed unit vector, and also denote by Λ^+ the vertices concerned in (1.35)

$$\Lambda^+ := \Lambda \cup \bigcup_{i=1}^{d} (\Lambda + e_i). \tag{1.36}$$

One motivation is that, for $n, m \in \mathbb{N}$ such that $n < m$, despite of (1.30), we observe that $\bigcup_{z \in \mathcal{Z}_{m,n}} (z + \square_n)^* \subsetneq (\square_m)^*$. On the other hand, (1.35) provides a better partition structure for bonds

$$\overline{(\square_m)^*} = \bigsqcup_{z \in \mathcal{Z}_{m,n}} \overline{(z + \square_n)^*}. \tag{1.37}$$

For disjoint sets $\Lambda, \Lambda' \subseteq \mathbb{Z}^d, \Lambda \cap \Lambda' = \emptyset$, we define $(\Lambda, \Lambda')^*$ as the set of bonds between Λ and Λ'

$$(\Lambda, \Lambda')^* := (\mathbb{Z}^d)^* \smallsetminus (\Lambda^* \sqcup (\Lambda')^*) \tag{1.38}$$

Especially, $(\Lambda, \Lambda^c)^*$ is the set of bonds connecting Λ and its complement.

1.4.2. *Probability spaces.* For every $\Lambda \subseteq \mathbb{Z}^d$, we denote by $\mathscr{F}_\Lambda$ the σ-algebra generate by the $(\eta_x)_{x \in \Lambda}$ and we write $\mathscr{F}$ short for $\mathscr{F}_{\mathbb{Z}^d}$. Given $\rho \in (0,1)$ as the density of particle, and make use of $\mathbb{P}_\rho$ as the Bernoulli product measure $\mathrm{Ber}(\rho)^{\otimes \mathbb{Z}^d}$ on $\mathcal{X}$, thus $(\mathcal{X}, \mathscr{F}, \mathbb{P}_\rho)$ is the triplet of probability space most used in this paper. For the expectation under $\mathbb{P}_\rho$, we use the notation $\langle \cdot \rangle_\rho$ or $\mathbb{E}_\rho[\cdot]$. We make use of $\mathbb{P}_{\rho,\Lambda}, \langle \cdot \rangle_{\rho,\Lambda}$ when we restrict our measure on $(\eta_x)_{x \in \Lambda}$. We also denote by $\mathbb{P}_{\Lambda,N,\zeta}$ and $\langle \cdot \rangle_{\Lambda,N,\zeta}$ for the probability and expectation under the canonical ensemble, i.e. N particles distributed uniformly on different sites of Λ with the configuration ζ on Λ^c. We usually omit ζ and just write them as $\mathbb{P}_{\Lambda,N}$ and $\langle \cdot \rangle_{\Lambda,N}$.

1.4.3. *Function spaces.* For every $1 \leqslant p \leqslant \infty$, we denote by $\|\cdot\|_{L^p}$ or $\|\cdot\|_p$ the L^p norm over the probability space $(\mathcal{X}, \mathscr{F}, \mathbb{P}_\rho)$, and denote by $L^p(\mathcal{X}, \mathscr{F}, \mathbb{P}_\rho)$ or shortly L^p the set of random variables with finite norm. For any $\Lambda \subseteq \mathbb{Z}^d$, let $\mathcal{F}_0(\Lambda)$ be the set of $\mathscr{F}_\Lambda$-measurable local functions. We also define the Sobolev norm $H^1(\Lambda)$ that

$$\|f\|_{H^1(\Lambda)}^2 = \left\langle f^2 \right\rangle_\rho + \sum_{b \in \Lambda^*} \left\langle (\pi_b f)^2 \right\rangle_\rho . \tag{1.39}$$

For every local functions, we can calculate its $H^1(\Lambda)$ norm, and we also use $H^1(\Lambda)$ to represent the set of functions with finite $H^1(\Lambda)$ norm. Despite of the natural definition of $\mathcal{F}_0(\Lambda)$, the function space $\mathcal{F}_0(\Lambda^-)$ is the proper analogue of the function space H_0^1 in the domain Λ. To see this, we can verify the following identity easily

$$\forall \Lambda \subseteq \Lambda' \subseteq \mathbb{Z}^d, f \in \mathcal{F}_0(\Lambda^-), \qquad \|f\|_{H^1(\Lambda)} = \|f\|_{H^1(\Lambda')} . \tag{1.40}$$

This is the important extension property of H_0^1 function, but a general $\mathcal{F}_0(\Lambda)$ function does not necessarily satisfy it. We will not use the notation $H_0^1(\Lambda)$ in the paragraphs for the conciseness of notation, while we keep in mind that $\mathcal{F}_0(\Lambda^-)$ plays the same role.

Viewing the discussion above, we define the space of harmonic functions with respect to the Kawasaki dynamics

$$\mathcal{A}(\Lambda) := \{u \in H^1(\Lambda) : \forall v \in \mathcal{F}_0(\Lambda^-), \langle v(-\mathcal{L}_\Lambda u) \rangle_\rho = 0\}. \tag{1.41}$$

Note that $u \in \mathcal{A}(\Lambda)$ does *not* imply that $u \in \mathcal{F}_0(\Lambda)$ and it can have dependence on the configuration outside Λ.

1.4.4. *Operators.* The translation operator, exchange operator and Kawasaki operator are respectively defined in (1.4), (1.5), (1.2) and (1.3). For $\eta \in \mathcal{X}$ and $\Lambda \subseteq \mathbb{Z}^d$, we define $(\eta \llcorner \Lambda)$ as the configuration restricted on Λ that

$$\forall x \in \mathbb{Z}^d, \qquad (\eta \llcorner \Lambda)_x := \eta_x \mathbf{1}_{\{x \in \Lambda\}}. \tag{1.42}$$

We sometimes identify $\eta \in \mathcal{X}$ as $\eta = \sum_{x \in \mathbb{Z}^d} \eta_x \delta_x$ for the convenience to manipulate.

The affine function defined by

$$\ell_p(\eta) := \sum_{x \in \mathbb{Z}^d} (p \cdot x)\, \eta_x, \tag{1.43}$$

is just a formal sum as there are infinite terms, while $\pi_b \ell_p$ is well-defined as

$$\forall b = \{x, y\} \in (\mathbb{Z}^d)^*, \qquad (\pi_b \ell_p)(\eta) := p \cdot (y - x)(\eta_x - \eta_y). \tag{1.44}$$

A rigorous version of (1.43) is a sum restricted on the finite set $\Lambda \subseteq \mathbb{Z}^d$

$$\ell_{p,\Lambda}(\eta) := \sum_{x \in \Lambda} (p \cdot x)\, \eta_x. \tag{1.45}$$

In Kawasaki dynamics, we define the tangent field along the direction e_i at x for $u : \mathcal{X} \to \mathbb{R}$ as

$$\nabla_{x,e_i} u := (\pi_{x,x+e_i} u)(\pi_{x,x+e_i} \ell_{e_i}). \tag{1.46}$$

Some simple calculation gives us

$$(\pi_{x,x+e_i} u)(\pi_{x,x+e_i} \ell_{e_i})(\eta) = \left(u(\eta^{x,x+e_i}) - u(\eta)\right)(\eta_x - \eta_{x+e_i}), \tag{1.47}$$

so the term is non-zero if and only if $(\eta_x, \eta_{x+e_i}) = (1,0)$ or $(\eta_x, \eta_{x+e_i}) = (0,1)$. Moreover, for both two non-zero cases, they evaluates the change that a particle jumps from x to $x + e_i$. Similarly, we define the gradient field of u at x as

$$\nabla_x u := (\nabla_{x,e_1} u, \nabla_{x,e_2} u, \cdots, \nabla_{x,e_d} u). \tag{1.48}$$

For every $p \in \mathbb{R}^d$, we also obtain that

$$p \cdot \nabla_x u = \sum_{i=1}^{d} (\pi_{x,x+e_i} u)(\pi_{x,x+e_i} \ell_p). \tag{1.49}$$

The Glauber operator appears naturally in some steps of analysis. We denote by η^x the flip operator at x that

$$(\eta^x)_z = \begin{cases} \eta_z, & z \neq x; \\ 1 - \eta_z, & z = x. \end{cases} \tag{1.50}$$

Then the Glauber derivative for $f : \mathcal{X} \to \mathbb{R}$ is defined by

$$\pi_x f := f(\eta^x) - f(\eta). \tag{1.51}$$

Clearly, $(\pi_x f)^2$ is independent of η_x.

1.4.5. *Constants.* We usually use C to represent a positive finite constant and $C(\cdots)$ to indicate its dependence with other parameters. The value of C may change from line to line. The following constants will be fixed and used throughout the paper.

- $d \in \mathbb{N}_+ = \{1, 2, 3, \cdots\}$ for the dimension of lattice.
- λ for the upper bound of the jump rate, i.e. $c_b(\eta) \leqslant \lambda$ for any $b \in (\mathbb{Z}^d)^*, \eta \in \mathcal{X}$.
- $\mathbf{r}$ for the radius of dependence of jump rate as indicated in Hypothesis 1.1.

2. Analytic tools

In this part, we collect all the necessary analytic tools in this paper. The two main results of this section are the modified Caccioppoli inequality (Proposition 2.6), and the weighted multiscale Poincaré inequality (Proposition 2.12 and then Proposition 3.9).

2.1. Analytic tools on Kawasaki dynamics.

2.1.1. *Glauber operator meets Kawasaki operator.* Our first inequality comes from the observation in [48, eq.(B.2)], which states that we can exchange the site of the Glauber derivative by paying the error of the Kawasaki operator.

Lemma 2.1. *Recall the L^2 function space defined in Section 1.4.3, then we have*

$$\|\pi_x f\|_{L^2} \leqslant \|\pi_y f\|_{L^2} + \frac{1}{\sqrt{2\chi(\rho)}} \|\pi_{x,y} f\|_{L^2} . \tag{2.1}$$

Proof. As we know, $(\pi_x f)^2$ does not depend on η_x, thus we decompose $\|\pi_x f\|_{L^2}$ with respect to the state of η_y

$$\|\pi_x f\|_{L^2} = \left\langle (\pi_x f)^2 \mathbf{1}_{\{\eta_y=1\}} + (\pi_x f)^2 \mathbf{1}_{\{\eta_y=0\}} \right\rangle_\rho^{\frac{1}{2}} .$$

We denote by $\widetilde{\eta} = (\eta_z)_{z\in\mathbb{Z}^d\setminus\{x,y\}}$ and $F(\eta_x,\eta_y,\widetilde{\eta}) = f(\eta)$. Then we have

$$\begin{aligned}
&\|\pi_x f\|_{L^2}\\
&\quad= \left(\int_{\mathcal{X}} p(F(1,1,\widetilde{\eta}) - F(0,1,\widetilde{\eta}))^2 + (1-p)(F(1,0,\widetilde{\eta}) - F(0,0,\widetilde{\eta}))^2 \, \mathrm{d}\mathbb{P}_\rho(\widetilde{\eta})\right)^{\frac{1}{2}} .
\end{aligned}$$

We apply the triangle inequality for this norm. The trick is that we only replace the term involving $\eta_x \neq \eta_y$. For example, in the terms $(F(1,1,\widetilde{\eta}) - F(0,1,\widetilde{\eta}))^2$, we replace $F(0,1,\widetilde{\eta})$ by $F(1,0,\widetilde{\eta})$. This follows exactly the spirit of the Kawasaki operator $\pi_{x,y}$ and we obtain

$$\begin{aligned}
&\|\pi_x f\|_{L^2}\\
&\leqslant \left(\int_{\mathcal{X}} p(F(1,1,\widetilde{\eta}) - F(1,0,\widetilde{\eta}))^2 + (1-p)(F(0,1,\widetilde{\eta}) - F(0,0,\widetilde{\eta}))^2 \, \mathrm{d}\mathbb{P}_\rho(\widetilde{\eta})\right)^{\frac{1}{2}}\\
&\qquad + \left(\int_{\mathcal{X}} p(F(0,1,\widetilde{\eta}) - F(1,0,\widetilde{\eta}))^2 + (1-p)(F(1,0,\widetilde{\eta}) - F(0,1,\widetilde{\eta}))^2 \, \mathrm{d}\mathbb{P}_\rho(\widetilde{\eta})\right)^{\frac{1}{2}}\\
&= \left(\int_{\mathcal{X}} p(F(1,1,\widetilde{\eta}) - F(1,0,\widetilde{\eta}))^2 + (1-p)(F(0,1,\widetilde{\eta}) - F(0,0,\widetilde{\eta}))^2 \, \mathrm{d}\mathbb{P}_\rho(\widetilde{\eta})\right)^{\frac{1}{2}}\\
&\qquad + \left(\int_{\mathcal{X}} (F(1,0,\widetilde{\eta}) - F(0,1,\widetilde{\eta}))^2 \, \mathrm{d}\mathbb{P}_\rho(\widetilde{\eta})\right)^{\frac{1}{2}} .
\end{aligned}$$

We notice the identity in the last equation

$$\|\pi_y f\|_{L^2} = \left(\int_{\mathcal{X}} p(F(1,1,\widetilde{\eta}) - F(1,0,\widetilde{\eta}))^2 + (1-p)(F(0,1,\widetilde{\eta}) - F(0,0,\widetilde{\eta}))^2 \, \mathrm{d}\mathbb{P}_\rho(\widetilde{\eta})\right)^{\frac{1}{2}} ,$$

$$\|\pi_{x,y} f\|_{L^2} = \left(\int_{\mathcal{X}} 2p(1-p)(F(1,0,\widetilde{\eta}) - F(0,1,\widetilde{\eta}))^2 \, \mathrm{d}\mathbb{P}_\rho(\widetilde{\eta})\right)^{\frac{1}{2}} ,$$

then we conclude the desired result. □

2.1.2. *Spectral inequality.* The spectral inequality is an important tool to analyze Markov processes. In this part, we resume several spectral inequalities from the literature.

The most used spectral inequality is the one for independent random variables known as Efron–Stein inequality. Here we state it and also its reverse version.

Lemma 2.2 (Efron–Stein inequality). *Let $(X_i)_{1\leqslant i\leqslant n}$ be i.i.d. random variables taking value on E, and we denote by*

$$\mathbb{E}_{(i)}[\cdot] := \int_E (\cdot)\,\mathrm{d}\mathbb{P}_{X_i}, \qquad \mathbb{E}_{(-i)}[\cdot] := \int_{E^{n-1}} (\cdot) \prod_{1\leqslant j\leqslant n, j\neq i} \mathrm{d}\mathbb{P}_{X_j}, \tag{2.2}$$

and $\mathbb{V}\mathrm{ar}_{(i)}, \mathbb{V}\mathrm{ar}_{(-i)}$ for the corresponding variances. Then for a random variable $f(X_1, X_2, \cdots, X_n)$, we have

$$\sum_{i=1}^n \mathbb{V}\mathrm{ar}_{(i)}[\mathbb{E}_{(-i)}[f]] \leqslant \mathbb{V}\mathrm{ar}[f] \leqslant \sum_{i=1}^n \mathbb{E}_{(-i)}[\mathbb{V}\mathrm{ar}_{(i)}[f]]. \tag{2.3}$$

Proof. The upper bound is the classical Efron–Stein inequality, and one can find its proof in [18, Theorem 3.1]. The lower bound, which could be seen as a reverse Efron–Stein inequality, is less well-known, but follows exactly the same strategy of proof. The authors learn the lower bound at first in [73]. □

A direct corollary of Efron–Stein inequality is the spectral inequality of the Glauber operator (1.50) under product Bernoulli measure.

Corollary 2.3 (Spectral inequality for Glauber dynamics). *For any $\Lambda \subseteq \mathbb{Z}^d$, we have*

$$\mathbb{V}\mathrm{ar}_{\rho,\Lambda}[f] \leqslant \chi(\rho) \sum_{x\in\Lambda} \left\langle (\pi_x f)^2 \right\rangle_{\rho,\Lambda}. \tag{2.4}$$

Proof. We apply the classical Efron–Stein inequality and the upper bound is the right-hand side of (2.4). □

With some more treatment, we can also obtain the spectral inequality for the Kawasaki operator (1.3) under product Bernoulli measure.

Lemma 2.4. *For any bounded set $\Lambda \subseteq \mathbb{Z}^d$ and $f \in \mathcal{F}_0(\Lambda^-)$, we have*

$$\mathbb{V}\mathrm{ar}_\rho[f] \leqslant \mathrm{diam}(\Lambda)^2 \sum_{b\in\Lambda^*} \left\langle (\pi_b f)^2 \right\rangle_\rho. \tag{2.5}$$

Proof. We apply at first the spectral inequality for Glauber dynamics

$$\mathbb{V}\mathrm{ar}_\rho[f] \leqslant \chi(\rho) \sum_{x\in\Lambda} \left\langle (\pi_x f)^2 \right\rangle_\rho. \tag{2.6}$$

We fix a direction in the canonical basis e_i, then for every $x \in \Lambda$, there exists a positive integer ℓ depending on x,

$$\ell(x) := \min\{k \in \mathbb{N}_+ : x + ke_i \in \partial\Lambda\}.$$

Then we apply Lemma 2.1

$$\|\pi_x f\|_{L^2} \leqslant \|\pi_{x+e_i} f\|_{L^2} + \frac{1}{\sqrt{2\chi(\rho)}} \|\pi_{x,x+e_i} f\|_{L^2}.$$

We sum this inequality along the path $x \to x + e_i \to x + 2e_i \cdots \to x + \ell(x)e_i$

$$\|\pi_x f\|_{L^2} \leqslant \left\|\pi_{x+\ell(x)e_i} f\right\|_{L^2} + \frac{1}{\sqrt{2\chi(\rho)}} \sum_{j=1}^{\ell(x)} \left\|\pi_{x+(j-1)e_i, x+je_i} f\right\|_{L^2}.$$

Notice that $x + \ell(x)e_i \in \partial\Lambda$ and $f \in \mathcal{F}_0(\Lambda^-)$, so f does not depend on $\eta_{x+\ell(x)e_i}$ and $\pi_{x+\ell(x)e_i} f = 0$. This implies

$$\|\pi_x f\|_{L^2} \leqslant \frac{1}{\sqrt{2\chi(\rho)}} \sum_{j=1}^{\ell(x)} \left\|\pi_{x+(j-1)e_i, x+je_i} f\right\|_{L^2}. \tag{2.7}$$

We put (2.7) back to (2.6) and apply Cauchy–Schwarz inequality

$$\begin{aligned}\mathbb{V}\mathrm{ar}_\rho[f] &\leqslant \frac{1}{2}\sum_{x\in\Lambda} \ell(x) \sum_{j=1}^{\ell(x)} \left\langle (\pi_{x+(j-1)e_i, x+je_i} f)^2 \right\rangle_\rho \\ &\leqslant \frac{1}{2}\operatorname{diam}(\Lambda) \sum_{x\in\Lambda} \sum_{j=1}^{\ell(x)} \left\langle (\pi_{x+(j-1)e_i, x+je_i} f)^2 \right\rangle_\rho.\end{aligned}$$

Here the factor $\chi(\rho)$ in (2.7) and (2.6) compensates, and we also make use of the fact $\ell(x) \leqslant \operatorname{diam}(\Lambda)$. We now exchange the order of the sum $\sum_{x\in\Lambda}\sum_{j=1}^{\ell(x)}$

$$\mathbb{V}\mathrm{ar}_\rho[f] \leqslant \frac{1}{2}\operatorname{diam}(\Lambda) \sum_{b\in\Lambda^*} \sum_{x\in\Lambda : \exists j\in\mathbb{N}_+, x+je_i\in b} \left\langle (\pi_b f)^2 \right\rangle_\rho.$$

Because every bond b can be counted at most $\operatorname{diam}(\Lambda)$ times along the direction e_i, we obtain the desired result. □

In [58, Theorem 1], Lu and Yau proved a generalized version of the spectral inequality for the Glauber dynamics. We do not need that one in this paper, but we will make use of [58, Theorem 2], the spectral inequality for Kawasaki dynamics under canonical ensemble.

Lemma 2.5 (Theorem 2, [58])**.** *There exists a positive constant $C = C(d)$, such that for any $L \in \mathbb{N}_+$ and any $N \in \mathbb{N}_+, N \leqslant |\Lambda_L|$, we have*

$$\mathbb{V}\mathrm{ar}_{\Lambda_L,N}[f] \leqslant CL^2 \sum_{b\in(\Lambda_L)^*} \left\langle (\pi_b f)^2 \right\rangle_{\Lambda_L,N}. \tag{2.8}$$

2.1.3. *Modified Caccioppoli inequality.* The modified Caccioppoli inequality is a key input to gain the convergence rate in the interacting particle systems. It is at first proved in [35, Proposition 3.9] and here we present its version in Kawasaki dynamics. The conditional expectation operator will be used in the following paragraphs. For $\Lambda \subseteq \mathbb{Z}^d$ and $f \in L^1$, we define

$$\mathsf{A}_\Lambda f := \mathbb{E}_\rho\left[f \,|\mathscr{F}_\Lambda\right]. \tag{2.9}$$

Concretely, it is calculated as

$$\mathsf{A}_\Lambda f(\eta) = \int_{\mathcal{X}} f(\eta \llcorner \Lambda + \eta' \llcorner \Lambda^c)\, \mathrm{d}\mathbb{P}_\rho(\eta').$$

We usually denote by $\mathsf{A}_L f \equiv \mathsf{A}_{\Lambda_L} f$ for short.

Proposition 2.6 (Modified Caccioppoli inequality)**.** *There exist $\theta(d,\lambda) \in (0,1)$, finite positive constants $C(d,\lambda)$, and $R_0(d,\lambda,\mathbf{r})$ such that for every $L \geqslant R_0$ and $u \in \mathcal{A}(\Lambda_{3L})$ (defined in (1.41)), we have*

$$\frac{1}{|\Lambda_L|}\left\langle \mathsf{A}_{L+2\mathbf{r}} u(-\mathcal{L}_{\Lambda_L} \mathsf{A}_{L+2\mathbf{r}} u)\right\rangle_\rho \leqslant \frac{CL^{-2}}{|\Lambda_{3L}|}\left\langle u^2\right\rangle_\rho + \frac{\theta}{|\Lambda_{3L}|}\left\langle u(-\mathcal{L}_{\Lambda_{3L}} u)\right\rangle_\rho. \tag{2.10}$$

Its proof is similar to [35, Proposition 3.9], which can be summarized as following three steps.

(1) Test the harmonic function $u \in \mathcal{A}(\Lambda_{3L})$ with its cutoff version $\mathsf{A}_L u$.

(2) Obtain the L^2-term using the quadratic variation of the martingale $(\mathsf{A}_n u)_{n\in\mathbb{N}_+}$.
(3) Bootstrap the result from its weak version with a correct normalization factor of volume.

The first and second step can be seen as the main difference between particle system and PDE setting, where we make the cutoff in order to reduce the influence of particles from the boundary, and we also need the nice L^2-isometry of martingale to recover the L^2 term of u.

In Kawasaki dynamics, the quadratic variation structure is less obvious compared to the setting of continuous configuration space in [35]. Let us make some explicit calculation at first. It is clear that

$$\begin{aligned} &\forall b \in (\Lambda_n)^*, \qquad \pi_b \mathsf{A}_n f = \mathsf{A}_n \pi_b f, \\ &\forall b \in (\Lambda_n^c)^*, \qquad \pi_b \mathsf{A}_n f = 0. \end{aligned} \tag{2.11}$$

That is to say, the operators A_n and π_b are commutative when the bond b stays in $(\Lambda_n)^*$, and the influence is 0 when b is outside Λ_n. When $b \in (\Lambda_n, \Lambda_n^c)^*$, the situation is subtle as we will see the perturbation

$$\forall x \in \partial\Lambda_n, y \notin \Lambda_n, y \sim x, \qquad \pi_{x,y}\mathsf{A}_n f(\eta) = \mathsf{A}_n f(\eta^{x,y}) - \mathsf{A}_n f(\eta).$$

Since we apply the conditional expectation, the information of η_x is no longer useful in $\mathsf{A}_n f(\eta^{x,y})$ and we have

$$\mathsf{A}_n f(\eta^{x,y}) = \mathsf{A}_n f(\eta \llcorner (\Lambda_n \setminus \{x\}) + \eta_y \delta_x).$$

Thus, near the boundary the Kawasaki operator is like resampling the state at x and we have

$$\forall x \in \partial\Lambda_n, y \notin \Lambda_n, y \sim x, \qquad \left\langle (\pi_{x,y}\mathsf{A}_n f)^2 \right\rangle_\rho = 2\chi(\rho) \left\langle (\pi_x \mathsf{A}_n f)^2 \right\rangle_\rho . \tag{2.12}$$

By the spectral inequality (2.4), we know that

$$\chi(\rho) \sum_{x\in\partial\Lambda_n} \left\langle (\pi_x \mathsf{A}_n f)^2 \right\rangle_\rho \geqslant \left\langle (\mathsf{A}_n f - \mathsf{A}_{n-2} f)^2 \right\rangle_\rho .$$

The inequality is not on the desired direction, because we hope to give an upper bound for the boundary perturbation. For this reason, we would like to study how to control this Glauber derivative near the boundary at first.

Lemma 2.7. *For $\mathsf{A}_n \equiv \mathsf{A}_{\Lambda_n}$ defined in (2.9) and $f \in L^2$, the following estimate holds*

$$\sum_{b\in(\Lambda_n,\Lambda_n^c)^*} \left\langle (\pi_b \mathsf{A}_n f)^2 \right\rangle_\rho \leqslant 4 \left(\sum_{b\in(\Lambda_n,\Lambda_n^c)^*} \left\langle (\mathsf{A}_{n+2}\pi_b f)^2 \right\rangle_\rho + \left\langle (\mathsf{A}_{n+2} f - \mathsf{A}_n f)^2 \right\rangle_\rho \right). \tag{2.13}$$

Proof. The left-hand side can be expressed with (2.12)

$$\begin{aligned} \sum_{b\in(\Lambda_n,\Lambda_n^c)^*} \left\langle (\pi_b \mathsf{A}_n f)^2 \right\rangle_\rho &= \sum_{x\in\partial\Lambda_n, y\notin\Lambda_n, y\sim x} \left\langle (\pi_{x,y}\mathsf{A}_n f)^2 \right\rangle_\rho \\ &= \sum_{x\in\partial\Lambda_n, y\notin\Lambda_n, y\sim x} 2\chi(\rho) \left\langle (\pi_x \mathsf{A}_n f)^2 \right\rangle_\rho . \end{aligned} \tag{2.14}$$

We notice that

$$\begin{aligned} \pi_x \mathsf{A}_n f &= \mathsf{A}_n f(\eta^x) - \mathsf{A}_n f(\eta) \\ &= \int_{\{0,1\}} \left(\mathsf{A}_{\Lambda_n\sqcup\{y\}} f(\eta^x) - \mathsf{A}_{\Lambda_n\sqcup\{y\}} f(\eta) \right) \mathrm{d}\mathbb{P}_\rho(\eta_y) \\ &= \int_{\{0,1\}} \pi_x \mathsf{A}_{\Lambda_n\sqcup\{y\}} f \,\mathrm{d}\mathbb{P}_\rho(\eta_y). \end{aligned} \tag{2.15}$$

Using Jensen's inequality, we have

$$\left\langle (\pi_x \mathsf{A}_n f)^2 \right\rangle_\rho \leqslant \left\langle \left(\pi_x \mathsf{A}_{\Lambda_n \sqcup \{y\}} f\right)^2 \right\rangle_\rho \tag{2.16}$$

Then we apply Lemma 2.1 to $\mathsf{A}_{\Lambda_n \sqcup \{y\}} f$ and obtain that

$$\left\| \pi_x \mathsf{A}_{\Lambda_n \sqcup \{y\}} f \right\|_{L^2} \leqslant \left\| \pi_y \mathsf{A}_{\Lambda_n \sqcup \{y\}} f \right\|_{L^2} + \frac{1}{\sqrt{2\chi(\rho)}} \left\| \pi_{x,y} \mathsf{A}_{\Lambda_n \sqcup \{y\}} f \right\|_{L^2}. \tag{2.17}$$

We put (2.17) and (2.16) back to (2.14), and obtain that

$$\begin{aligned} &\sum_{b \in (\Lambda_n, \Lambda_n^c)^*} \left\langle (\pi_b \mathsf{A}_n f)^2 \right\rangle_\rho \\ &\qquad \leqslant \sum_{x \in \partial \Lambda_n, y \notin \Lambda_n, y \sim x} 4\chi(\rho) \left(\left\langle \left(\pi_y \mathsf{A}_{\Lambda_n \sqcup \{y\}} f\right)^2 \right\rangle_\rho + \frac{1}{2\chi(\rho)} \left\langle \left(\pi_{x,y} \mathsf{A}_{\Lambda_n \sqcup \{y\}} f\right)^2 \right\rangle_\rho \right). \end{aligned} \tag{2.18}$$

For the first term on right-hand side, it is exactly the fluctuation on $\partial \Lambda_{n+2}$, so we apply the reverse Efron–Stein inequality (the first inequality of (2.2)) to $\mathsf{A}_{n+2} f$ under the expectation over the $\{\eta_y\}_{y \in \partial \Lambda_{n+2}}$

$$\begin{aligned} \sum_{y \in \partial \Lambda_{n+2}} \chi(\rho) \left\langle \left(\pi_y \mathsf{A}_{\Lambda_n \sqcup \{y\}} f\right)^2 \right\rangle_{\rho, \partial \Lambda_{n+2}} &= \sum_{y \in \partial \Lambda_{n+2}} \mathbb{V}\mathrm{ar}_{\rho, \partial \Lambda_{n+2}} [\mathbb{E}_{\rho, \partial \Lambda_{n+2}} [\mathsf{A}_{n+2} f \,|\eta_y]] \\ &\leqslant \mathbb{V}\mathrm{ar}_{\rho, \partial \Lambda_{n+2}} [\mathsf{A}_{n+2} f] \\ &= \left\langle (\mathsf{A}_{n+2} f - \mathsf{A}_n f)^2 \right\rangle_{\rho, \partial \Lambda_{n+2}}. \end{aligned}$$

Recall that $\langle \cdot \rangle_{\rho, \partial \Lambda_{n+2}}$ is defined in Section 1.4.2, and note that $\chi(\rho)$ also appears similarly in (2.4) in a reversed inequality. We also use the identity $\mathbb{E}_{\rho, \partial \Lambda_{n+2}} [\mathsf{A}_{n+2} f \,|\eta_y] = \mathsf{A}_{\Lambda_n \sqcup \{y\}} f$ here. Then we take the expectation of other variables to yield the estimate of the first term on the right-hand sideof (2.18)

$$\sum_{x \in \partial \Lambda_n, y \notin \Lambda_n, y \sim x} 4\chi(\rho) \left\langle \left(\pi_y \mathsf{A}_{\Lambda_n \sqcup \{y\}} f\right)^2 \right\rangle_\rho \leqslant 4 \left\langle (\mathsf{A}_{n+2} f - \mathsf{A}_n f)^2 \right\rangle_\rho.$$

For the second term on the right-hand sideof (2.18), we apply (2.11) and once again Jensen's inequality to obtain that

$$\begin{aligned} \sum_{x \in \partial \Lambda_n, y \notin \Lambda_n, y \sim x} \left\langle \left(\pi_{x,y} \mathsf{A}_{\Lambda_n \sqcup \{y\}} f\right)^2 \right\rangle_\rho &= \sum_{x \in \partial \Lambda_n, y \notin \Lambda_n, y \sim x} \left\langle \left(\mathsf{A}_{\Lambda_n \sqcup \{y\}} \pi_{x,y} f\right)^2 \right\rangle_\rho \\ &\leqslant \sum_{b \in (\Lambda_n, \Lambda_n^c)^*} \left\langle (\mathsf{A}_{n+2} \pi_b f)^2 \right\rangle_\rho. \end{aligned}$$

This concludes the desired result. □

Once we develop Lemma 2.7, the rest of the proof follows that in [35, Proposition 3.9].

Proof of Proposition 2.6. The proof can be divided into three steps.

Step 1: construction of test function. In the first step, we do some preparation. Our object is to regularize u such that it becomes a function in $\mathcal{F}_0(\Lambda_{3L}^-)$. A very natural idea is to apply the conditional expectation operator (2.9), then the information outside $3L$ will be averaged. In order to make this cutoff more smooth, we propose the following regularized version

$$\mathsf{A}_{s,\ell} f := \frac{1}{\ell} \int_0^\ell \mathsf{A}_{s+t} f \, \mathrm{d}t. \tag{2.19}$$

Recall Λ_L is defined for $L \in \mathbb{R}_+$, so A_{s+t} does for $s, t \in \mathbb{R}_+$. Its Kawasaki derivative can be calculated using (2.11)

$$\pi_b \mathsf{A}_{s,\ell} f = \begin{cases} \mathsf{A}_{s,\ell}\pi_b f & \text{if } b \in (\Lambda_s)^*; \\ \frac{1}{\ell}\int_{\tau(b)+}^{s+\ell} \mathsf{A}_t \pi_b f \,\mathrm{d}t + \frac{(\tau_b - s)\wedge 1}{\ell}\pi_b \mathsf{A}_{\tau_b} f \mathbf{1}_{b\in\left(\Lambda_{\tau(b)}, \Lambda^c_{\tau(b)}\right)^*} & \text{if } b \in \left(\Lambda_s, \Lambda^c_{s+\ell}\right)^*; \\ 0 & \text{if } b \in (\Lambda^c_{s+\ell})^*. \end{cases} \tag{2.20}$$

Here the notation $\tau(b)$ is defined as

$$\tau(b) := \inf\{s \in \mathbb{R}_+ : b \in (\Lambda_s)^*\}. \tag{2.21}$$

From the definition of hypercube in Section 1.4.1, we know $b \in \Lambda_{\tau(b)+}$ but $b \notin \Lambda_{\tau(b)}$.

We will also make use of the following operator

$$\widetilde{\mathsf{A}}_{s,\ell} f := (\mathsf{A}_{s,\ell} \circ \mathsf{A}_{s,\ell})(f) = \frac{2}{\ell^2}\int_0^\ell (\ell - t)\mathsf{A}_{s+t} f \,\mathrm{d}t, \tag{2.22}$$

The motivation comes from the following identity that

$$\left\langle (\mathsf{A}_{s,\ell} f)^2 \right\rangle_\rho = \left\langle f(\widetilde{\mathsf{A}}_{s,\ell} f)\right\rangle_\rho = \frac{2}{\ell^2}\int_0^\ell (\ell - t)\left\langle (\mathsf{A}_{s+t} f)^2\right\rangle_\rho \mathrm{d}t. \tag{2.23}$$

Similar to (2.20), we also calculate its Kawasaki derivative as preparation

$$\pi_b \widetilde{\mathsf{A}}_{s,\ell} f = \tag{2.24}$$
$$\begin{cases} \mathsf{A}_{s,\ell}\pi_b f & \text{if } b \in (\Lambda_s)^*; \\ \frac{2}{\ell^2}\int_{\tau(b)+}^{s+\ell}(s+\ell-t)\mathsf{A}_t \pi_b f \,\mathrm{d}t & \\ \qquad + \frac{2}{\ell^2}\int_{(\tau(b)-2)\vee s}^{\tau(b)}(s+\ell-t)\pi_b \mathsf{A}_{\tau(b)} f \mathbf{1}_{b\in\left(\Lambda_{\tau(b)}, \Lambda^c_{\tau(b)}\right)^*} & \text{if } b \in \left(\Lambda_s, \Lambda^c_{s+\ell}\right)^*; \\ 0 & \text{if } b \in (\Lambda^c_{s+\ell})^*. \end{cases}$$

Step 2: week Caccioppoli inequality. We then prove the weak Caccioppoli inequality at first. Fix $\theta'(\lambda) := \frac{10\lambda}{1+10\lambda} \in (0,1)$; recall that $\lambda \geqslant 1$ is the constant in Hypothesis 1.1. For every $L > 0$, $s \geqslant L + 2\mathbf{r}, \ell > 1, s + \ell < 3L$ and $u \in \mathcal{A}(\Lambda_{s+\ell+2})$, we claim that

$$\begin{aligned} \ell^{-2}\left\langle (\mathsf{A}_s u)^2\right\rangle_\rho + \left\langle \mathsf{A}_{s,\ell} u(-\mathcal{L}_{\Lambda_L}\mathsf{A}_{s,\ell} u)\right\rangle_\rho & \\ \leqslant \theta'\left(\ell^{-2}\left\langle (\mathsf{A}_{s+\ell} u)^2\right\rangle_\rho + \langle u(-\mathcal{L}_{\Lambda_{s+\ell}} u)\rangle_\rho\right). & \end{aligned} \tag{2.25}$$

The main idea is to use the conditional expectation $\widetilde{\mathsf{A}}_{s,\ell} u$ given in (2.22), because it provides a cutoff that $\widetilde{\mathsf{A}}_{s,\ell} u \in \mathcal{F}_0(\Lambda^-_{s+\ell+2})$. Then we test it with u

$$0 = \left\langle \widetilde{\mathsf{A}}_{s,\ell} u(-\mathcal{L}_{s+\ell+2} u)\right\rangle_\rho = \sum_{b\in(\Lambda_{s+\ell+2})^*}\left\langle c_b(\pi_b \widetilde{\mathsf{A}}_{s,\ell} u)(\pi_b u)\right\rangle_\rho = \mathbf{I} + \mathbf{II} + \mathbf{III}.$$

Here decompose the right-hand side into the sum of three terms

$$\begin{aligned} \mathbf{I} &:= \sum_{b\in(\Lambda_{s-2\mathbf{r}})^*}\left\langle c_b(\pi_b \widetilde{\mathsf{A}}_{s,\ell} u)(\pi_b u)\right\rangle_\rho, \\ \mathbf{II} &:= \sum_{b\in(\Lambda_s)^*\setminus(\Lambda_{s-2\mathbf{r}})^*}\left\langle c_b(\pi_b \widetilde{\mathsf{A}}_{s,\ell} u)(\pi_b u)\right\rangle_\rho, \\ \mathbf{III} &:= \sum_{b\in(\Lambda_{s+\ell})^*\setminus(\Lambda_s)^*}\left\langle c_b(\pi_b \widetilde{\mathsf{A}}_{s,\ell} u)(\pi_b u)\right\rangle_\rho, \end{aligned} \tag{2.26}$$

and we have the estimate

$$|\mathbf{I}| \leqslant |\mathbf{II}| + |\mathbf{III}|. \tag{2.27}$$

The term **I** is the easiest one to treat

$$
\begin{aligned}
\mathbf{I} &= \sum_{b\in(\Lambda_{s-2\mathbf{r}})^*} \left\langle c_b(\pi_b \widetilde{\mathsf{A}}_{s,\ell} u)(\pi_b u)\right\rangle_\rho \\
&= \sum_{b\in(\Lambda_{s-2\mathbf{r}})^*} \left\langle (\pi_b \mathsf{A}_{s,\ell} u)\mathsf{A}_{s,\ell}(c_b \pi_b u)\right\rangle_\rho \\
&= \sum_{b\in(\Lambda_{s-2\mathbf{r}})^*} \left\langle c_b(\pi_b \mathsf{A}_{s,\ell} u)^2\right\rangle_\rho .
\end{aligned} \tag{2.28}
$$

From the first line to the second line, we use the fact $\widetilde{\mathsf{A}}_{s,\ell} = \mathsf{A}_{s,\ell} \circ \mathsf{A}_{s,\ell}$ and the reversibility of $\mathsf{A}_{s,\ell}$. From the second line to the third line, we use the fact c_b is $\mathscr{F}_{\Lambda_s}$-measurable when $b \in (\Lambda_{s-2\mathbf{r}})^*$ and $\mathsf{A}_{s,\ell}\pi_b = \pi_b \mathsf{A}_{s,\ell}$ from (2.24).

The identity (2.28) does not apply directly to **II**, because for $b \in (\Lambda_s)^* \setminus (\Lambda_{s-2\mathbf{r}})^*$, the jump rate c_b is no longer $\mathscr{F}_{\Lambda_s}$-measurable. Therefore, we make use of the exact expression (2.22)

$$
\begin{aligned}
|\mathbf{II}| &= \sum_{b\in(\Lambda_s)^*\setminus(\Lambda_{s-2\mathbf{r}})^*} \left| \frac{2}{\ell^2}\int_0^\ell (\ell - t)\left\langle c_b(\pi_b \mathsf{A}_{s+t} u)(\pi_b u)\right\rangle_\rho \,\mathrm{d}t \right| \\
&\leqslant \sum_{b\in(\Lambda_s)^*\setminus(\Lambda_{s-2\mathbf{r}})^*} \frac{\lambda}{\ell^2}\int_0^\ell (\ell - t)\left\langle (\pi_b \mathsf{A}_{s+t} u)^2 + (\pi_b u)^2\right\rangle_\rho \,\mathrm{d}t \\
&\leqslant \sum_{b\in(\Lambda_s)^*\setminus(\Lambda_{s-2\mathbf{r}})^*} \frac{2\lambda}{\ell^2}\int_0^\ell (\ell - t)\left\langle (\pi_b u)^2\right\rangle_\rho \,\mathrm{d}t \\
&= \sum_{b\in(\Lambda_s)^*\setminus(\Lambda_{s-2\mathbf{r}})^*} \lambda\left\langle (\pi_b u)^2\right\rangle_\rho .
\end{aligned} \tag{2.29}
$$

From the first line to the second line, we make use of Young's inequality and $c_b \leqslant \lambda$. From the second line to the third line, $\mathsf{A}_{s,\ell}\pi_b = \pi_b \mathsf{A}_{s,\ell}$ and $\left\langle (\mathsf{A}_{s+t}\pi_b u)^2\right\rangle_\rho \leqslant \left\langle (\pi_b u)^2\right\rangle_\rho$ is also applied thanks to Jensen's inequality.

The term **III** has two integrals following (2.24), which can be noted respectively by **III.1** and **III.2**. The first part is similar to (2.29)

$$
\begin{aligned}
|\mathbf{III.1}| &\leqslant \sum_{b\in(\Lambda_{s+\ell})^*\setminus(\Lambda_s)^*} \left| \frac{2}{\ell^2}\int_{\tau(b)+}^{s+\ell} (s+\ell - t)\left\langle c_b(\mathsf{A}_t \pi_b u)(\pi_b u)\right\rangle_\rho \,\mathrm{d}t \right| \\
&\leqslant \sum_{b\in(\Lambda_{s+\ell})^*\setminus(\Lambda_s)^*} \lambda\left\langle (\pi_b u)^2\right\rangle_\rho .
\end{aligned} \tag{2.30}
$$

The second part is the key to make appear the L^2 term

$$
\begin{aligned}
|\mathbf{III.2}| &\leqslant \sum_{b\in(\Lambda_{s+\ell})^*\setminus(\Lambda_s)^*} \left| \frac{2}{\ell^2}\int_{(\tau(b)-2)\vee s}^{\tau(b)} (s+\ell - t)\left\langle c_b(\pi_b \mathsf{A}_{\tau(b)} u)(\pi_b u)\right\rangle_\rho \mathbf{1}_{b\in\left(\Lambda_{\tau(b)},\Lambda^c_{\tau(b)}\right)^*} \,\mathrm{d}t \right| \\
&\leqslant \sum_{b\in(\Lambda_{s+\ell})^*\setminus(\Lambda_s)^*} \frac{2\lambda}{\ell}\left\langle \gamma^{-1}(\pi_b \mathsf{A}_{\tau(b)} u)^2 + \gamma(\pi_b u)^2\right\rangle_\rho \mathbf{1}_{b\in\left(\Lambda_{\tau(b)},\Lambda^c_{\tau(b)}\right)^*}.
\end{aligned}
$$

From the first line to the second line, Young's inequality is applied with $\gamma > 0$ to be fixed, together with the trivial bound $(s+\ell-t) \leqslant \ell$ for t defied above. We rearrange the sum, and obtain

$$
\begin{aligned}
|\mathbf{III.2}| &\leqslant \sum_{n=0}^{\lfloor \ell \rfloor - 2} \sum_{b \in (\Lambda_n, \Lambda_n^c)^*} \frac{2\lambda}{\ell} \left\langle \gamma^{-1}(\pi_b \mathsf{A}_n u)^2 + \gamma(\pi_b u)^2 \right\rangle_\rho \\
&\leqslant \sum_{n=0}^{\lfloor \ell \rfloor - 2} \left(\frac{8\lambda}{\gamma \ell} \left\langle (\mathsf{A}_{s+n+2} u - \mathsf{A}_{s+n} u)^2 \right\rangle_\rho + \sum_{b \in (\Lambda_n, \Lambda_n^c)^*} \left(\frac{8\lambda}{\gamma \ell} + \frac{2\gamma\lambda}{\ell} \right) \left\langle (\pi_b u)^2 \right\rangle_\rho \right) \qquad (2.31) \\
&= \frac{8\lambda}{\gamma \ell} \left\langle (\mathsf{A}_{s+\ell} u - \mathsf{A}_s u)^2 \right\rangle_\rho + \sum_{n=0}^{\lfloor \ell \rfloor - 2} \sum_{b \in (\Lambda_n, \Lambda_n^c)^*} \left(\frac{8\lambda}{\gamma \ell} + \frac{2\gamma\lambda}{\ell} \right) \left\langle (\pi_b u)^2 \right\rangle_\rho .
\end{aligned}
$$

Here we insert the estimate Lemma 2.7 in the second line to handle the perturbation of boundary term, and then make use of the orthogonal decomposition of martingale from the second line to the third line. We choose $\gamma = \ell$, and put (2.28), (2.29), (2.30) and (2.31) back to (2.27), which concludes that

$$
\begin{aligned}
&\sum_{b \in (\Lambda_{s-2\mathbf{r}})^*} \left\langle c_b (\pi_b \mathsf{A}_{s,\ell} u)^2 \right\rangle_\rho \qquad (2.32) \\
&\qquad \leqslant 10\lambda \left(\ell^{-2} \left\langle (\mathsf{A}_{s+\ell} u - \mathsf{A}_s u)^2 \right\rangle_\rho + \sum_{b \in (\Lambda_{s+\ell})^* \setminus (\Lambda_{s-2\mathbf{r}})^*} \left\langle c_b (\pi_b u)^2 \right\rangle_\rho \right).
\end{aligned}
$$

Then a "filling-hole" argument applies by adding $10\lambda \sum_{b \in (\Lambda_{s-2\mathbf{r}})^*} \left\langle c_b (\pi_b \mathsf{A}_{s,\ell} u)^2 \right\rangle_\rho$ on the two sides (Jensen's inequality is also applied to the right-hand side)

$$
\begin{aligned}
&(1 + 10\lambda) \sum_{b \in (\Lambda_{s-2\mathbf{r}})^*} \left\langle c_b (\pi_b \mathsf{A}_{s,\ell} u)^2 \right\rangle_\rho \qquad (2.33) \\
&\qquad \leqslant 10\lambda \left(\ell^{-2} \left\langle (\mathsf{A}_{s+\ell} u - \mathsf{A}_s u)^2 \right\rangle_\rho + \sum_{b \in (\Lambda_{s+\ell})^*} \left\langle c_b (\pi_b u)^2 \right\rangle_\rho \right).
\end{aligned}
$$

We note the martingale property of $(\mathsf{A}_s u)_{s \geqslant 0}$ and divide $(1 + 10\lambda)$ on the two sides to obtain (2.25) with $\theta' = \frac{10\lambda}{1+10\lambda}$.

Step 3: bootstrap. When we take L large enough, $\ell = L$, and $s = 2L$ in (2.25), we obtain

$$
\frac{1}{|\Lambda_L|} \langle \mathsf{A}_{2L} u (-\mathcal{L}_{\Lambda_L} \mathsf{A}_{2L} u) \rangle_\rho \leqslant 3^d \theta' \left(\frac{L^{-2}}{|\Lambda_{3L}|} \left\langle (\mathsf{A}_{3L} u)^2 \right\rangle_\rho + \frac{1}{|\Lambda_{3L}|} \langle u (-\mathcal{L}_{\Lambda_{3L}} u) \rangle_\rho \right). \qquad (2.34)
$$

The factor before the L^2 term is correct, but it misses a volume factor compared to the desired result (2.10), as we do not necessarily have $3^d \theta' < 1$. However, if we choose carefully $s = (1+\delta)L$ and $\ell = \delta L$, we obtain

$$
\begin{aligned}
&\frac{1}{|\Lambda_L|} \langle \mathsf{A}_L u (-\mathcal{L}_{\Lambda_L} \mathsf{A}_L u) \rangle_\rho \qquad (2.35) \\
&\quad \leqslant (1+2\delta)^d \theta' \left(\frac{(\delta L)^{-2}}{|\Lambda_{(1+2\delta)L}|} \left\langle (\mathsf{A}_{(1+2\delta)L} u)^2 \right\rangle_\rho + \frac{1}{|\Lambda_{(1+2\delta)L}|} \left\langle u (-\mathcal{L}_{\Lambda_{(1+2\delta)L}} u) \right\rangle_\rho \right).
\end{aligned}
$$

We can choose δ small such that $(1+2\delta)^d \theta' < 1$, and then iterate the Dirichlet energy term on the right-hand side, such that the domain increases progressively to Λ_{3L}. See [35, eq.(56)-(60)] for details, since this step is an algebraic iteration and independent of model. □

2.2. **Analytic tools on independent particles.** In this part, we develop the analytic tools on independent particles. We denote by $\widetilde{\mathcal{X}} := \mathbb{N}^{\mathbb{Z}^d}$ for the configuration space of independent particles on $\mathbb{Z}^d$, which does not have constraint for the number of particles on every site. For $\widetilde{\eta} \in \widetilde{\mathcal{X}}$ such that $\widetilde{\eta}_x \geqslant 1$, we define the jump operator

$$\widetilde{\eta}^{x,y} := \widetilde{\eta} - \delta_x + \delta_y,$$

and

$$(\widetilde{\pi}_{x,y}\widetilde{u})(\widetilde{\eta}) := \widetilde{u}(\widetilde{\eta}^{x,y}) - \widetilde{u}(\widetilde{\eta}).$$

The generator of the independent particles is

$$(\widetilde{\mathcal{L}}\widetilde{u})(\widetilde{\eta}) := \sum_{x \in \mathbb{Z}^d} \widetilde{\eta}_x \sum_{y \sim x} \widetilde{\pi}_{x,y}\widetilde{u}. \tag{2.36}$$

This generates a dynamic that every particle jumps independently with rate 1 to the nearest neighbor site.

Fix an $\alpha > 0$, we denote by $\mathrm{Poi}(\alpha)$ the Poisson distribution on $\mathbb{N}$ with mean α and $\widetilde{\mathbb{P}}_\alpha := \mathrm{Poi}(\alpha)^{\otimes \mathbb{Z}^d}$ the probability measure on $\widetilde{\mathcal{X}}$ which is stationary with respect to $\widetilde{\mathcal{L}}$ in (2.36), and $\langle\!\langle \cdot \rangle\!\rangle_\alpha$ its associated expectation. We usually write $\widetilde{\eta} \in \widetilde{\mathcal{X}}$ as a canonical random variable sampled under $\widetilde{\mathbb{P}}_\alpha$. We also denote by $\widetilde{\mathbb{P}}_{\Lambda,N}$ and $\langle\!\langle \cdot \rangle\!\rangle_{\Lambda,N}$ for the probability and expectation under the canonical ensemble, i.e. N particles distributed independently and uniformly in Λ.

2.2.1. *Mecke's identity.* The first lemma is about Mecke's identity, which simplifies some expectation under $\langle\!\langle \cdot \rangle\!\rangle_\alpha$ by adding one additional particle. This identity is inspired from the reference [55, Theorem 4.1].

Lemma 2.8 (Mecke's identity). *For any $\alpha > 0$ and $F : \widetilde{\mathcal{X}} \times \Lambda \to \mathbb{R}$ integrable under $\mathbb{P}_\alpha$, then for every $x \in \Lambda$, the following identity holds*

$$\langle\!\langle \widetilde{\eta}_x F(\widetilde{\eta}, x) \rangle\!\rangle_\alpha = \alpha \langle\!\langle F(\widetilde{\eta} + \delta_x, x) \rangle\!\rangle_\alpha . \tag{2.37}$$

Proof. We make the calculation directly

$$\begin{aligned}
\langle\!\langle \widetilde{\eta}_x F(\widetilde{\eta}, x) \rangle\!\rangle_\alpha &= e^{-\alpha} \sum_{k=0}^{\infty} \frac{\alpha^k}{k!} \langle\!\langle kF(\widetilde{\eta}, x) \,|\widetilde{\eta}_x = k \rangle\!\rangle_\alpha \\
&= \alpha e^{-\alpha} \sum_{k=1}^{\infty} \frac{\alpha^{(k-1)}}{(k-1)!} \langle\!\langle F(\widetilde{\eta}, x) \,|\widetilde{\eta}_x = k \rangle\!\rangle_\alpha \\
&= \alpha e^{-\alpha} \sum_{k=1}^{\infty} \frac{\alpha^{(k-1)}}{(k-1)!} \langle\!\langle F(\widetilde{\eta} + \delta_x, x) \,|\widetilde{\eta}_x = k-1 \rangle\!\rangle_\alpha \\
&= \alpha e^{-\alpha} \langle\!\langle F(\widetilde{\eta} + \delta_x, x) \rangle\!\rangle_\alpha .
\end{aligned}$$

□

The calculation over independent particle system has a close connection with the finite difference operator on lattice $\mathbb{Z}^d$ or $\mathbb{T}_L^d$. Given a function $\widetilde{u} : \widetilde{\mathcal{X}} \to \mathbb{R}$ which only dependents on the configuration in Λ^-, and using the expression $\widetilde{\eta} \llcorner \Lambda = \sum_{i=1}^N \delta_{x_i}$, then we have the following canonical projection

$$\widetilde{u}_N(x_1, x_2, \cdots, x_N) := \widetilde{u}(\widetilde{\eta}). \tag{2.38}$$

Moreover, $\widetilde{u}$ is a function on $\widetilde{\mathcal{X}}$ if and only if $\widetilde{u}_N$ is invariant under permutation for all $N \in \mathbb{N}_+$; see [35, Lemma A.1] for similar discussions.

We state some more properties using the expression (2.38). To better treat the high dimensional function, we define the following notation for shorthand,

$$\fint_{\Lambda^N} (\cdot) := \frac{1}{|\Lambda|^N} \sum_{x_1,\cdots,x_N \in \Lambda} (\cdot)\,,$$

then using the notation (2.38), we observe

$$\langle\!\langle \widetilde{u} \rangle\!\rangle_{\Lambda,N} = \fint_{\Lambda^N} \widetilde{u}_N. \tag{2.39}$$

For any integer $1 \leqslant i \leqslant N$ and $e \in U := \{e' \in \mathbb{Z}^d : |e'| = 1\}$, the finite difference operator $\mathcal{D}_{x_i,e}$ is defined for $\widetilde{u}_N$ as

$$(\mathcal{D}_{x_i,e}\widetilde{u}_N)(x_1,\cdots,x_N) := \widetilde{u}_N(x_1,\cdots,x_i+e,\cdots,x_N) - \widetilde{u}_N(x_1,\cdots,x_i,\cdots,x_N), \tag{2.40}$$

which is commutative in the sense

$$\forall 1 \leqslant i,j \leqslant N, \forall e,e' \in U, \qquad \mathcal{D}_{x_i,e}\mathcal{D}_{x_j,e'} = \mathcal{D}_{x_j,e'}\mathcal{D}_{x_i,e}. \tag{2.41}$$

Combing the canonical projection (2.38), the finite difference operator is related to the generator $\widetilde{\mathcal{L}}$ in (2.36) by the following identity

$$\widetilde{\mathcal{L}}\widetilde{u}(\widetilde{\eta}) = \sum_{i=1}^N \sum_{e\in U} \mathcal{D}_{x_i,e}\widetilde{u}_N(x_1,\cdots,x_N). \tag{2.42}$$

Then a lot of analytic tools on Euclidean space can be applied to independent particles.

2.2.2. *H^2-estimate.* In the following paragraphs, we will recall H^2 estimate for independent particles (indeed identity), which will be used in the multiscale Poincaré estimate later in Proposition 2.12. The proof follows [11, Lemma B.19] and [17, Proposition 3.10] after a careful review. We will also explain in detail in Remark 2.10 the difficulty met when developing the counterpart for Kawasaki dynamics.

Lemma 2.9 (Dimension-free H^2 estimate on torus)**.** *For any $L, N \in \mathbb{N}_+$ and any function $u, f : (\mathbb{T}_L^d)^N \to \mathbb{R}$ satisfying*

$$\sum_{i=1}^N \sum_{e\in U} \mathcal{D}_{x_i,e} u = f, \tag{2.43}$$

the following identity holds

$$\fint_{(\mathbb{T}_L^d)^N} \sum_{i,j=1}^N \sum_{e,e'\in U} (\mathcal{D}_{x_j,e'}\mathcal{D}_{x_i,e}u)^2 = 4 \fint_{(\mathbb{T}_L^d)^N} f^2. \tag{2.44}$$

Proof. It is easy to verify the identity

$$\sum_{i=1}^N \sum_{e\in U} \mathcal{D}_{x_i,e} = \frac{1}{2}\sum_{i=1}^N \sum_{e\in U} \mathcal{D}_{x_i,-e}\mathcal{D}_{x_i,e},$$

so (2.43) is equivalent to

$$\frac{1}{2}\sum_{i=1}^N \sum_{e\in U} \mathcal{D}_{x_i,-e}\mathcal{D}_{x_i,e} u = f.$$

We evaluate the L^2 sum of the two sides

$$
\begin{aligned}
\fint_{(\mathbb{T}_L^d)^N} f^2 &= \frac{1}{4} \fint_{(\mathbb{T}_L^d)^N} \left(\sum_{i=1}^N \sum_{e \in U} \mathcal{D}_{x_i,-e} \mathcal{D}_{x_i,e} u \right)^2 \\
&= \frac{1}{4} \fint_{(\mathbb{T}_L^d)^N} \left(\sum_{i=1}^N \sum_{e \in U} \mathcal{D}_{x_i,-e} \mathcal{D}_{x_i,e} u \right) \left(\sum_{j=1}^N \sum_{e' \in U} \mathcal{D}_{x_j,-e'} \mathcal{D}_{x_j,e'} u \right).
\end{aligned} \tag{2.45}
$$

On torus, it is easy to verify the integration by part formula for $u, v : (\mathbb{T}_L^d)^N \to \mathbb{R}$

$$
\forall 1 \leqslant i \leqslant N, e \in U, \qquad \fint_{(\mathbb{T}_L^d)^N} (\mathcal{D}_{x_i,e} u) v = \fint_{(\mathbb{T}_L^d)^N} u (\mathcal{D}_{x_i,-e} v).
$$

We apply it and the commutativity (2.41) to the right-hand side of (2.45)

$$
\begin{aligned}
\fint_{(\mathbb{T}_L^d)^N} \left(\mathcal{D}_{x_i,-e} \mathcal{D}_{x_i,e} u\right) \left(\mathcal{D}_{x_j,-e'} \mathcal{D}_{x_j,e'} u\right) &= \fint_{(\mathbb{T}_L^d)^N} \left(\mathcal{D}_{x_i,e} u\right) \left(\mathcal{D}_{x_i,e} \mathcal{D}_{x_j,-e'} \mathcal{D}_{x_j,e'} u\right) \\
&= \fint_{(\mathbb{T}_L^d)^N} \left(\mathcal{D}_{x_i,e} u\right) \left(\mathcal{D}_{x_j,-e'} \mathcal{D}_{x_i,e} \mathcal{D}_{x_j,e'} u\right) \\
&= \fint_{(\mathbb{T}_L^d)^N} \left(\mathcal{D}_{x_j,e'} \mathcal{D}_{x_i,e} u\right) \left(\mathcal{D}_{x_i,e} \mathcal{D}_{x_j,e'} u\right) \\
&= \fint_{(\mathbb{T}_L^d)^N} \left(\mathcal{D}_{x_i,e} \mathcal{D}_{x_j,e'} u\right)^2.
\end{aligned}
$$

We put it back to the right-hand side of (2.45) and conclude (2.44). □

Remark 2.10. If one hopes to recover a similar identity on Kawasaki dynamics for $u, f : \mathcal{X} \to \mathbb{R}$, such that $\sum_{b \in (\mathbb{T}_L^d)^*} \pi_b u = f$, then we also have $\pi_b = \frac{1}{2} \pi_b \pi_b$ and integration by part formula. However, the main difficulty appears in the commutativity. The identity

$$
\pi_b \pi_{b'} = \pi_{b'} \pi_b,
$$

holds when $b \cap b' = \varnothing$. Otherwise, when b, b' shares common endpoint, as the symmetry group is not Abelian, some exotic term will generate and pose challenge.

We then extend the result above to the discrete Poisson equation on cube Λ_L. Here we add the Neumann boundary condition (2.46), and the indicator in (2.47) excludes the second-order finite difference outside Λ_L^+.

Corollary 2.11 (Dimension-free H^2 estimate on cube)**.** *For any $L, N \in \mathbb{N}_+$ and any function $u, f : (\Lambda_L)^N \to \mathbb{R}$ satisfying*

$$
\sum_{i=1}^N \sum_{e \in U} \mathcal{D}_{x_i,e} u = f,
$$

with the Neumann boundary condition

$$
\mathcal{D}_{x_i,e} u \mathbf{1}_{\{x_i \in \partial \Lambda_L, x_i + e \notin \Lambda_L\}} = 0, \tag{2.46}
$$

we have

$$
\fint_{(\Lambda_L)^N} \sum_{i,j=1}^N \sum_{e,e' \in U} (\mathcal{D}_{x_j,e'} \mathcal{D}_{x_i,e} u)^2 \mathbf{1}_{\{x_i = x_j, e = e', x_i + 2e \notin \Lambda_L\}} = 4 \fint_{(\Lambda_L)^N} f^2. \tag{2.47}
$$

Proof. The main idea of proof is periodization. Only in this proof, we work on half integers. With a translation and scaling, we set the cube of size L is $\square$ =

$(\mathbb{Z}+\frac{1}{2})^d \cap [0,L]^d$ and set $\widetilde{\square} = (\mathbb{Z}+\frac{1}{2})^d \cap [-L,L]^d$. We still denote by $u, f : (\square)^N \to \mathbb{R}$ and extend them to $\widetilde{u}, \widetilde{f} : (\widetilde{\square})^N \to \mathbb{R}$ by mirror symmetry

$$x_1, \cdots, x_N \in (\widetilde{\square})^N, \qquad \widetilde{u}(x_1, \cdots, x_N) := u(|x_1|, \cdots, |x_N|). \tag{2.48}$$

Next, we identify the opposite sides of the cube $[-L,L]^d$ together to get a torus, and view $\widetilde{\square}$ as a lattice torus $\mathbb{T}^d_{2L}$. Then functions $\widetilde{u}, \widetilde{f}$ on $\widetilde{\square}$ can be regarded as functions on $\mathbb{T}^d_{2L}$. Moreover, the Neaumann boundary condition and the mirror symmetry implies that on the whole $\mathbb{T}^d_{2L}$, we have

$$\sum_{i=1}^N \sum_{e \in U} \mathcal{D}_{x_i, e} \widetilde{u} = \widetilde{f},$$

Therefore, Lemma 2.9 applies to obtain

$$\fint_{(\mathbb{T}^d_{2L})^N} \sum_{i,j=1}^N \sum_{e,e' \in U} (\mathcal{D}_{x_j,e'} \mathcal{D}_{x_i,e} \widetilde{u})^2 = 4 \fint_{(\mathbb{T}^d_{2L})^N} \widetilde{f}^2.$$

Each side counts 2^d times in the integration over $\square$, which gives us

$$\fint_{\square^N} \sum_{i,j=1}^N \sum_{e,e' \in U} (\mathcal{D}_{x_j,e'} \mathcal{D}_{x_i,e} \widetilde{u})^2 = 4 \fint_{\square^N} \widetilde{f}^2.$$

We realize that the second-order derivative of $\widetilde{u}$ near the boundary vanishes due to the Neaumann boundary condition and the mirror symmetry, then we conclude (2.47).

$\square$

2.2.3. *Multiscale Poincaré inequality.* In this part, we introduce the notion of spatial average and use it to develop some kind of multiscale Poincaré inequality.

We define the gradient by adding one more particle

$$\begin{aligned} (\partial_k \widetilde{u})(\widetilde{\eta}, x) &:= \widetilde{u}(\widetilde{\eta} + \delta_{x+e_k}) - \widetilde{u}(\widetilde{\eta} + \delta_x), \\ (\widetilde{\nabla} \widetilde{u})(\widetilde{\eta}, x) &:= ((\partial_1 \widetilde{u})(\widetilde{\eta}, x), (\partial_2 \widetilde{u})(\widetilde{\eta}, x), \cdots, (\partial_d \widetilde{u})(\widetilde{\eta}, x)). \end{aligned} \tag{2.49}$$

Recalling the definition of the enlarged domain in (1.36), we define the filtration

$$\widetilde{\mathcal{G}}_{\Lambda^+} := \sigma\left(\sum_{x \in \Lambda^+} \widetilde{\eta}_x, \{\widetilde{\eta}_y, y \in (\Lambda^+)^c\} \right). \tag{2.50}$$

Using $\mathcal{Z}_{m,n}$ and $\mathcal{Z}_n$ defined in Section 1.4.1, we also define the spatial average operator

$$\mathsf{S}_n(\widetilde{\nabla} \widetilde{u})(\widetilde{\eta}, x) := \sum_{z \in \mathcal{Z}_n} \left(\frac{1}{|\square_n|} \sum_{y \in z + \square_n} \left\langle\!\!\left\langle (\widetilde{\nabla} \widetilde{u})(\widetilde{\eta}, y) \,|\, \widetilde{\mathcal{G}}_{(z+\square_n)^+} \right\rangle\!\!\right\rangle \right) \mathbf{1}_{\{x \in z + \square_n\}}. \tag{2.51}$$

That is, this operator makes spatial average over the added particle and over the local configuration. The enlarged domain $(z+\square_n)^+$ is needed to include all the sites to add particle in $(\widetilde{\nabla}\widetilde{u}(\mu, x))_{x \in z+\square_n}$. Then note that $\left\langle\!\left\langle \cdot \,|\widetilde{\mathcal{G}}_{(z+\square_n)^+} \right\rangle\!\right\rangle_\alpha$ does not dependent on $\alpha > 0$, so we drop the density α. This leads to the multiscale Poincaré inequality.

Proposition 2.12 (Multiscale Poincaré inequality)**.** *There exists a finite positive constant $C = C(d)$ such that for all function $\widetilde{u} : \widetilde{\mathcal{X}} \to \mathbb{R}$ such that $\left\langle\!\left\langle \widetilde{u} \,|\widetilde{\mathcal{G}}_{\square_m^+} \right\rangle\!\right\rangle = 0$, we have*

$$\left\langle\!\!\left\langle \frac{1}{|\square_m|} \widetilde{u}^2 \right\rangle\!\!\right\rangle_\alpha^{\frac{1}{2}} \leqslant C \alpha^{\frac{1}{2}} \sum_{n=1}^m 3^n \left(\frac{1}{|\mathcal{Z}_{m,n}|} \sum_{z \in \mathcal{Z}_{m,n}} \left\langle\!\!\left\langle \left| \mathsf{S}_n \widetilde{\nabla} \widetilde{u} \right|^2 (\widetilde{\eta}, z) \right\rangle\!\!\right\rangle_\alpha \right)^{\frac{1}{2}} \tag{2.52}$$

Proof. We are only interested in the case that $\widetilde{u}$ is integrable under $\langle\!\langle \cdot \rangle\!\rangle_\alpha$. The proof follows [35, Proposition 3.5] and we give a scratch here. Let $\widetilde{w}$ solve the following equation

$$\sum_{x\in\square_m} \widetilde{\eta}_x \sum_{y\sim x} \widetilde{\pi}_{x,y}\widetilde{w} = \widetilde{u},$$

in the sense of the Poisson equation with the same Neumann boundary condition (2.46) for $\widetilde{w}$ in Corollary 2.11 and by the identity (2.42). This will give us the H^2 estimate for $\widetilde{w}$ using the projection. Then we have

$$\begin{aligned}
\frac{1}{|\square_m|}\left\langle\!\!\left\langle \widetilde{u}^2 \right\rangle\!\!\right\rangle_\alpha &= \frac{1}{|\square_m|}\sum_{i=1}^d \sum_{x\in\square_m} \langle\!\langle \widetilde{\eta}_x (\widetilde{\pi}_{x,x+e_i}\widetilde{u})(\widetilde{\pi}_{x,x+e_i}\widetilde{w}) \rangle\!\rangle_\alpha \\
&= \frac{\alpha}{|\square_m|}\sum_{x\in\square_m} \left\langle\!\!\left\langle (\widetilde{\nabla}\widetilde{u})(\widetilde{\nabla}\widetilde{w})(\widetilde{\eta},x) \right\rangle\!\!\right\rangle_\alpha .
\end{aligned}$$

Here we use the Mecke's identity in Lemma 2.8. Then we add the local averages

$$\begin{aligned}
\frac{1}{|\square_m|}\left\langle\!\!\left\langle \widetilde{u}^2 \right\rangle\!\!\right\rangle_\alpha = \frac{\alpha}{|\square_m|}\sum_{x\in\square_m} &\left\langle\!\!\left\langle (\widetilde{\nabla}\widetilde{u})(\widetilde{\nabla}\widetilde{w} - \mathsf{S}_0\widetilde{\nabla}\widetilde{w})(\widetilde{\eta},x) \right\rangle\!\!\right\rangle_\alpha \\
&+ \sum_{n=0}^{m-1}\frac{\alpha}{|\square_m|}\sum_{x\in\square_m} \left\langle\!\!\left\langle (\mathsf{S}_n\widetilde{\nabla}\widetilde{u})(\mathsf{S}_n\widetilde{\nabla}\widetilde{w} - \mathsf{S}_{n+1}\widetilde{\nabla}\widetilde{w})(\widetilde{\eta},x) \right\rangle\!\!\right\rangle_\alpha \\
&+ \frac{\alpha}{|\square_m|}\sum_{x\in\square_m} \left\langle\!\!\left\langle (\mathsf{S}_m\widetilde{\nabla}\widetilde{u})(\mathsf{S}_m\widetilde{\nabla}\widetilde{w})(\widetilde{\eta},x) \right\rangle\!\!\right\rangle_\alpha .
\end{aligned}$$

We just focus on one scale, which gives us

$$\begin{aligned}
&\left| \frac{1}{|\square_m|}\sum_{x\in\square_m} \left\langle\!\!\left\langle (\mathsf{S}_n\widetilde{\nabla}\widetilde{u})(\mathsf{S}_n\widetilde{\nabla}\widetilde{w} - \mathsf{S}_{n+1}\widetilde{\nabla}\widetilde{w})(\widetilde{\eta},x) \right\rangle\!\!\right\rangle_\alpha \right| \\
&= \left| \frac{1}{|\mathcal{Z}_{m,n}|}\sum_{z\in\mathcal{Z}_{m,n}} \left\langle\!\!\left\langle (\mathsf{S}_n\widetilde{\nabla}\widetilde{u})(\mathsf{S}_n\widetilde{\nabla}\widetilde{w} - \mathsf{S}_{n+1}\widetilde{\nabla}\widetilde{w})(\widetilde{\eta},z) \right\rangle\!\!\right\rangle_\alpha \right| \\
&\leqslant \left(\frac{1}{|\mathcal{Z}_{m,n}|}\sum_{z\in\mathcal{Z}_{m,n}} \left\langle\!\!\left\langle \left|\mathsf{S}_n\widetilde{\nabla}\widetilde{u}\right|^2(\widetilde{\eta},z) \right\rangle\!\!\right\rangle_\alpha \right)^{\frac{1}{2}} \left(\frac{1}{|\mathcal{Z}_{m,n}|}\sum_{z\in\mathcal{Z}_{m,n}} \left\langle\!\!\left\langle \left|\mathsf{S}_n\widetilde{\nabla}\widetilde{w} - \mathsf{S}_{n+1}\widetilde{\nabla}\widetilde{w}\right|^2(\widetilde{\eta},z) \right\rangle\!\!\right\rangle_\alpha \right)^{\frac{1}{2}} \\
&\leqslant C\alpha^{-\frac{1}{2}}3^n \left(\frac{1}{|\mathcal{Z}_{m,n}|}\sum_{z\in\mathcal{Z}_{m,n}} \left\langle\!\!\left\langle \left|\mathsf{S}_n\widetilde{\nabla}\widetilde{u}\right|^2(\widetilde{\eta},z) \right\rangle\!\!\right\rangle_\alpha \right)^{\frac{1}{2}} \left\langle\!\!\left\langle \frac{1}{|\square_m|}\widetilde{u}^2 \right\rangle\!\!\right\rangle_\alpha^{\frac{1}{2}} .
\end{aligned}$$

From the third line to the forth line, we use the Poincaré inequality of independent particles for the term $\left|\mathsf{S}_n\widetilde{\nabla}\widetilde{w} - \mathsf{S}_{n+1}\widetilde{\nabla}\widetilde{w}\right|^2$, and this will give the factor 3^n. The output is the second-order derivatives of $\widetilde{w}$, which will be bounded by $\left\langle\!\left\langle \widetilde{u}^2 \right\rangle\!\right\rangle_\alpha$ using the dimension-free H^2 estimate in Corollary 2.11. Here the factor $\alpha^{-\frac{1}{2}}$ comes from another application of Mecke's identity (2.37), and this concludes (2.52). □

3. Coarse-grained lifting

This part introduces the key technique to handle the constraint of particle numbers in Kawasaki dynamics, which is the coarse-grained lifting to independent particles. Let $\widetilde{\eta} \in \widetilde{\mathcal{X}} = \mathbb{N}^{\mathbb{Z}^d}$ stand for the configuration of independent particles, and $\eta \in \mathcal{X} = \{0,1\}^{\mathbb{Z}^d}$ for the configuration in Kawasaki dynamics. We aim to embed $\mathcal{X}$ into $\widetilde{\mathcal{X}}$, and a natural idea is to define the following projection $[\,] : \widetilde{\mathcal{X}} \to \mathcal{X}$ that

$$\forall x \in \mathbb{Z}^d, \qquad [\widetilde{\eta}]_x \coloneqq \mathbf{1}_{\{\widetilde{\eta}_x \geqslant 1\}}. \tag{3.1}$$

Therefore, $[\widetilde{\eta}]$ only indicates whether the site is occupied, but does not care the exact number of particles. This projection operator also induces an extension for every function $u : \mathcal{X} \to \mathbb{R}$ by pull-back that

$$[u] : \widetilde{\mathcal{X}} \to \mathbb{R}, \qquad [u](\widetilde{\eta}) \coloneqq u([\widetilde{\eta}]), \tag{3.2}$$

and we call $[u]$ the coarsened function of u. In the following paragraphs, we will use $[\widetilde{\eta}]$ to represent the configuration of Kawasaki dynamics, and explore several identities under this projection/coarsen.

3.1. **Grand canonical ensemble.** Let the configuration of independent particles $\widetilde{\eta}$ follow the law $\widetilde{\mathbb{P}}_\alpha$, which is independent Poisson distribution of parameter $\alpha > 0$. In order to make $[\widetilde{\eta}]$ of the same law as sampled from $\mathbb{P}_\rho$, we make a specific choice between the parameters such that $e^{-\alpha} = 1 - \rho$, i.e.

$$\forall \rho \in (0,1), \qquad \alpha(\rho) \coloneqq -\log(1-\rho). \tag{3.3}$$

Under this specific choice of parameter, we can see the coarsen function $[u]$ as a *lift* of u thanks of the following proposition.

Proposition 3.1 (Coarsen-grained lifting)**.** *Given $\rho \in (0,1)$ and $\alpha(\rho)$ defined as (3.3) and $\widetilde{\eta}$ sampled from $\widetilde{\mathbb{P}}_{\alpha(\rho)}$, then $[\widetilde{\eta}] \in \mathcal{X}$ follows the law $\mathbb{P}_\rho$. As a consequence, for every $u : \mathcal{X} \to \mathbb{R}$ integrable under $\mathbb{P}_\rho$, its coarsen function satisfies*

$$\langle\!\langle [u] \rangle\!\rangle_{\alpha(\rho)} = \langle u \rangle_\rho . \tag{3.4}$$

Proof. With the choice of the parameter (3.3) and the definition of the projection operator (3.1), we have

$$\begin{aligned}
\forall x \in \mathbb{Z}^d, \qquad & \widetilde{\mathbb{P}}_{\alpha(\rho)}[[\widetilde{\eta}]_x = 0] = \widetilde{\mathbb{P}}_{\alpha(\rho)}[\widetilde{\eta}_x = 0] = e^{-\alpha(\rho)} = 1 - \rho, \\
& \widetilde{\mathbb{P}}_{\alpha(\rho)}[[\widetilde{\eta}]_x = 1] = \widetilde{\mathbb{P}}_{\alpha(\rho)}[\widetilde{\eta}_x \geqslant 1] = 1 - \widetilde{\mathbb{P}}_{\alpha(\rho)}[\widetilde{\eta}_x = 0] = \rho.
\end{aligned}$$

Therefore, $[\widetilde{\eta}]_x$ follows the Bernoulli law with parameter ρ. Since $(\widetilde{\eta}_x)_{x\in\mathbb{Z}^d}$ are i.i.d. random variables, $[\widetilde{\eta}]$ has the same law as $\mathbb{P}_\rho$. Combing this fact and (3.2), the identity (3.4) is a direct corollary

$$\langle\!\langle [u] \rangle\!\rangle_{\alpha(\rho)} = \langle\!\langle u([\widetilde{\eta}]) \rangle\!\rangle_{\alpha(\rho)} = \langle u(\eta) \rangle_\rho .$$

□

Although we can couple the static configuration between two systems and obtain the nice identity (3.4), similar result cannot be extended to the Dirichlet energy and we cannot expect a similar identity like

$$\sum_{x\in\Lambda} \sum_{y\in\Lambda, y\sim x} \left\langle\!\!\left\langle \widetilde{\eta}_x (\widetilde{\pi}_{x,y}[u])^2 \right\rangle\!\!\right\rangle_{\alpha(\rho)} = C(\rho) \sum_{x,y\in\Lambda, x\sim y} \left\langle (\pi_{x,y} u)^2 \right\rangle_\rho . \tag{3.5}$$

On the other hand, the coarse-grained lifting can be very useful when evaluating the spatial average of the gradient. We establish the following identity, which can be used as the gradient coupling. We recall the definition of the tangent field ∇_{x,e_i} for Kawasaki dynamics defined in (1.46).

Proposition 3.2 (Gradient coupling)**.** *For every $\rho \in (0,1), \Lambda \subseteq \mathbb{Z}^d$ and $u : \mathcal{X} \to \mathbb{R}$, the following identity holds for every $i \in \{1, \cdots, d\}$*

$$\sum_{x\in\Lambda} \langle\!\langle \widetilde{\eta}_x \widetilde{\pi}_{x,x+e_i}[u] \rangle\!\rangle_{\alpha(\rho)} = \frac{\alpha(\rho)}{2\rho} \sum_{x\in\Lambda} \langle \nabla_{x,e_i} u \rangle_\rho . \tag{3.6}$$

Proposition 3.2 is the result of the following lemmas. We state them separately as they can be useful in other proofs. The first lemma is a similar version of Mecke's identity in Kawasaki dynamics. Here, instead of adding a particle, the added particle is understood as "forcing the site occupied".

Lemma 3.3 (Mecke's identity in Kawasaki dynamics)**.** *For every $\rho \in (0,1)$ and $u : \mathcal{X} \to \mathbb{R}$, the following identity holds*

$$\langle \nabla_{x,e_i} u \rangle_\rho = 2\rho \left(\langle u \,|\eta_{x+e_i} = 1 \rangle_\rho - \langle u \,|\eta_x = 1 \rangle_\rho \right). \tag{3.7}$$

Proof. Recall the identity (1.47) and the term is non-zero if and only if $(\eta_x, \eta_{x+e_i}) = (1,0)$ or $(\eta_x, \eta_{x+e_i}) = (0,1)$. Then we obtain that

$$\langle (\pi_{x,x+e_i} u)(\pi_{x,x+e_i} \ell_{e_i}) \rangle_\rho = 2 \left\langle u(\eta) \mathbf{1}_{\{(\eta_x, \eta_{x+e_i}) = (0,1)\}} - u(\eta) \mathbf{1}_{\{(\eta_x, \eta_{x+e_i}) = (1,0)\}} \right\rangle_\rho .$$

Here noticing that the term $(\eta_x, \eta_{x+e_i}) = (1,1)$ is canceled on the right-hand side of the equation, we can rewrite it as

$$\begin{aligned}
&\left\langle u(\eta) \mathbf{1}_{\{(\eta_x, \eta_{x+e_i}) = (0,1)\}} - u(\eta) \mathbf{1}_{\{(\eta_x, \eta_{x+e_i}) = (1,0)\}} \right\rangle_\rho \\
&= \left\langle u(\eta) \mathbf{1}_{\{(\eta_x, \eta_{x+e_i}) = (0,1)\}} + u(\eta) \mathbf{1}_{\{(\eta_x, \eta_{x+e_i}) = (1,1)\}} \right\rangle_\rho \\
&\qquad - \left\langle u(\eta) \mathbf{1}_{\{(\eta_x, \eta_{x+e_i}) = (1,0)\}} + u(\eta) \mathbf{1}_{\{(\eta_x, \eta_{x+e_i}) = (1,1)\}} \right\rangle_\rho \\
&= \left\langle u(\eta) \mathbf{1}_{\{\eta_{x+e_i} = 1\}} - u(\eta) \mathbf{1}_{\{\eta_x = 1\}} \right\rangle_\rho .
\end{aligned}$$

Therefore, we obtain the identity

$$\begin{aligned}
\langle (\pi_{x,x+e_i} u)(\pi_{x,x+e_i} \ell_{e_i}) \rangle_\rho &= 2 \left\langle u(\eta) \mathbf{1}_{\{\eta_{x+e_i} = 1\}} - u(\eta) \mathbf{1}_{\{\eta_x = 1\}} \right\rangle_\rho \\
&= 2\rho \left(\langle u \,|\eta_{x+e_i} = 1 \rangle_\rho - \langle u \,|\eta_x = 1 \rangle_\rho \right).
\end{aligned}$$

□

The second lemma makes the bridge between Lemma 2.8 and Lemma 3.3, and is a corollary from Proposition 3.1.

Lemma 3.4 (Change of variable)**.** *For any $\rho \in (0,1)$ and $x, y \in \mathbb{Z}^d$, let $\eta \in \mathcal{X}$ be sampled from $\mathbb{P}_\rho$ and $\widetilde{\eta} \in \widetilde{\mathcal{X}}$ sampled from $\widetilde{\mathbb{P}}_{\alpha(\rho)}$, then we have*

$$([\widetilde{\eta} + \delta_x], [\widetilde{\eta} + \delta_y]) \overset{(d)}{=} (\eta \vee \delta_x, \eta \vee \delta_y), \tag{3.8}$$

where $\eta \vee \delta_x \in \mathcal{X}$ is defined as $(\eta \vee \delta_x)_x = \eta_x \vee 1$ and $(\eta \vee \delta_x)_z = \eta_z$ for $z \neq x$.

Proof. The proof follows the observation that

$$[\widetilde{\eta} + \delta_x] = [\widetilde{\eta}] \vee \delta_x.$$

Thus, we apply Proposition 3.1

$$([\widetilde{\eta} + \delta_x], [\widetilde{\eta} + \delta_y]) = ([\widetilde{\eta}] \vee \delta_x, [\widetilde{\eta}] \vee \delta_y) \overset{(d)}{=} (\eta \vee \delta_x, \eta \vee \delta_y).$$

□

Proof of Proposition 3.2. We combine the results in Lemma 2.8, Lemma 3.3 and Lemma 3.4

$$\sum_{x \in \Lambda} \langle\!\langle \widetilde{\eta}_x \widetilde{\pi}_{x,x+e_i} [u] \rangle\!\rangle_{\alpha(\rho)} \overset{(2.37)}{=} \alpha(\rho) \sum_{x \in \Lambda} \langle\!\langle [u](\widetilde{\eta} + \delta_{x+e_i}) - [u](\widetilde{\eta} + \delta_x) \rangle\!\rangle_{\alpha(\rho)}$$

$$\overset{(3.2)}{=} \alpha(\rho) \sum_{x\in\Lambda} \langle\!\langle u([\widetilde{\eta}+\delta_{x+e_i}]) - u([\widetilde{\eta}+\delta_x]) \rangle\!\rangle_{\alpha(\rho)}$$
$$\overset{(3.8)}{=} \alpha(\rho) \sum_{x\in\Lambda} \langle u(\eta\vee\delta_{x+e_i}) - u(\eta\vee\delta_x)\rangle_\rho$$
$$\overset{(3.7)}{=} \frac{\alpha(\rho)}{2\rho} \sum_{x\in\Lambda} \langle (\pi_{x,x+e_i}u)(\pi_{x,x+e_i}\ell_{e_i})\rangle_\rho .$$

Here from the third line to the forth line, we also use the fact $\langle u\,|\eta_x = 1\rangle_\rho = \langle u(\eta\vee\delta_x)\rangle_\rho$. □

At the end of this subsection, we give another application of Lemma 3.3. We recall the notation Λ^- defined in (1.34).

Corollary 3.5. *For every bounded set $\Lambda \subseteq \mathbb{Z}^d$ and $u \in \mathcal{F}_0(\Lambda^-)$, we have*

$$\sum_{x\in\Lambda} \langle \nabla_x u\rangle_\rho = 0. \tag{3.9}$$

Proof. We just focus on the gradient field along one direction e_i and apply (3.7)

$$\begin{aligned}\sum_{x\in\Lambda} \langle \nabla_{x,e_i} u\rangle_\rho &= 2\rho \sum_{x\in\Lambda} \left(\langle u(\eta\vee\delta_{x+e_i})\rangle_\rho - \langle u(\eta\vee\delta_x)\rangle_\rho\right) \\ &= 2\rho\left(\sum_{x\in\Lambda, x+e_i\notin\Lambda} \langle u(\eta\vee\delta_x)\rangle_\rho - \sum_{x\in\Lambda, x-e_i\notin\Lambda} \langle u(\eta\vee\delta_x)\rangle_\rho\right) \\ &= 2\rho\left(\sum_{x\in\Lambda, x+e_i\notin\Lambda} \langle u(\eta)\rangle_\rho - \sum_{x\in\Lambda, x-e_i\notin\Lambda} \langle u(\eta)\rangle_\rho\right) \\ &= 0.\end{aligned}$$

In the first line, we also use the observation $\langle u\,|\eta_x = 1\rangle_\rho = \langle u(\eta\vee\delta_x)\rangle_\rho$. Then the proof is similar to the discrete Stokes' formula that all the terms except those on the boundary cancel, which yields the second line. The condition $u \in \mathcal{F}_0(\Lambda^-)$ implies that $u(\eta\vee\delta_x) = u(\eta)$ and passes the result from the second line to the third line. Finally, as the terms in $\sum_{x\in\Lambda, x+e_i\notin\Lambda}$ and $\sum_{x\in\Lambda, x-e_i\notin\Lambda}$ are coupled, we obtain 0. □

3.2. **Canonical ensemble.**

We establish a gradient coupling similar to Proposition 3.2 under the canonical ensemble. Recall Λ^+ defined in (1.36).

Proposition 3.6 (Gradient coupling)**.** *For every $M \in \mathbb{N}, \Lambda \subseteq \mathbb{Z}^d$ and $u : \mathcal{X} \to \mathbb{R}$ a $\mathcal{F}_{\Lambda^+}$-measurable function, let $(\widetilde{\eta}_x)_{x\in\Lambda^+}$ be sampled from the canonical ensemble of independent particles $\widetilde{\mathbb{P}}_{\Lambda^+,M}$ and $(P_{\Lambda,M,N})_{N\in\mathbb{N}}$ be the probability of the number of occupied sites defined as*

$$P_{\Lambda,M,N} := \widetilde{\mathbb{P}}_{\Lambda^+,M}\left[\sum_{z\in\Lambda^+\smallsetminus\{y\}} \mathbf{1}_{\{\widetilde{\eta}_z>0\}} = N-1\right], \qquad y \in \Lambda^+, \tag{3.10}$$

then the following identity holds for every $i \in \{1,\cdots,d\}$

$$\sum_{x\in\Lambda} \langle\!\langle \partial_i[u](\widetilde{\eta},x)\rangle\!\rangle_{\Lambda^+,M} = \sum_{N=1}^{M} \frac{|\Lambda^+|P_{\Lambda,M,N}}{2N} \sum_{x\in\Lambda} \langle \nabla_{x,e_i} u\rangle_{\Lambda^+,N} . \tag{3.11}$$

Remark 3.7. One can check that (3.10) is well-defined and does not depend on the choice of y. One can pick a specific y and put it back to (3.11). This will make the identity there a little strange, but it transfers the symmetry used in the proof.

To prove this proposition, we also give a Mecke's identity like Lemma 3.3 under the canonical ensemble.

Lemma 3.8 (Mecke's identity under canonical ensemble)**.** *Given $u : \mathcal{X} \to \mathbb{R}$, then for every $N \in \mathbb{N}$, a subset $\Lambda \subseteq \mathbb{Z}^d$ and $x \in \Lambda$, the following identity holds*

$$\langle \nabla_{x,e_i} u \rangle_{\Lambda^+,N} = \frac{2N}{|\Lambda^+|} \left(\langle u \,|\eta_{x+e_i} = 1 \rangle_{\Lambda^+,N} - \langle u \,|\eta_x = 1 \rangle_{\Lambda^+,N} \right). \tag{3.12}$$

Proof. The proof follows the similar strategy as Lemma 3.3

$$\begin{aligned}
\langle (\pi_{x,x+e_i} u)(\pi_{x,x+e_i} \ell_{e_i}) \rangle_{\Lambda^+,N} &= 2 \left\langle u(\eta) \mathbf{1}_{\{(\eta_x,\eta_{x+e_i})=(0,1)\}} - u(\eta) \mathbf{1}_{\{(\eta_x,\eta_{x+e_i})=(1,0)\}} \right\rangle_{\Lambda^+,N} \\
&= 2 \left\langle u(\eta) \mathbf{1}_{\{\eta_{x+e_i}=1\}} - u(\eta) \mathbf{1}_{\{\eta_x=1\}} \right\rangle_{\Lambda^+,N} \\
&= \frac{2N}{|\Lambda^+|} \left(\langle u \,|\eta_{x+e_i} = 1 \rangle_{\Lambda^+,N} - \langle u \,|\eta_x = 1 \rangle_{\Lambda^+,N} \right).
\end{aligned}$$

Here from the first line to the second line, we use the fact that the case $(\eta_x, \eta_{x+e_i}) = (1,1)$ cancels. From the second line to the third line, we notice that $x, x+e_i \in \Lambda^+$, so

$$\mathbb{P}_{\Lambda^+,N}[\eta_x = 1] = \mathbb{P}_{\Lambda^+,N}[\eta_{x+e_i} = 1] = \frac{\binom{|\Lambda^+|-1}{(N-1)}}{\binom{|\Lambda^+|}{N}} = \frac{N}{|\Lambda^+|}.$$

□

Proof of Proposition 3.6. Notice that, for every $x \in \Lambda^+$, we have

$$\begin{aligned}
&\langle\!\langle [u](\widetilde{\eta} + \delta_x) \rangle\!\rangle_{\Lambda^+,M} \\
&= \langle\!\langle u([\widetilde{\eta} + \delta_x]) \rangle\!\rangle_{\Lambda^+,M} \\
&= \sum_{N=1}^{M} \widetilde{\mathbb{P}}_{\Lambda^+,M} \left[\sum_{z \in \Lambda^+} \mathbf{1}_{\{\widetilde{\eta}_z + \delta_x > 0\}} = N \right] \left\langle\!\!\left\langle u([\widetilde{\eta} + \delta_x]) \,\middle|\, \sum_{z \in \Lambda^+} \mathbf{1}_{\{\widetilde{\eta}_z + \delta_x > 0\}} = N \right\rangle\!\!\right\rangle_{\Lambda^+,M} \qquad (3.13) \\
&= \sum_{N=1}^{M} P_{\Lambda,M,N} \langle u \,|\eta_x = 1 \rangle_{\Lambda^+,N}.
\end{aligned}$$

Here we apply the definition of the projection operator (3.1) from the first line to the second line, and the total probability formula from the second line to the third line. The passage from the third line to the forth line requires some work. Firstly, we recall the identity (3.10)

$$\widetilde{\mathbb{P}}_{\Lambda^+,M} \left[\sum_{z \in \Lambda^+} \mathbf{1}_{\{\widetilde{\eta}_z + \delta_x > 0\}} = N \right] = \widetilde{\mathbb{P}}_{\Lambda^+,M} \left[\sum_{z \in \Lambda^+ \setminus \{x\}} \mathbf{1}_{\{\widetilde{\eta}_z > 0\}} = N - 1 \right] = P_{\Lambda,M,N}.$$

We also remark that $P_{\Lambda,M,N}$ is well-defined and does not dependent on the excluded site x. Secondly, we notice that, for every choice of vertex set $V \subseteq \Lambda^+ \setminus \{x\}$ such that $|V| = N - 1$, in the following set

$$\left\{ \sum_{z \in \Lambda^+} \widetilde{\eta}_z = M : \widetilde{\eta}_z > 0 \text{ for all } z \in V, \text{ and } \widetilde{\eta}_z = 0 \text{ for all } z \in \Lambda^+ \setminus (\{x\} \cup V) \right\},$$

the number of configurations is the same. Therefore, the configuration $[\widetilde{\eta} + \delta_x]$ contains a particle on x, and the other occupied positions are uniformly distributed conditioned on the number of occupied sites. This generalizes the coarse-grained lifting under the canonical ensemble, and gives the passage concerning the conditional expectation in the third line of (3.13).

Applying (3.13), we obtain

$$\begin{aligned}
&\sum_{x\in\Lambda} \langle\!\langle [u](\widetilde{\eta}+\delta_{x+e_i}) - [u](\widetilde{\eta}+\delta_x)\rangle\!\rangle_{\Lambda^+,M} \\
&= \sum_{N=1}^{M} P_{\Lambda,M,N} \sum_{x\in\Lambda} \left(\langle u\,|\eta_{x+e_i}=1\rangle_{\Lambda^+,N} - \langle u\,|\eta_x = 1\rangle_{\Lambda^+,N} \right) \\
&= \sum_{N=1}^{M} \frac{|\Lambda^+|P_{\Lambda,M,N}}{2N} \sum_{x\in\Lambda} \langle (\pi_{x,x+e_i} u)(\pi_{x,x+e_i}\ell_{e_i})\rangle_{\Lambda^+,N}\,.
\end{aligned}$$

Here from the second line to the third line, we also use (3.12). This yields the desired result. □

3.3. **Weighted multiscale Poincaré inequality.** As the main result of this subsection, we combine the coarse-grained lifting and gradient coupling to obtain the weighted multiscale Poincaré inequality on Kawasaki dynamics. Here we define $\mathcal{G}_{\Lambda^+}$ similar to (2.50)

$$\mathcal{G}_{\Lambda^+} := \sigma\left(\sum_{x\in\Lambda^+} \eta_x, \{\eta_y, y\in(\Lambda^+)^c\} \right). \tag{3.14}$$

Proposition 3.9 (Weighted multiscale Poincaré inequality)**.** *There exists a finite positive constant $C = C(d,\rho)$ such that for all functions $u:\mathcal{X}\to\mathbb{R}$ such that $\left\langle u\,|\mathcal{G}_{\square_m^+}\right\rangle = 0$, the following estimate is established*

$$\begin{aligned}
&\left\langle \frac{1}{|\square_m|} u^2 \right\rangle_\rho^{\frac12} \\
&\leqslant C\alpha^{\frac12} \sum_{n=0}^{m} 3^n \left(\frac{1}{|\mathcal{Z}_{m,n}|} \sum_{z\in\mathcal{Z}_{m,n}} \left\langle \left(\frac{|\square_n^+|}{2\mathbf{N}_{z,n}}\right)^2 \left| \frac{1}{|\square_n|} \sum_{x\in z+\square_n} \langle \nabla_x u\rangle_{z+\square_n^+,\mathbf{N}_{z,n}} \right|^2 \right\rangle_\rho \right)^{\frac12}.
\end{aligned} \tag{3.15}$$

Here the random variable $\mathbf{N}_{z,n}$ is defined as $\mathbf{N}_{z,n} := 1 + \sum_{x\in(z+\square_n^+)\setminus\{z\}} \eta_x$.

Proof. Step 1: forward procedure. In order to better estimate its L^2 norm, we need the multiscale Poincaré inequality (2.52). Since this is only established in independent particle systems, we make use of the coarse-grained lifting developed in Proposition 3.1 as a bridge, that is

$$\begin{aligned}
\left\langle \frac{1}{|\square_m|} u^2 \right\rangle_\rho^{\frac12} &= \left\langle\!\!\left\langle \frac{1}{|\square_m|} [u]^2 \right\rangle\!\!\right\rangle_{\alpha(\rho)}^{\frac12} \\
&\leqslant C\alpha^{\frac12}(\rho) \sum_{n=1}^{m} 3^n \left(\frac{1}{|\mathcal{Z}_{m,n}|} \sum_{z\in\mathcal{Z}_{m,n}} \left\langle\!\!\left\langle \left|\mathsf{S}_n \widetilde{\nabla}[u]\right|^2 (\widetilde{\eta}, z) \right\rangle\!\!\right\rangle_{\alpha(\rho)} \right)^{\frac12}.
\end{aligned}$$

We also remark that $\left\langle u\,|\mathcal{G}_{\square_m^+}\right\rangle = 0$ ensures the condition $\left\langle [u]\,|\widetilde{\mathcal{G}}_{\square_m^+}\right\rangle = 0$ to apply (2.52).

Step 2: backward procedure. We focus on one term $\left\langle\!\!\left\langle \left(\mathsf{S}_n \widetilde{\nabla}[u]\right)^2 (\widetilde{\eta}, z) \right\rangle\!\!\right\rangle_{\alpha(\rho)}$, which is the spatial average in the independent particles. Our main task in this step is to bring $(\mathsf{S}_n\widetilde{\nabla}[u])(\widetilde{\eta},z)$ back to the original Kawasaki dynamics from the coarsen operator. We apply the definition of S_n in (2.51)

$$\left\langle\!\!\left\langle \left|\mathsf{S}_n \widetilde{\nabla}[u]\right|^2 (\widetilde{\eta}, z) \right\rangle\!\!\right\rangle_{\alpha(\rho)} \tag{3.16}$$

$$= \sum_{M=0}^{\infty} \widetilde{\mathbb{P}}_{\alpha(\rho)} \left[\sum_{y \in z + \square_n^+} \widetilde{\eta}_y = M \right] \left\langle\!\!\left\langle \left| \frac{1}{|\square_n|} \sum_{x \in z + \square_n} \left\langle\!\!\left\langle \widetilde{\nabla}[u](\widetilde{\eta}, x) \right\rangle\!\!\right\rangle_{z+\square_n^+, M} \right|^2 \right\rangle\!\!\right\rangle_{\alpha(\rho)}.$$

More precisely, by $\left\langle\!\left\langle \widetilde{\nabla}[u](\widetilde{\eta}, x) \right\rangle\!\right\rangle_{z+\square_n^+, M}$ we mean

$$\left\langle\!\!\left\langle \widetilde{\nabla}[u](\widetilde{\eta}, x) \right\rangle\!\!\right\rangle_{z+\square_n^+, M} = \int_{\widetilde{\mathcal{X}}} \widetilde{\nabla}[u](\widetilde{\eta}' \llcorner (z + \square_n^+) + \widetilde{\eta} \llcorner (z + \square_n^+)^c, x) \, \mathrm{d}\widetilde{\mathbb{P}}_{z+\square_n^+, M}(\widetilde{\eta}').$$

Then we apply the gradient coupling (3.11) to the term $\widetilde{\eta}'$ in the right-hand side

$$\begin{aligned}
&\left(\frac{1}{|\square_n|} \sum_{x \in z + \square_n} \langle\!\langle \partial_i [u](\widetilde{\eta}, x) \rangle\!\rangle_{z+\square_n^+, M} \right)^2 \\
&= \left(\sum_{N=1}^{M} P_{z+\square_n, M, N} \frac{|\square_n^+|}{2N} \right. \\
&\qquad \left. \frac{1}{|\square_n|} \sum_{x \in z + \square_n} \langle (\nabla_{x, e_i} u)(\eta \llcorner (z + \square_n^+) + [\widetilde{\eta}] \llcorner (z + \square_n^+)^c) \rangle_{z+\square_n^+, N} \right)^2 \\
&\leqslant \sum_{N=1}^{M} P_{z+\square_n, M, N} \left(\frac{|\square_n^+|}{2N} \right)^2 \\
&\qquad \left(\frac{1}{|\square_n|} \sum_{x \in z + \square_n} \langle (\nabla_{x, e_i} u)(\eta \llcorner (z + \square_n^+) + [\widetilde{\eta}] \llcorner (z + \square_n^+)^c) \rangle_{z+\square_n^+, N} \right)^2.
\end{aligned}$$

From the second line to the third line, we make use of Jensen's inequality. We also remark that, our function is not $\mathcal{F}_{z+\square_n^+}$-measurable, thus we keep $[\widetilde{\eta}] \llcorner (z + \square_n^+)^c$ after the local average. We put this result back to (3.16) and apply Proposition 3.1 to $[\widetilde{\eta}] \llcorner (z + \square_n^+)^c$, which gives us

$$\left\langle\!\!\left\langle \left(\frac{1}{|\square_n|} \sum_{x \in z + \square_n} \langle (\nabla_{x, e_i} u)(\eta \llcorner (z + \square_n^+) + [\widetilde{\eta}] \llcorner (z + \square_n^+)^c) \rangle_{z+\square_n^+, N} \right)^2 \right\rangle\!\!\right\rangle_{\alpha(\rho)} = \underbrace{\left\langle \left(\frac{1}{|\square_n|} \sum_{x \in z + \square_n} \langle \nabla_{x, e_i} u \rangle_{z+\square_n^+, N} \right)^2 \right\rangle_{\rho}}_{=:F(N)}. \tag{3.17}$$

Using the definition (3.10) with a specific choice $y = z$ there, we have

$$P_{z+\square_n, M, N} = \widetilde{\mathbb{P}}_{z+\square_n^+, M} \left[\sum_{x \in (z + \square_n^+) \setminus \{z\}} \mathbf{1}_{\{\widetilde{\eta}_x > 0\}} = N - 1 \right].$$

Then we apply Fubuni's lemma to the double sum

$$\begin{aligned}
&\left\langle\!\!\left\langle (\mathsf{S}_n(\partial_i [u])(\widetilde{\eta}, z))^2 \right\rangle\!\!\right\rangle_{\alpha(\rho)} \\
&\leqslant \sum_{M=0}^{\infty} \widetilde{\mathbb{P}}_{\alpha(\rho)} \left[\sum_{y \in z + \square_n^+} \widetilde{\eta}_y = M \right] \sum_{N=1}^{M} P_{z+\square_n, M, N} \left(\frac{|\square_n^+|}{2N} \right)^2 F(N) \\
&= \sum_{N=1}^{\infty} \left(\sum_{M \geqslant N} \widetilde{\mathbb{P}}_{\alpha(\rho)} \left[\sum_{y \in z + \square_n^+} \widetilde{\eta}_y = M \right] \widetilde{\mathbb{P}}_{z+\square_n^+, M} \left[\sum_{x \in (z + \square_n^+) \setminus \{z\}} [\widetilde{\eta}]_x = N - 1 \right] \right) \left(\frac{|\square_n^+|}{2N} \right)^2 F(N)
\end{aligned}$$

$$= \sum_{N=1}^{\infty} \widetilde{\mathbb{P}}_{\alpha(\rho)} \left[\sum_{x \in (z+\square_n^+) \setminus \{z\}} [\widetilde{\eta}]_x = N-1 \right] \left(\frac{|\square_n^+|}{2N} \right)^2 F(N)$$
$$= \sum_{N=1}^{\infty} \mathbb{P}_{\rho} \left[\sum_{x \in (z+\square_n^+) \setminus \{z\}} \eta_x = N-1 \right] \left(\frac{|\square_n^+|}{2N} \right)^2 F(N)$$

The identity from the third line to the forth line comes from the definition of conditional probability. From the forth line to the fifth line, we apply the coarse-grained lifting in Proposition 3.1 once again. Using the notation of $\mathbf{N}_{z,n}$, the last term becomes $\left\langle \left(\frac{|\square_n^+|}{2\mathbf{N}_{z,n}} \right)^2 F(\mathbf{N}_{z,n}) \right\rangle_{\rho}$, and it gives us

$$\tag{3.18} \left\langle\!\!\left\langle (\mathsf{S}_n(\partial_i[u])(\widetilde{\eta}, z))^2 \right\rangle\!\!\right\rangle_{\alpha(\rho)} \leqslant \left\langle \left(\frac{|\square_n^+|}{2\mathbf{N}_{z,n}} \right)^2 \left(\frac{1}{|\square_n|} \sum_{x \in z+\square_n} \langle (\nabla_{x,e_i} u) \rangle_{z+\square_n^+, \mathbf{N}_{z,n}} \right)^2 \right\rangle_{\rho}.$$

This result is as expected, since we retract the spatial average under the Kawasaki dynamics with a weight of particle numbers. □

4. Subadditive quantities

In this section, we define several subadditive quantities and develop their elementary properties.

4.1. Subadditive quantities $\overline{\nu}$ and $\overline{\nu}_*$.

For every finite set $\Lambda \subseteq \mathbb{Z}^d$ and $p, q \in \mathbb{R}^d$, we define the quantities

$$\tag{4.1} \begin{aligned} \overline{\nu}(\rho, \Lambda, p) &:= \inf_{v \in \ell_{p,\Lambda^+} + \mathcal{F}_0(\Lambda^-)} \left\{ \frac{1}{2\chi(\rho)|\Lambda|} \sum_{b \in \overline{\Lambda^*}} \left\langle \frac{1}{2} c_b (\pi_b v)^2 \right\rangle_{\rho} \right\}, \\ \overline{\nu}_*(\rho, \Lambda, q) &:= \sup_{v \in \mathcal{F}_0} \left\{ \frac{1}{2\chi(\rho)|\Lambda|} \sum_{b \in \overline{\Lambda^*}} \left\langle (\pi_b \ell_q)(\pi_b v) - \frac{1}{2} c_b (\pi_b v)^2 \right\rangle_{\rho} \right\}. \end{aligned}$$

Recall that the affine function ℓ_{p,Λ^+} defined in (1.45), and Λ^-, $\overline{\Lambda^*}$ defined respectively in (1.34), (1.35). Compared to (1.20), here we add the factor $\chi(\rho)$ in the normalization, which will make the notation lighter in the homogenization step.

We record some elementary properties satisfied by $\overline{\nu}$ and $\overline{\nu}_*$.

Proposition 4.1 (Elementary properties of $\overline{\nu}$ and $\overline{\nu}_*$)**.** *The following properties hold for every bounded $\Lambda \subseteq \mathbb{Z}^d$ and $p, p', q, q' \in \mathbb{R}^d$.*

(1) There exists a unique solution for the optimization problem of $\overline{\nu}(\rho, \Lambda, p)$ satisfying $\langle v - \ell_{p,\Lambda^+} \rangle_{\rho} = 0$; we denote it by $v(\cdot, \rho, \Lambda, p)$. For the optimization problem of $\overline{\nu}_*(\rho, \Lambda.q)$, there exists a unique maximizer $u(\cdot, \Lambda, q)$ being independent of ρ and belonging to $\mathcal{F}_0(N_{\mathbf{r}}(\Lambda^+))$ satisfying $\mathbb{E}_{\rho}[u|\mathcal{G}_{\Lambda^+}] = 0$, where $\mathcal{G}_{\Lambda^+}$ is defined in (3.14) and $N_{\mathbf{r}}(\Lambda^+)$ is defined as

$$N_{\mathbf{r}}(\Lambda^+) := \{x \in \mathbb{Z}^d : \mathrm{dist}(x, \Lambda^+) \leqslant \mathbf{r}\}.$$

Moreover, the two optimizers are both harmonic in the sense $v(\cdot, \rho, \Lambda, p), u(\cdot, \Lambda, q) \in \mathcal{A}(\Lambda)$, where $\mathcal{A}(\Lambda)$ is defined in (1.41).

(2) There exist two $d \times d$ symmetric matrices $\overline{D}(\rho, \Lambda)$ and $\overline{D}_*(\rho, \Lambda)$ such that for every $p, q \in \mathbb{R}^d$

$$\overline{\nu}(\rho, \Lambda, p) = \frac{1}{2} p \cdot \overline{D}(\rho, \Lambda) p, \qquad \overline{\nu}_*(\rho, \Lambda, q) = \frac{1}{2} q \cdot \overline{D}_*^{-1}(\rho, \Lambda) q \tag{4.2}$$

and these matrices satisfy $\mathsf{Id} \leqslant \overline{D}(\rho, \Lambda), \overline{D}_*(\rho, \Lambda) \leqslant \lambda \mathsf{Id}$. Moreover, for ever $p', q' \in \mathbb{R}^d$, we have

$$\begin{aligned} p' \cdot \overline{D}(\rho, \Lambda) p &= \left\langle \frac{1}{2\chi(\rho)|\Lambda|} \sum_{b \in \overline{\Lambda^*}} c_b (\pi_b \ell_{p'})(\pi_b v(\cdot, \rho, \Lambda, p)) \right\rangle_\rho, \\ q' \cdot \overline{D}_*^{-1}(\rho, \Lambda) q &= \left\langle \frac{1}{2\chi(\rho)|\Lambda|} \sum_{b \in \overline{\Lambda^*}} (\pi_b \ell_{q'})(\pi_b u(\cdot, \Lambda, q)) \right\rangle_\rho . \end{aligned} \tag{4.3}$$

(3) For every $v' \in \ell_{p,\Lambda^+} + \mathcal{F}_0(\Lambda^-)$, we have

$$\left\langle \frac{1}{2\chi(\rho)|\Lambda|} \sum_{b \in \overline{\Lambda^*}} \frac{1}{2} c_b (\pi_b (v - v'))^2 \right\rangle_\rho = \left\langle \frac{1}{2\chi(\rho)|\Lambda|} \sum_{b \in \overline{\Lambda^*}} \frac{1}{2} c_b (\pi_b v')^2 \right\rangle_\rho - \overline{\nu}(\rho, \Lambda, p), \tag{4.4}$$

where $v = v(\cdot, \rho, \Lambda, p)$ is the minimizer defined in (1).

Similarly, for every $u' \in \mathcal{F}_0$ and the maximiser $u = u(\cdot, \Lambda, q)$ defined in (1), we have

$$\begin{aligned} &\left\langle \frac{1}{2\chi(\rho)|\Lambda|} \sum_{b \in \overline{\Lambda^*}} \frac{1}{2} c_b (\pi_b (u - u'))^2 \right\rangle_\rho \\ &\qquad = \overline{\nu}_*(\rho, \Lambda, q) - \left\langle \frac{1}{2\chi(\rho)|\Lambda|} \sum_{b \in \overline{\Lambda^*}} \left(-\frac{1}{2} c_b (\pi_b u')^2 + (\pi_b \ell_q)(\pi_b u') \right) \right\rangle_\rho . \end{aligned} \tag{4.5}$$

(4) For any partition of verticies $\Lambda = \bigsqcup_{i=1}^m \Lambda_i$, we have

$$\overline{\nu}(\rho, \Lambda, p) \leqslant \sum_{i=1}^m \frac{|\Lambda_i|}{|\Lambda|} \overline{\nu}(\rho, \Lambda_i, p), \tag{4.6}$$

$$\overline{\nu}_*(\rho, \Lambda, p) \leqslant \sum_{i=1}^m \frac{|\Lambda_i|}{|\Lambda|} \overline{\nu}_*(\rho, \Lambda_i, p). \tag{4.7}$$

In particular, we have $\overline{\nu}(\rho, \square_{n+1}, q) \leqslant \overline{\nu}(\rho, \square_n, q)$ and $\overline{\nu}_*(\rho, \square_{n+1}, q) \leqslant \overline{\nu}_*(\rho, \square_n, q)$ for every $n \in \mathbb{N}_+$.

Proof. The proof of this proposition is elementary and standard. We give the complete proof of (1), which concerns some details. For other statements, we sketch the main idea of the proof. Readers may look for details in a similar setting in [35, Proposition 4.1].

(1) We first study $\overline{\nu}(\rho, \Lambda, p)$. By a variational calculus, the minimizer of the equation can be characterized by the equation that for any $\phi \in \mathcal{F}_0(\Lambda^-)$,

$$\left\langle \sum_{b \in \overline{\Lambda^*}} c_b (\pi_b (v - \ell_{p,\Lambda^+}))(\pi_b \phi) \right\rangle_\rho = \left\langle \sum_{b \in \overline{\Lambda^*}} c_b (\pi_b (-\ell_{p,\Lambda^+}))(\pi_b \phi) \right\rangle_\rho . \tag{4.8}$$

We can define its solution in the space

$$V = \{ f \in \mathcal{F}_0(\Lambda^-) : \langle f \rangle_\rho = 0 \}.$$

The coercivity is ensured by Poincaré inequality (2.5) and we can apply Lax–Milgram theorem to get the unique minimizer $v(\cdot, \rho, \Lambda, p)$. Moreover, (4.8) implies that $v(\cdot, \rho, \Lambda, p) \in \mathcal{A}(\Lambda)$.

Then we turn to $\overline{\nu}_*(\rho,\Lambda,q)$. A first observation is that the maximizer can be found in $\mathcal{F}_0(N_{\mathbf{r}}(\Lambda^+))$. Because for any $u\in\mathcal{F}_0$, its conditional expectation $\mathsf{A}_{N_{\mathbf{r}}(\Lambda^+)}u$ (defined in (2.9)) reaches a larger value for the functional. More precisely,

$$\begin{aligned}
&\left\langle \sum_{b\in\overline{\Lambda^*}}\left(-\frac{1}{2}c_b(\pi_b\mathsf{A}_{N_{\mathbf{r}}(\Lambda^+)}u)^2+(\pi_b\ell_q)(\pi_b\mathsf{A}_{N_{\mathbf{r}}(\Lambda^+)}u)\right)\right\rangle_\rho\\
&=\left\langle \sum_{b\in\overline{\Lambda^*}}\left(-\frac{1}{2}c_b(\mathsf{A}_{N_{\mathbf{r}}(\Lambda^+)}\pi_b u)^2+(\pi_b\ell_q)(\pi_b\mathsf{A}_{N_{\mathbf{r}}(\Lambda^+)}u)\right)\right\rangle_\rho\\
&\geqslant\left\langle \sum_{b\in\overline{\Lambda^*}}\left(-\frac{1}{2}c_b(\pi_b u)^2+(\pi_b\ell_q)(\pi_b u)\right)\right\rangle_\rho,
\end{aligned}\tag{4.9}$$

where we use the locality of c_b and (2.11) from the first line to the second line, and Jensen's inequality from the second line to the third line.

Similar to the discussion above, the maximizer of the functional can be characterized by the variational equation that for any $\phi\in\mathcal{F}_0$,

$$\left\langle \sum_{b\in\overline{\Lambda^*}} c_b(\pi_b u)(\pi_b\phi)\right\rangle_\rho=\left\langle \sum_{b\in\overline{\Lambda^*}}(\pi_b\ell_q)(\pi_b\phi)\right\rangle_\rho\tag{4.10}$$

Notice that for any $\mathcal{G}_{\Lambda^+}$ measurable function ϕ', the difference $\pi_b\phi'$ becomes zero and thus the above equation automatically holds. By replacing ϕ by $\phi-\mathbb{E}_\rho[\phi|\mathcal{G}_{\Lambda^+}]$ we may only solve the problem for all ϕ in the space $W=\{f\in\mathcal{F}_0:\mathbb{E}_\rho[f|\mathcal{G}_{\Lambda^+}]=0\}$. Moreover, testing the equation with $\phi\mathbf{1}_{\{\sum_{x\in\Lambda^+}\eta_x=N\}}\mathbf{1}_{\{\eta_y=\varepsilon_y,\text{ for finitely many } y\in(\Lambda^+)^c\}}$ for arbitrary $N\in\mathbb{N}$ and $\varepsilon_y\in\{0,1\}$, we acturally reinforce the equation (4.10) into a stronger way

$$\mathbb{E}_\rho\left[\sum_{b\in\Lambda^*} c_b(\pi_b u)(\pi_b\phi)\mid\mathcal{G}_{\Lambda^+}\right]=\mathbb{E}_\rho\left[\sum_{b\in\Lambda^*}(\pi_b\ell_q)(\pi_b\phi)\mid\mathcal{G}_{\Lambda^+}\right].\tag{4.11}$$

and we may seek for the solution in the space W. In this space the Poincaré inequality (2.8) ensures the coercivity, so Lax–Milgram theorem applies and we get the unique maximizer $u(\cdot,\Lambda,q)$. We emphasize that after taking conditional expectation with respect to proper $\mathcal{G}_{\Lambda^+}$, the equation (4.11) is independent of the choice of ρ. Thus the optimizer is found separately for every conditional expectation $\mathbb{E}[\cdot|\mathcal{G}_{\Lambda^+}]$ and in particular, u is a optimizer for all ρ. Moreover, testing (4.10) with $\phi\in\mathcal{F}_0(\Lambda^-)$, the right hand side becomes zero thanks to (3.9) and we get $u\in\mathcal{A}(\Lambda)$.

(2) We test (4.8) with $v(\cdot,\rho,\Lambda,p')-\ell_{p',\Lambda^+}$ and get

$$\left\langle \sum_{b\in\overline{\Lambda^*}} c_b(\pi_b v(\cdot,\rho,\Lambda,p))(\pi_b v(\cdot,\rho,\Lambda,p'))\right\rangle_\rho=\left\langle \sum_{b\in\overline{\Lambda^*}} c_b(\pi_b v(\cdot,\rho,\Lambda,p))(\pi_b\ell_{p',\Lambda^+})\right\rangle_\rho,\tag{4.12}$$

which gives the linearity. To obtain the bound of $\overline{D}(\rho,\Lambda)$, we use the condition $1\leqslant c_b\leqslant\lambda$ in Hypothesis 1.1,

$$\begin{aligned}
&\inf_{v\in\ell_{p,\Lambda}+\mathcal{F}_0(\Lambda^-)}\left\langle\frac{1}{4\chi(\rho)|\Lambda|}\sum_{b\in\overline{\Lambda^*}}(\pi_b v)^2\right\rangle_\rho\\
&\leqslant\inf_{v\in\ell_{p,\Lambda}+\mathcal{F}_0(\Lambda^-)}\left\langle\frac{1}{4\chi(\rho)|\Lambda|}\sum_{b\in\overline{\Lambda^*}}c_b(\pi_b v)^2\right\rangle_\rho=\frac{1}{2}p\cdot\overline{D}(\rho,\Lambda)\cdot p\\
&\leqslant\inf_{v\in\ell_{p,\Lambda}+\mathcal{F}_0(\Lambda^-)}\left\langle\frac{1}{4\chi(\rho)|\Lambda|}\sum_{b\in\overline{\Lambda^*}}\lambda(\pi_b v)^2\right\rangle_\rho.
\end{aligned}$$

We can observe that ℓ_{p,Λ^+} is the minimizer for $\inf_{v \in \ell_{p,\Lambda^+} + \mathcal{F}_0(\Lambda^-)} \left\langle \frac{1}{4\chi(\rho)|\Lambda|} \sum_{b \in \Lambda^*} (\pi_b v)^2 \right\rangle_\rho$, whose energy is precisely p^2, so this proves the first part of the proposition.

The similar argument works for $\overline{\nu}_*(\rho, \Lambda, q)$. Testing (4.10) with $u(\cdot, \Lambda, q')$ yields the linearity. Concerning the bound for $\overline{D}_*$, we use the bound for c_b to obtain:

$$\begin{aligned}
&\sup_{v \in \mathcal{F}_0} \frac{1}{2\chi(\rho)|\Lambda|} \left\langle -\frac{\lambda}{2} \sum_{b \in \Lambda^*} (\pi_b v)^2 + \sum_{b \in \Lambda^*} (\pi_b \ell_{q,\Lambda^+})(\pi_b v) \right\rangle_\rho \\
&\leqslant \overline{\nu}_*(\rho, \Lambda, q) \\
&\leqslant \sup_{v \in \mathcal{F}_0} \frac{1}{2\chi(\rho)|\Lambda|} \left\langle -\frac{1}{2} \sum_{b \in \Lambda^*} (\pi_b v)^2 + \sum_{b \in \Lambda^*} (\pi_b \ell_{q,\Lambda^+})(\pi_b v) \right\rangle_\rho .
\end{aligned}$$

One can see in the lower bound, the maximizer is $\ell_{\frac{q}{\lambda},\Lambda^+}$, while in the upper bound, the maximizer is ℓ_{q,Λ^+}. This gives the bound we want.

(3) This is a direct calculation. We test (4.8) with $(v' - \ell_{p,\Lambda^+})$ to get the first equation, and test the equation (4.10) with u' to get the second equation.

(4) For the quantity $\overline{\nu}(\rho, \Lambda, p)$, we use

$$v' = \ell_{p,\Lambda^+} + \sum_{i=1}^{n} (v(\cdot, \rho, \Lambda_i, p) - \ell_{p,\Lambda_i^+})$$

is a sub-minimizer of $\overline{\nu}(\rho, \Lambda, p)$ to prove the subadditivity of $\overline{\nu}$. Concerning the quantity $\overline{\nu}_*(\rho, \Lambda, q)$, we can not "glue" the local optimizers, but we use the fact that $u(\cdot, \Lambda, q)$ is a sub-maximizer for every $\overline{\nu}_*(\rho, \Lambda, q)$ to get the subadditivity of $\overline{\nu}_*$. We also highlight the identity (1.37), which helps avoid the boundary layer and is the motivation we use $\overline{\Lambda^*}$ in the definition (4.1).

□

Remark 4.2. Intuitively, the last part of the proof of item (1) actually says that the maximizer $u(\cdot, \Lambda, q)$ can be found in the following way: first, fix the number of particles inside Λ^+ (denoted by n) and the environment outside Λ^+ (denoted by ζ), and seek for a maximizer $u_{n,\zeta}$ in this condition. Then the maximizer is exactly obtained by pasting all the $u_{n,\zeta}$'s.

Viewing the subadditivity (4.6), we have the following corollary.

Corollary 4.3. *For every $\rho \in (0,1)$, the following limit is well-defined*

$$\overline{D}(\rho) := \lim_{m \to \infty} \overline{D}(\rho, \square_m). \tag{4.13}$$

Recall the L^∞ norm over $\mathcal{X}$ defined as $\|F\|_\infty := \sup_{\eta \in \mathcal{X}} |F(\eta)|$. Here we give an upper bound estimate of $v(\cdot, \rho, \Lambda, p)$ and $u(\cdot, \Lambda, q)$ defined in Proposition 4.1.

Lemma 4.4 (L^∞ estimate)**.** *For any connected domain Λ of diameter L and $p, q \in B_1$, there exists a constant $C(\lambda, d)$ such that*

$$\|v(\cdot, \rho, \Lambda, p)\|_\infty + \|u(\cdot, \Lambda, q)\|_\infty \leqslant C L^{d+2} \log L. \tag{4.14}$$

The proof relies on the mixing time of our non-gradient dynamic; see Appendix B. Note that for $1 \leqslant p < \infty$, we have a better estimate that

$$\|v(\cdot, \rho, \Lambda, p)\|_p + \|u(\cdot, \Lambda, q)\|_p \leqslant C L^{d+2}.$$

See Remark B.3 for details.

4.2. **Master quantities J.** We continue to explore the dual property between $\overline{\nu}$ and $\overline{\nu}_*$ in this subsection. For every bonded $\Lambda \subseteq \mathbb{Z}^d$ and $p, q \in \mathbb{R}^d$, we define the quantities

$$J(\rho, \Lambda, p, q) := \overline{\nu}(\rho, \Lambda, p) + \overline{\nu}_*(\rho, \Lambda, q) - p \cdot q. \tag{4.15}$$

We first describe J with a variational problem.

Lemma 4.5. *(1) For each $p, q \in \mathbb{R}^d$, we have the variational representation*

$$J(\rho, \Lambda, p, q) := \sup_{w \in \mathcal{A}(\Lambda)} J(\rho, \Lambda, p, q; w), \tag{4.16}$$

where $J(\rho, \Lambda, p, q; w)$ is a functional defined as

$$\begin{aligned} &J(\rho, \Lambda, p, q; w) \\ &:= \left\langle \frac{1}{2\chi(\rho)|\Lambda|} \sum_{b \in \overline{\Lambda^*}} \left(-\frac{1}{2} c_b (\pi_b w)^2 - c_b (\pi_b \ell_p)(\pi_b w) + (\pi_b \ell_q)(\pi_b w) \right) \right\rangle_\rho . \end{aligned} \tag{4.17}$$

One maximizer is $(u(\cdot, \Lambda, q) - v(\cdot, \rho, \Lambda, p))$ with v, u defined in (1) of Proposition 4.1.

(2) The master quantity J is always positive, i.e. $J(\rho, \Lambda, p, q) \geqslant 0$.

Proof. (1) In the following paragraph, we use $v(\cdot, \rho, \Lambda, p)$ to denote the minimizer in the definition of $\overline{\nu}$ and write $v = v(\cdot, \rho, \Lambda, p)$ for short. Since we have deduced that the maximizer of $\overline{\nu}_*$ can be found in $\mathcal{A}(\Lambda)$ in (1) of Proposition 4.1, we may write

$$\begin{aligned} J(\rho, \Lambda, p, q) = &\left\langle \frac{1}{4\chi(\rho)|\Lambda|} \sum_{b \in \overline{\Lambda^*}} c_b (\pi_b v)^2 \right\rangle_\rho \\ &+ \sup_{u \in \mathcal{A}(\Lambda)} \left\langle \frac{1}{2\chi(\rho)|\Lambda|} \sum_{b \in \overline{\Lambda^*}} \left(-\frac{1}{2} c_b (\pi_b u)^2 + (\pi_b \ell_q)(\pi_b u) \right) \right\rangle_\rho \\ &- p \cdot q. \end{aligned}$$

Since $(v - \ell_{p,\Lambda^+}) \in \mathcal{F}_0(\Lambda^-)$, we have

$$\left\langle \sum_{b \in \overline{\Lambda^*}} (\pi_b (v - \ell_{p,\Lambda^+}))(\pi_b \ell_{q,\Lambda^+}) \right\rangle_\rho = \sum_{x \in \Lambda} \langle q \cdot \nabla_x (v - \ell_{p,\Lambda^+}) \rangle_\rho = 0 \tag{4.18}$$

Here the first equality comes from (1.46), (1.48), (1.49) and the second equality comes from (3.9). Moreover, for any $u \in \mathcal{A}(\Lambda)$, by its definition (1.41) and noting $(v - \ell_{p,\Lambda^+}) \in \mathcal{F}_0(\Lambda^-)$ again, we have

$$\left\langle \sum_{b \in \overline{\Lambda^*}} c_b (\pi_b u)(\pi_b v) \right\rangle_\rho = \left\langle \sum_{b \in \overline{\Lambda^*}} c_b (\pi_b u)(\pi_b \ell_{p,\Lambda^+}) \right\rangle_\rho$$

and in particular this equation is true when taking $u = v$.

Combining these results, we obtain

$$\begin{aligned} &J(\rho, \Lambda, p, q) \\ &= \sup_{u \in \mathcal{A}(\Lambda)} \Bigg(\left\langle \frac{1}{4\chi(\rho)|\Lambda|} \sum_{b \in \overline{\Lambda^*}} c_b (\pi_b v)^2 \right\rangle_\rho + \left\langle \frac{1}{2\chi(\rho)|\Lambda|} \sum_{b \in \overline{\Lambda^*}} \left(-\frac{1}{2} c_b (\pi_b u)^2 + (\pi_b \ell_q)(\pi_b u) \right) \right\rangle_\rho \\ &\quad - \left\langle \frac{1}{2\chi(\rho)|\Lambda|} \sum_{b \in \overline{\Lambda^*}} (\pi_b \ell_p)(\pi_b \ell_q) \right\rangle_\rho \Bigg) \\ &= \sup_{u \in \mathcal{A}(\Lambda)} \left\langle \frac{1}{2\chi(\rho)|\Lambda|} \sum_{b \in \overline{\Lambda^*}} \left(-\frac{1}{2} c_b (\pi_b (u - v))^2 - c_b (\pi_b \ell_p) \pi_b (u - v) + (\pi_b \ell_q) \pi_b (u - v) \right) \right\rangle_\rho \end{aligned}$$

$$= \sup_{w \in \mathcal{A}(\Lambda)} \left\langle \frac{1}{2\chi(\rho)|\Lambda|} \sum_{b \in \overline{\Lambda^*}} \left(-\frac{1}{2} c_b (\pi_b w)^2 - c_b(\pi_b \ell_p)(\pi_b w) + (\pi_b \ell_q)(\pi_b w) \right) \right\rangle_\rho .$$

From the definition of $u(\cdot, \Lambda, q)$, we conclude that $w = u(\cdot, \Lambda, q) - v(\cdot, \rho, \Lambda, p)$ is exactly a maximizer.

(2) We test the functional in the definition of $\overline{\nu}_*(\rho, \Lambda, q)$ with the minimizer $v = v(\cdot, \rho, \Lambda, p)$ of $\overline{\nu}(\rho, \Lambda, p)$ and obtain

$$\begin{aligned} \overline{\nu}_*(\rho, \Lambda, q) &\geqslant \left\langle \frac{1}{2\chi(\rho)|\Lambda|} \sum_{b \in \overline{\Lambda^*}} \left(-\frac{1}{2} c_b (\pi_b v)^2 + (\pi_b \ell_q)(\pi_b v) \right) \right\rangle_\rho \\ &= -\overline{\nu}(\rho, \Lambda, p) + \left\langle \frac{1}{2\chi(\rho)|\Lambda|} \sum_{b \in \overline{\Lambda^*}} (\pi_b \ell_q)(\pi_b \ell_p) \right\rangle_\rho \\ &= -\overline{\nu}(\rho, \Lambda, p) + p \cdot q, \end{aligned}$$

which proves $J(\rho, \Lambda, p, q) \geqslant 0$. Here from the first line to the second line, we also use the identity (4.18). □

Now we explain why J is convenient in estimating the convergence rate.

Lemma 4.6. *(1) For any bounded $\Lambda \subseteq \mathbb{Z}^d$ and $\rho \in (0,1)$, we have $\overline{D}(\rho, \Lambda) \geqslant \overline{D}_*(\rho, \Lambda)$.*

(2) There exists a constant $C(d, \lambda)$ such that for every symmetric matrix $\widetilde{D}$ satisfying $|\xi|^2 \leqslant \xi \cdot \widetilde{D}\xi \leqslant \lambda |\xi|^2$, we have

$$|\widetilde{D} - \overline{D}(\rho, \Lambda)| + |\widetilde{D} - \overline{D}_*(\rho, \Lambda)| \leqslant C \sup_{|p|=1} J(\rho, \Lambda, p, \widetilde{D}p)^{\frac{1}{2}} \tag{4.19}$$

Proof. (1) Recall $J(\rho, \Lambda, p, q) \geqslant 0$ from (2) of Lemma 4.5, and insert $q = \overline{D}_*(\rho, \Lambda)p$ to obtain

$$\begin{aligned} 0 \leqslant & J(\rho, \Lambda, p, \overline{D}_*(\rho, \Lambda)p) \\ = & \frac{1}{2} p \cdot \overline{D}(\rho, \Lambda)p + \frac{1}{2} p \cdot \overline{D}_*(\rho, \Lambda)p - p \cdot \overline{D}_*(\rho, \Lambda)p \\ = & \frac{1}{2} p \cdot \overline{D}(\rho, \Lambda)p - \frac{1}{2} p \cdot \overline{D}_*(\rho, \Lambda)p, \end{aligned}$$

so we have $\overline{D}(\rho, \Lambda) \geqslant \overline{D}_*(\rho, \Lambda)$.

(2) Using the property $\overline{D}(\rho, \Lambda) \geqslant \overline{D}_*(\rho, \Lambda)$, we have

$$\begin{aligned} J(\rho, \Lambda, p, q) = & \frac{1}{2} p \cdot \overline{D}(\rho, \Lambda)p + \frac{1}{2} q \cdot \overline{D}_*^{-1}(\rho, \Lambda)q - p \cdot q \\ \geqslant & \frac{1}{2} p \cdot \overline{D}(\rho, \Lambda)p + \frac{1}{2} q \cdot \overline{D}^{-1}(\rho, \Lambda)q - p \cdot q \\ = & \frac{1}{2} (\overline{D}(\rho, \Lambda)p - q) \cdot \overline{D}^{-1}(\rho, \Lambda)(\overline{D}(\rho, \Lambda)p - q). \end{aligned}$$

Setting $|p| = 1$ and $q = \widetilde{D}p$, we conclude that

$$|\overline{D}(\rho, \Lambda) - \widetilde{D}| \leqslant C \sup_{|p|=1} J(\rho, \Lambda, p, \widetilde{D}p)^{\frac{1}{2}}.$$

The proof of the statement concerning $|\overline{D}_*(\rho, \Lambda) - \widetilde{D}|$ is similar. □

This lemma allows us to control the convergence rate of $\overline{D}(\rho, \square_m)$ by showing the convergence rate of $J(\rho, \square_m, p, \overline{D}_*(\rho, \square_m)p)$. Finally, we resume some more properties of J similar to the properties for $\overline{\nu}$ and $\overline{\nu}_*$.

Proposition 4.7 (Elementary properties of J)**.** *For every $\rho \in (0,1)$, every bounded $\Lambda \in \mathbb{Z}^d$ and every $p, q \in \mathbb{R}^d$, the quantity $J(\rho, \Lambda, p, q)$ defined as (4.15) satisfies the following properties:*

(1) First order variation and optimizer: the optimization problem in (4.16) admits a unique solution $v(\cdot, \rho, \Lambda, p, q) \in \mathcal{A}(\Lambda)$ such that $\mathbb{E}_\rho[v(\cdot, \rho, \Lambda.p, q)|\mathcal{G}_{\Lambda^+}] = 0$, which can be expressed in terms of the optimizers of $\overline{\nu}$ and $\overline{\nu}_*$ as

$$v(\cdot, \rho, \Lambda, p, q) = u(\cdot, \Lambda, q) - v(\cdot, \rho, \Lambda, q) - \mathbb{E}_\rho[u(\cdot, \Lambda, q) - v(\cdot, \rho, \Lambda, p)|\mathcal{G}_{\Lambda^+}]. \tag{4.20}$$

This solution $v(\cdot, \rho, \Lambda, p, q)$ satisfies that for every $w \in \mathcal{A}(\Lambda)$,

$$\left\langle \sum_{b \in \overline{\Lambda^*}} (c_b \pi_b v(\cdot, \rho, \Lambda, p, q))(\pi_b w) \right\rangle_\rho = \left\langle \sum_{b \in \overline{\Lambda^*}} (-c_b(\pi_b \ell_p)(\pi_b w) + (\pi_b \ell_q)(\pi_b w)) \right\rangle_\rho, \tag{4.21}$$

and the mapping $(p, q) \mapsto v(\cdot, \rho, \Lambda, p, q)$ is linear.

(2) Quadratic response: we have an quadratic expression for J

$$J(\rho, \Lambda, p, q) = \left\langle \frac{1}{4\chi(\rho)|\Lambda|} \sum_{b \in \overline{\Lambda^*}} c_b (\pi_b v(\cdot, \rho, \Lambda, p, q))^2 \right\rangle_\rho. \tag{4.22}$$

For every $w \in \mathcal{A}(\Lambda)$, with the functional defined in (4.17), we have

$$\left\langle \frac{1}{4\chi(\rho)|\Lambda|} \sum_{b \in \overline{\Lambda^*}} c_b (\pi_b (w - v(\cdot, \rho, \Lambda, p, q)))^2 \right\rangle_\rho = J(\rho, \Lambda, p, q) - J(\rho, \Lambda, p, q; w). \tag{4.23}$$

(3) For any partition of verticies $\Lambda = \bigsqcup_{i=1}^m \Lambda_i$, we have

$$J(\rho, \Lambda, p, q) \leqslant \sum_{i=1}^m \frac{|\Lambda_i|}{|\Lambda|} J(\rho, \Lambda_i, p, q). \tag{4.24}$$

(4) Slope property: with the gradient operator ∇_x defined in (1.48), the optimizer $v(\cdot, \rho, \Lambda, p, q)$ satisfies the slop property

$$\left\langle \frac{1}{2\chi(\rho)|\Lambda|} \sum_{x \in \Lambda} \nabla_x v(\cdot, \rho, \Lambda, p, q) \right\rangle_\rho = \overline{D}_*^{-1}(\rho, \Lambda) q - p. \tag{4.25}$$

Proof. (1) The equation (4.21) comes directly from the first order variation calculus. The proof of existance and uniqueness of the solution $v(\cdot, \rho, \Lambda, p, q)$ is similar to the one for $\overline{\nu}_*(\rho, \Lambda, q)$.

One can check directly that $v(\cdot, \rho, \Lambda, p_1, q_1) + v(\cdot, \rho, \Lambda, p_2, q_2)$ is the solution for the problem (4.21) with parameter $(p_1 + p_2, q_1 + q_2)$ and it also satisfies the conditional expectation condition, so we have $v(\cdot, \rho, \Lambda, p_1 + p_2, q_1 + q_2) = v(\cdot, \rho, \Lambda, p_1, q_1) + v(\cdot, \rho, \Lambda, p_2, q_2)$ by the uniqueness, which implies $(p, q) \mapsto v(\cdot, \rho, \Lambda, p, q)$ is linear.

The exact expression (4.20) follows directly from the fact that $(u(\cdot, \Lambda, q) - v(\cdot, \rho, \Lambda, p))$ is a maximizer in (4.16), and we add the necessary regularization condition.

(2) We put $v = v(\cdot, \rho, \Lambda, p, q)$ in the first order variation (4.21) to get

$$\left\langle \sum_{b \in \Lambda^*} (c_b(\pi_b v)^2 + c_b(\pi_b \ell_p)(\pi_b v) - (\pi_b \ell_{q, \Lambda^+})(\pi_b v)) \right\rangle_\rho = 0.$$

Then we plug this into (4.16) to get the quadratic expression. The quadratic response (4.23) follows directly from (4.5) by testing with $u = w - v(\cdot, \rho, \Lambda, p)$.

(3) This is a consequence of (4.6) and (4.7).

(4) Using the exact expression of $v(\cdot,\rho,\Lambda,p,q)$ in (4.20), we have

$$\begin{aligned}&\left\langle \frac{1}{2\chi(\rho)|\Lambda|}\sum_{x\in\Lambda}\nabla_x v(\cdot,\rho,\Lambda,p,q)\right\rangle_\rho \\ &= \left\langle \frac{1}{2\chi(\rho)|\Lambda|}\sum_{x\in\Lambda}\nabla_x u(\cdot,\Lambda,q)\right\rangle_\rho - \left\langle \frac{1}{2\chi(\rho)|\Lambda|}\sum_{x\in\Lambda}\nabla_x v(\cdot,\rho,\Lambda,q)\right\rangle_\rho \\ &= \overline{D}_*^{-1}(\rho,\Lambda)q - p.\end{aligned}$$

Here in the second line, we apply (4.3) to the first term and (4.18) to the second term. □

5. Renormalization under grand canonical ensemble

Based on the diffusion matrix $\overline{D}(\rho,\Lambda), \overline{D}_*(\rho,\Lambda)$ defined in Proposition 4.1 and using the Einstein relation (1.10), we define the conductivity and its dual quantity on $(0,1)$ as

$$\bar{\mathbf{c}}(\rho,\Lambda) \coloneqq 2\chi(\rho)\overline{D}(\rho,\Lambda), \qquad \bar{\mathbf{c}}_*(\rho,\Lambda) \coloneqq 2\chi(\rho)\overline{D}_*(\rho,\Lambda). \tag{5.1}$$

By default, when $\rho\in\{0,1\}$, we set $\bar{\mathbf{c}}(\rho,\Lambda) = \bar{\mathbf{c}}_*(\rho,\Lambda) = 0$. Especially, $\bar{\mathbf{c}}(\rho,\Lambda)$ coincides with the definition (1.20). Then Corollary 4.3 also defines a limit

$$\bar{\mathbf{c}}(\rho) \coloneqq 2\chi(\rho)\overline{D}(\rho) = \lim_{m\to\infty}\bar{\mathbf{c}}(\rho,\square_m). \tag{5.2}$$

In this section, we are ready to prove the convergence rate.

Proposition 5.1. *There exists an exponent $\gamma_1(d,\lambda,\mathbf{r}) > 0$ and a positive constant $C(d,\lambda,\mathbf{r}) < \infty$ such that for every $L\in\mathbb{N}_+$,*

$$\sup_{\rho\in[0,1]}\left(|\bar{\mathbf{c}}(\rho,\Lambda_L) - \bar{\mathbf{c}}(\rho)| + |\bar{\mathbf{c}}_*(\rho,\Lambda_L) - \bar{\mathbf{c}}(\rho)|\right) \leqslant CL^{-\gamma_1}. \tag{5.3}$$

Although Proposition 5.1 is stated for $\bar{\mathbf{c}}(\rho,\Lambda_L)$ and $\bar{\mathbf{c}}_*(\rho,\Lambda_L)$, we will still rely on $\overline{D}(\rho,\Lambda)$ and $\overline{D}_*(\rho,\Lambda)$ in the intermediate steps as they are already normalized; see (2) of Proposition 4.1. Note that $\chi(\rho)$ in $\bar{\mathbf{c}}(\rho,\Lambda_L)$ helps to prove the uniform convergence in $\rho\in[0,1]$. Throughout this section, we define the following shorthand expression

$$D_n \coloneqq \overline{D}_*(\rho,\square_n). \tag{5.4}$$

By the subadditive quantity (4.6) and (4.7), we know that $\overline{D}(\rho,\square_m)$ and $\overline{D}_*^{-1}(\rho,\square_m)$ are decreasing. Therefore, it suffices to show the convergence rate for $|\overline{D}_*(\rho,\square_m) - \overline{D}(\rho,\square_m)|$, which is reduced to the decay rate of $\sup_{|p|=1} J(\rho,\square_n,p,D_m p)$ after applying (4.19) with $\widetilde{D} = D_m$.

In this section, we will heavily use the master quantity $J(\rho,\Lambda,p,q)$ defined in (4.16) and its optimizer $v(\cdot,\rho,\Lambda,p,q)$ defined in (1) of Proposition 4.7. The notation $\mathcal{Z}_{m,n} = 3^n\mathbb{Z}^d\cap\square_m$ is usually involved to make comparison between J in different scales. We will also use the following gap of the master quantities in the proof very often

$$\tau_n \coloneqq \sup_{p,q\in B_1}\left(J(\rho,\square_n,p,q) - J(\rho,\square_{n+1},p,q)\right), \tag{5.5}$$

where $B_1 \coloneqq \{p\in\mathbb{R}^d : |p| = 1\}$.

Our first lemma makes use of the elliptic regularity, especially the modified Caccioppoli inequality (2.10), in our master quantity.

Lemma 5.2. *There exists a finite positive constant $C(d,\lambda)$ such that, for every $m \in \mathbb{N}_+$ satisfying $3^m > R_0$ for the constant in Proposition 2.6, we have the following estimate*

$$J(\rho, \square_m, p, D_m p)^{\frac{1}{2}} \leqslant C\tau_m^{\frac{1}{2}} + C3^{-m} \left\langle \frac{1}{2\chi(\rho)|\square_{m+1}|} v(\cdot, \rho, \square_{m+1}, p, D_m p)^2 \right\rangle_\rho^{\frac{1}{2}}. \tag{5.6}$$

Proof. We write $v_m = v(\cdot, \rho, \square_m, p, D_m p)$ as a shortened version in the proof. Recall the operator A_L defined in (2.9), and denote by

$$L := 3^m + 2\mathbf{r}, \tag{5.7}$$

as the length of localization in this proof. Then, by (4.22), we decompose our quantity $J(\rho, \square_m, p, D_m p)$ as

$$\begin{aligned} &J(\rho, \square_m, p, D_m p)^{\frac{1}{2}} \leqslant (\mathbf{I})^{\frac{1}{2}} + (\mathbf{II})^{\frac{1}{2}}, \\ &\mathbf{I} := \left\langle \frac{1}{4\chi(\rho)|\square_m|} \sum_{b \in \overline{\square^*_m}} c_b(\pi_b(v_m - \mathsf{A}_L v_{m+1}))^2 \right\rangle_\rho, \\ &\mathbf{II} := \left\langle \frac{1}{4\chi(\rho)|\square_m|} \sum_{b \in \overline{\square^*_m}} c_b(\pi_b \mathsf{A}_L v_{m+1})^2 \right\rangle_\rho. \end{aligned} \tag{5.8}$$

We justify at first $\mathsf{A}_L v_{m+1} \in \mathcal{A}(\square_m)$ by testing an arbitrary function $f \in \mathcal{F}_0(\square^-_m)$

$$\begin{aligned} \sum_{b \in \overline{\square^*_m}} \langle c_b(\pi_b \mathsf{A}_L v_{m+1})(\pi_b f) \rangle_\rho &= \sum_{b \in \overline{\square^*_m}} \langle \mathsf{A}_L \left(c_b(\pi_b v_{m+1})(\pi_b f) \right) \rangle_\rho \\ &= \sum_{b \in \overline{\square^*_m}} \langle c_b(\pi_b v_{m+1})(\pi_b f) \rangle_\rho = 0. \end{aligned}$$

Here in the first line, since c_b and f are both $\mathscr{F}_{\Lambda_L}$-measurable, they commute with A_L operator. In the second line, we use the fact that $v_{m+1} \in \mathcal{A}(\square_{m+1}) \subseteq \mathcal{A}(\square_m)$ from (1) of Proposition 4.7.

We now turn to the two terms in (5.8). As we have proved $\mathsf{A}_L v_{m+1} \in \mathcal{A}(\square_m)$, we apply (4.23) to $\mathbf{I}$ and get

$$\mathbf{I} = J(\rho, \square_m, p, D_m p) - J(\rho, \square_m, p, D_m p; \mathsf{A}_L v_{m+1}). \tag{5.9}$$

We investigate the expression of $J(\rho, \square_m, p, D_m p; \mathsf{A}_L v_{m+1})$ by (4.17). By Jensen's inequality, we have

$$\left\langle \frac{1}{2\chi(\rho)|\square_m|} \sum_{b \in \overline{\square^*_m}} \frac{1}{2} c_b(\pi_b \mathsf{A}_L v_{m+1})^2 \right\rangle_\rho \leqslant \left\langle \frac{1}{2\chi(\rho)|\square_m|} \sum_{b \in \overline{\square^*_m}} \frac{1}{2} c_b(\pi_b v_{m+1})^2 \right\rangle_\rho,$$

and by the definition of conditional expectation, we also have

$$\begin{aligned} &\left\langle \frac{1}{2\chi(\rho)|\square_m|} \sum_{b \in \overline{\square^*_m}} (c_b(\pi_b \ell_p)(\pi_b \mathsf{A}_L v_{m+1}) - (\pi_b \ell_{D_m p})(\pi_b \mathsf{A}_L v_{m+1})) \right\rangle_\rho \\ &= \left\langle \frac{1}{2\chi(\rho)|\square_m|} \sum_{b \in \overline{\square^*_m}} (c_b(\pi_b \ell_p)(\pi_b v_{m+1}) - (\pi_b \ell_{D_m p})(\pi_b v_{m+1})) \right\rangle_\rho. \end{aligned}$$

These imply that $J(\rho, \square_m, p, D_m p; \mathsf{A}_L v_{m+1}) \geqslant J(\rho, \square_m, p, D_m p; v_{m+1})$. Putting it back to (5.9) and using (4.23) again, we have

$$\mathbf{I} \leqslant J(\rho, \square_m, p, D_m p) - J(\rho, \square_m, p, D_m p; v_{m+1})$$

$$= \left\langle \frac{1}{4\chi(\rho)|\square_m|} \sum_{b\in\square_m^*} c_b(\pi_b(v_m - v_{m+1}))^2 \right\rangle_\rho .$$

We claim that the right hand side of the inequality can be controlled by $C\tau_m$ defined in (5.5) with some constant only depending on the dimension d. It suffices to apply (4.23) to $z + \square_m$ with $z \in \mathcal{Z}_{m+1,m}$. We denote by $v_{m,z} = v(\cdot, \rho, z + \square_m, p, D_m p)$ to simplify the notation in the following calculation

$$\begin{aligned}
&\left\langle \frac{1}{4\chi(\rho)|\square_m|} \sum_{b\in\square_m^*} c_b(\pi_b(v_m - v_{m+1}))^2 \right\rangle_\rho \\
&\leqslant \sum_{z\in\mathcal{Z}_{m+1,m}} \left\langle \frac{1}{4\chi(\rho)|\square_m|} \sum_{b\in(z+\square_m)^*} c_b(\pi_b(v_{m,z} - v_{m+1}))^2 \right\rangle_\rho \\
&= \sum_{z\in\mathcal{Z}_{m+1,m}} (J(\rho, z+\square_m, p, D_m p) - J(\rho, z+\square_m, p, D_m p; v_{m+1})) \\
&= 3^d(J(\rho, \square_m, p, D_m p) - J(\rho, \square_{m+1}, p, D_m p)).
\end{aligned}$$

This concludes that

$$\mathbf{I} \leqslant 3^d \lambda \tau_m. \tag{5.10}$$

Concerning **II**, as we have proved $\mathsf{A}_L v_{m+1} \in \mathcal{A}(\square_m)$ and $3^m > R_0$ in Proposition 2.6, we can apply the modified Caccioppoli inequality (2.10) to get

$$\mathbf{II} \leqslant \left\langle \frac{C3^{-2m}}{2\chi(\rho)|\square_{m+1}|} v_{m+1}^2 \right\rangle_\rho + \theta J(\rho, \square_{m+1}, p, D_m p). \tag{5.11}$$

Here the factor $\theta \in (0,1)$ and the constant C all come from Proposition 2.6, and only dependent on d, λ.

Combining the two estimates (5.10) and (5.11), we have

$$\begin{aligned}
&J(\rho, \square_m, p, D_m p)^{\frac12} \\
&\qquad \leqslant \theta^{\frac12} J(\rho, \square_m, p, D_m p)^{\frac12} + C3^{-m} \left\langle \frac{1}{2\chi(\rho)|\square_{m+1}|} v(\square_{m+1})^2 \right\rangle_\rho^{\frac12} + C\tau_m^{\frac12}.
\end{aligned}$$

Using the fact $\theta \in (0,1)$, we rearrange the expression above and conclude (5.6). $\square$

5.1. **Flatness estimate.** This part gives a flatness estimate and this is the main challenge compared to its counterpart in [35, Lemma 5.1]. The whole Section 3 is devoted to overcome the technical difficulty here.

Proposition 5.3 (L^2-flatness estimate)**.** *There exists an exponent $\beta(d) \in (0, \frac14)$ and a constant $C(d, \lambda, \mathbf{r}) < \infty$ such that for every $p, q \in B_1$ and $m \in \mathbb{N}$,*

$$\begin{aligned}
&\frac{1}{2\chi(\rho)|\square_{m+1}|} \left\langle \left(v(\rho, \square_{m+1}, p, q) - \ell_{D_m^{-1}q - p, \square_{m+1}}\right)^2 \right\rangle_\rho \\
&\qquad \leqslant C\alpha(\rho) 3^{2m} \left(\rho^{-2} \left(3^{-\beta m} + \sum_{n=0}^{m} 3^{-\beta(m-n)} \tau_n \right) + \chi^{-2}(\rho) 3^{(4d+6)m} e^{-\frac{\rho^2 3^{\frac{m}{4}}}{2}} \right).
\end{aligned} \tag{5.12}$$

Proof. In the proof, we write v_{m+1} as a shorthand of $v(\rho, \square_{m+1}, p, q)$. We apply the weighted multiscale Poincaré inequality Proposition 3.9 to obtain

$$\left\langle \frac{1}{2\chi(\rho)|\square_{m+1}|}\left(v_{m+1} - \ell_{D_m^{-1}q-p,\square_{m+1}}\right)^2 \right\rangle_\rho^{\frac{1}{2}} \leqslant C\alpha^{\frac{1}{2}}(\rho)\sum_{n=1}^{m} 3^n \left(\frac{1}{|\mathcal{Z}_{m+1,n}|}\sum_{z\in\mathcal{Z}_{m+1,n}}\left\langle\left(\frac{|\square_n^+|}{2\mathbf{N}_{z,n}}\right)^2 I_{z,n}\right\rangle_\rho\right)^{\frac{1}{2}}, \tag{5.13}$$

where the term $I_{z,n}$ is the shorthand of

$$I_{z,n} := \frac{1}{2\chi(\rho)}\left|\frac{1}{|\square_n|}\sum_{x\in z+\square_n}\left\langle\nabla_x\left(v_{m+1} - \ell_{D_m^{-1}q-p}\right)\right\rangle_{z+\square_n^+,\mathbf{N}_{z,n}}\right|^2,$$

and recall (1.48) for ∇_x. We need to treat the terms $I_{z,n}$ and study at first the case of large scales.

Case 1.1: large scale $n > \frac{m}{2(d+1)}$; high density $\mathbf{N}_{z,n} \geqslant \frac{1}{2}\rho|\square_n^+|$. For this case, the weight is of typical size and does not degenerate, so we have

$$\begin{aligned}
&\left\langle\left(\frac{|\square_n^+|}{2\mathbf{N}_{z,n}}\right)^2 I_{z,n}\mathbf{1}_{\{\mathbf{N}_{z,n}\geqslant\frac{1}{2}\rho|\square_n^+|\}}\right\rangle_\rho \\
&\leqslant \left\langle\left(\frac{1}{\rho}\right)^2 I_{z,n}\mathbf{1}_{\{\mathbf{N}_{z,n}\geqslant\frac{1}{2}\rho|\square_n^+|\}}\right\rangle_\rho \\
&\leqslant \left(\frac{1}{\rho}\right)^2\langle I_{z,n}\rangle_\rho.
\end{aligned} \tag{5.14}$$

Here we use the fact $I_{z,n} \geqslant 0$, so we can drop the indicator in the third line. Then we use the comparison of Dirichlet energy with $v_{n,z} = v(\rho, z+\square_n, p, q)$

$$\begin{aligned}
\langle I_{z,n}\rangle_\rho \leqslant{}& \frac{3}{2\chi(\rho)}\left\langle\left|\frac{1}{|\square_n|}\sum_{x\in z+\square_n}\left\langle\nabla_x\left(v_{n,z} - \ell_{D_n^{-1}q-p}\right)\right\rangle_{z+\square_n^+,\mathbf{N}_z}\right|^2\right\rangle_\rho \\
&+ \frac{3}{2\chi(\rho)}\left\langle\left|\frac{1}{|\square_n|}\sum_{x\in z+\square_n}\left\langle\nabla_x\left(v_{n,z} - v_{m+1}\right)\right\rangle_{z+\square_n^+,\mathbf{N}_z}\right|^2\right\rangle_\rho \\
&+ 3|D_m^{-1}q - D_n^{-1}q|^2.
\end{aligned} \tag{5.15}$$

We will estimate the first term by Lemma 5.4 below, which studies the decay of variance

$$\begin{aligned}
&\frac{1}{2\chi(\rho)}\left\langle\left|\frac{1}{|\square_n|}\sum_{x\in z+\square_n}\left\langle\nabla_x\left(v_{n,z} - \ell_{D_n^{-1}q-p}\right)\right\rangle_{z+\square_n^+,\mathbf{N}_z}\right|^2\right\rangle_\rho \\
&\leqslant \left\langle\frac{1}{2\chi(\rho)}\left|\frac{1}{|\square_n|}\sum_{x\in z+\square_n}\nabla_x\left(v_{n,z} - \ell_{D_n^{-1}q-p}\right)\right|^2\right\rangle_\rho \\
&\leqslant C3^{-\beta n} + C\sum_{k=0}^{n-1}3^{-\beta(n-k)}\tau_k.
\end{aligned}$$

Here from the first line to the second line, we use Jensen's inequality. Notice that the exponent β only depends on d and the constant C depends on $\lambda, \mathbf{r}, d$ as stated in Lemma 5.4.

The second term in (5.15) can be controlled by the gap of the subadditive quantity using (4.23) as the proof of Lemma 5.6

$$
\begin{aligned}
&\frac{1}{|\mathcal{Z}_{m+1,n}|}\sum_{z\in\mathcal{Z}_{m+1,n}}\frac{1}{2\chi(\rho)}\left\langle\left|\frac{1}{|\square_n|}\sum_{x\in z+\square_n}\langle\nabla_x(v_{n,z}-v_{m+1})\rangle_{z+\square_n^+,\mathbf{N}_z}\right|^2\right\rangle_\rho\\
&\leqslant\frac{1}{|\mathcal{Z}_{m+1,n}|}\sum_{z\in\mathcal{Z}_{m+1,n}}\frac{1}{2\chi(\rho)}\left\langle\frac{1}{|\square_n|}\sum_{x\in z+\square_n}|\nabla_x(v_{n,z}-v_{m+1})|^2\right\rangle_\rho\\
&\leqslant J(\rho,\square_n,p,q)-J(\rho,\square_{m+1},p,q)\\
&\leqslant\sum_{k=n}^{m}\tau_k.
\end{aligned}
$$

The third term in (5.15) is also naturally bounded

$$
|D_m^{-1}q-D_n^{-1}q|^2\leqslant C(\lambda)\sum_{k=n}^{m-1}\tau_k.
$$

Combine these three terms and take in account of the factor ρ^{-2} from (5.14), we obtain

$$
\begin{aligned}
\text{(5.16)}\quad&\frac{1}{|\mathcal{Z}_{m+1,n}|}\sum_{z\in\mathcal{Z}_{m+1,n}}\left\langle\left(\frac{|\square_n^+|}{2\mathbf{N}_{z,n}}\right)^2 I_{z,n}\mathbf{1}_{\{\mathbf{N}_{z,n}\geqslant\frac{1}{2}\rho|\square_n^+|\}}\right\rangle_\rho\\
&\qquad\leqslant C\rho^{-2}\left(3^{-\beta n}+\sum_{k=0}^{n-1}3^{-\beta(n-k)}\tau_k+\sum_{k=n}^{m}\tau_k\right).
\end{aligned}
$$

Case 1.2: large scale $n>\frac{m}{2(d+1)}$*; low density* $1\leqslant\mathbf{N}_{z,n}<\frac{1}{2}\rho|\square_n^+|$. For this case, we will get a very large factor from the weight using the trivial bound

$$
\left(\frac{|\square_n^+|}{2\mathbf{N}_{z,n}}\right)^2\leqslant 3^{2dn}. \tag{5.17}
$$

However, this case is quite rare by Hoeffding inequality (see [18, Theorem 2.8])

$$
\mathbb{P}_\rho\left[\mathbf{N}_{z,n}<\frac{\rho|\square_n^+|}{2}\right]\leqslant\exp\left(-\frac{\rho^2|\square_n^+|}{2}\right).
$$

We also make use of the L^∞ estimate from Lemma 4.4, and obtain

$$
\begin{aligned}
&\left\langle\left(\frac{|\square_n^+|}{2\mathbf{N}_{z,n}}\right)^2 I_{z,n}\mathbf{1}_{\{1\leqslant\mathbf{N}_{z,n}<\frac{1}{2}\rho|\square_n^+|\}}\right\rangle_\rho\leqslant\chi^{-2}(\rho)3^{2dn}\|v_{m+1}\|_\infty^2\,\mathbb{P}_\rho\left[\mathbf{N}_{z,n}<\frac{\rho|\square_n^+|}{2}\right]\\
\text{(5.18)}\quad&\qquad\leqslant C(\lambda,d)\chi^{-2}(\rho)3^{2dn}3^{2(d+3)m}\exp\left(-\frac{\rho^2|\square_n^+|}{2}\right)\\
&\qquad\leqslant C(\lambda,d)\chi^{-2}(\rho)3^{2dn}3^{2(d+3)m}\exp\left(-\frac{\rho^2 3^{\frac{m}{4}}}{2}\right).
\end{aligned}
$$

In the last line, we use the condition $n>\frac{m}{2(d+1)}$ to bound the rare probability.

Now we turn to the case of small scales.

Case 2: small scale $0\leqslant n\leqslant\frac{m}{2(d+1)}$. For this case, we cannot expect too much spatial average cancellation, while the concentration effect is not strong enough to beat the L^∞ norm like in (5.18). On the other hand, the factors from the Poincaré inequality and the weight is still not so large. Therefore, we use Jensen's inequality and the trivial bound (5.17), and pay directly the price of the volume $|\square_n^+|$

$$\left\langle \left(\frac{|\square_n^+|}{2\mathbf{N}_{z,n}}\right)^2 I_{z,n}\right\rangle_\rho \leqslant 3^{2dn}\left\langle \frac{1}{2\chi(\rho)}\frac{1}{|\square_n|}\sum_{x\in z+\square_n}|\nabla_x v_{m+1}-\ell_{D_m^{-1}q-p}|^2\right\rangle_\rho .$$

The contribution from all these scales are

$$\begin{aligned}
&\alpha^{\frac12}(\rho)\sum_{n=0}^{\lfloor\frac{m}{2(d+1)}\rfloor}3^n\left(\frac{1}{|\mathcal{Z}_{m+1,n}|}\sum_{z\in\mathcal{Z}_{m+1,n}}\left\langle\left(\frac{|\square_n^+|}{2\mathbf{N}_{z,n}}\right)^2 I_{z,n}\right\rangle_\rho\right)^{\frac12}\\
&\leqslant \alpha^{\frac12}(\rho)\sum_{n=0}^{\lfloor\frac{m}{2(d+1)}\rfloor}3^{(d+1)n}\left\langle\frac{1}{2\chi(\rho)}\frac{1}{|\square_{m+1}|}\sum_{x\in\square_{m+1}}|\nabla_x v_{m+1}-\ell_{D_m^{-1}q-p}|^2\right\rangle_\rho^{\frac12}\\
&\leqslant C(\lambda)\alpha^{\frac12}(\rho)3^{\frac{m}{2}}.
\end{aligned}\tag{5.19}$$

Here from the second line to the third line, we just use the Dirichlet energy estimate of $\nabla_x v_{m+1}$ in (2) of Proposition 4.1.

Plugging (5.16), (5.18) and (5.19) into (5.13), we have

$$\begin{aligned}
&\left\langle\frac{1}{2\chi(\rho)|\square_{m+1}|}(v_{m+1}-\ell_{D_m^{-1}q-p,\square_{m+1}})^2\right\rangle_\rho^{\frac12}\\
&\leqslant C(\lambda)\alpha^{\frac12}(\rho)\left(3^{\frac{m}{2}}+\sum_{n=\lfloor\frac{m}{2(d+1)}\rfloor}^{m}3^n\left(\rho^{-2}\left(3^{-\beta n}+\sum_{k=0}^{n-1}3^{-\beta(n-k)}\tau_k+\sum_{k=n}^{m}\tau_k\right)\right.\right.\\
&\qquad\left.\left.+\chi^{-2}(\rho)3^{2dn}3^{2(d+3)m}\exp\left(-\frac{\rho^2 3^{\frac{m}{4}}}{2}\right)\right)^{\frac12}\right).
\end{aligned}$$

Square this equation with Cauchy–Schwartz inequality, and shrink β, then we prove (5.12).

□

5.2. **Variance decay of the averaged gradient.** In this part, we prove the variance decay of gradient. Since its proof only uses the spatial independence, all its parameters are independent of ρ. Especially, the decay exponent β only depends on the dimension d.

Lemma 5.4 (Variance decay)**.** *There exist $\beta(d)>0$ and $C(d,\lambda,\mathbf{r})<\infty$ such that for every $p,q\in B_1$ and $n\in\mathbb{N}$, we have*

$$\frac{1}{2\chi(\rho)}\left\langle\left|\frac{1}{|\square_n|}\sum_{x\in\square_n}\nabla_x\left(v(\rho,\square_n,p,q)-\ell_{D_n^{-1}q-p}\right)\right|^2\right\rangle_\rho \leqslant C3^{-\beta n}+C\sum_{k=0}^{n-1}3^{-\beta(n-k)}\tau_k.\tag{5.20}$$

Proof. Since ρ,p,q does not change, we write v_n as a shorthand for $v(\rho,\square_n,p,q)$ and $v_{n-1,z}$ as a shorthand of $v(\rho,z+\square_{n-1},p,q)$. We use a comparison between scale n and scale $(n-1)$

$$\left\langle\frac{1}{2\chi(\rho)}\left|\frac{1}{|\square_n|}\sum_{x\in\square_n}\nabla_x\left(v_n-\ell_{D_n^{-1}q-p}\right)\right)^2\right\rangle_\rho^{\frac12}$$

$$
\begin{aligned}
&\leqslant \left\langle \frac{1}{2\chi(\rho)} \left| \frac{1}{|\square_n|} \sum_{z \in \mathcal{Z}_{n,n-1}} \sum_{x \in z + \square_{n-1}} \nabla_x \left(v_{n-1,z} - \ell_{D_{n-1}^{-1} q - p} \right) \right|^2 \right\rangle_\rho^{\frac{1}{2}} \\
&\quad + \left\langle \frac{1}{2\chi(\rho)} \left| \frac{1}{|\square_n|} \sum_{z \in \mathcal{Z}_{n,n-1}} \sum_{x \in z + \square_{n-1}} \nabla_x \left(v_{n-1,z} - v_n \right) \right|^2 \right\rangle_\rho^{\frac{1}{2}} \\
&\quad + \lambda |D_n^{-1} q - D_{n-1}^{-1} q|.
\end{aligned}
$$

Here the last term comes from the integration of affine function, and $\chi(\rho)|\square_n|$ is the correct factor to normalize the integration. We deal with the three terms separately. The third term is naturally bounded

$$
|D_n^{-1} q - D_{n-1}^{-1} q| \leqslant C(d, \lambda) \tau_{n-1}^{\frac{1}{2}}.
$$

For the second term, we use Jensen's inequality and (4.23) to obtain

$$
\begin{aligned}
&\left\langle \frac{1}{2\chi(\rho)} \left| \frac{1}{|\square_n|} \sum_{z \in \mathcal{Z}_{n,n-1}} \sum_{x \in z + \square_{n-1}} \nabla_x \left(v_{n-1,z} - v_n \right) \right|^2 \right\rangle_\rho \\
&\leqslant \left\langle \frac{1}{2\chi(\rho)} \frac{1}{|\square_n|} \sum_{z \in \mathcal{Z}_{n,n-1}} \sum_{x \in z + \square_{n-1}} \left| \nabla_x \left(v_{n-1,z} - v_n \right) \right|^2 \right\rangle_\rho \\
&\leqslant \tau_{n-1}.
\end{aligned}
$$

For the first term, we define

$$
X_z := \frac{1}{|\square_{n-1}|} \sum_{x \in z + \square_{n-1}} \nabla_x \left(v(z + \square_{n-1}, p, q)) - \ell_{D_{n-1}^{-1} q - p} \right),
$$

and expand the square

$$
\begin{aligned}
&\left\langle \left| \frac{1}{|\square_n|} \sum_{z \in \mathcal{Z}_{n,n-1}} \sum_{x \in z + \square_{n-1}} \nabla_x \left(v_{n-1,z} - \ell_{D_{n-1}^{-1} q - p} \right) \right|^2 \right\rangle_\rho \\
&= \left\langle \left| \frac{1}{|\mathcal{Z}_{n,n-1}|} \sum_{z \in \mathcal{Z}_{n,n-1}} X_z \right|^2 \right\rangle_\rho \\
&= \frac{1}{|\mathcal{Z}_{n,n-1}|^2} \sum_{z, w \in \mathcal{Z}_{n,n-1}} \langle X_z \cdot X_w \rangle_\rho .
\end{aligned}
$$

Note that X_z is a $\mathbb{R}^d$-valued random vector, and here $X_z \cdot X_w$ is the inner product between X_z and X_w.

For sufficiently large n such that $3^n > 10\mathbf{r}$, there exist two cubes $z + \square_{n-1}$ and $w + \square_{n-1}$ with distance greater than $2\mathbf{r}$. Recall that $X_z \in \mathcal{F}_0(N_{\mathbf{r}}(z + \square_n^+))$ from (1) of Proposition 4.1. Then we can use the local property and independence for such pair X_z, X_w to get

$$
\langle X_z \cdot X_w \rangle_\rho = \langle X_z \rangle_\rho \cdot \langle X_w \rangle_\rho = 0,
$$

where the last equal sign follows from the average slope property (4.25). For other pairs, we just use Cauchy–Schwartz inequality and stationary property, which concludes

$$
\left\langle \left| \frac{1}{|\mathcal{Z}_{n,n-1}|} \sum_{z \in \mathcal{Z}_{n,n-1}} X_z \right|^2 \right\rangle_\rho \leqslant \frac{3^{2d} - 1}{3^{2d}} \left\langle |X_0|^2 \right\rangle_\rho .
$$

If we denote left hand side of (5.20) by σ_n^2, then our calculation above gives

$$\sigma_n \leqslant \theta\sigma_{n-1} + C\tau_{n-1}^{\frac{1}{2}} \tag{5.21}$$

for some positive constant $\theta := \frac{3^{2d}-1}{3^{2d}} < 1$. This holds for all $n \geqslant n_0$ such that $3^{n_0} > 10\mathbf{r}$, and we do iteration to obtain

$$\begin{aligned}\forall n > n_0, \qquad \sigma_n &\leqslant \theta^{n-n_0}\sigma_{n_0} + C(d,\lambda)\sum_{k=n_0}^{n-1}\theta^{n-k}\tau_k^{\frac{1}{2}} \\ &\leqslant C(d,\lambda,\mathbf{r})\left(\theta^n + \sum_{k=0}^{n-1}\theta^{n-k}\tau_k^{\frac{1}{2}}\right).\end{aligned}$$

Here we also use the trivial bound $\sigma_0 \leqslant \lambda$, and shift the constant. Since there are only finite cases $1 \leqslant n \leqslant n_0$, we square the equation above and prove (5.20). $\square$

5.3. **Iterations and track of parameters.** In this part, we resume the previous steps and conclude the proof. Especially, here we need to track the dependence of the density ρ in our convergence. Recall (5.1), and combine the result (4.19) with $\widetilde{D} = D_m$ defined in (5.4), we obtain

$$\begin{aligned}&|\overline{\mathbf{c}}(\rho,\square_m) - \overline{\mathbf{c}}_*(\rho,\square_m)| \\ &= 2\chi(\rho)\left|\overline{D}(\rho,\square_m) - \overline{D}_*(\rho,\square_m)\right| \\ &\leqslant C(d,\lambda)\chi(\rho)\left(\sup_{|p|=1} J(\rho,\square_m,p,D_m p)^{\frac{1}{2}}\right).\end{aligned} \tag{5.22}$$

Therefore, we need to study the uniform convergence of $\chi^2(\rho)J(\rho,\square_m,p,D_m p)$, where the compresssibility $\chi(\rho)$ helps control the convergence near two endpoints. We state the following lemma, which controls $\chi^2(\rho)J(\rho,\square_m,p,D_m p)$ by the energy gap τ_n defined in (5.5) uniformly in ρ.

Lemma 5.5. *There exist an exponent $\kappa(d) > 0$ and a constant $C(d,\lambda,\mathbf{r}) < \infty$ such that for every $p \in B_1$ and $\rho \in (0,1)$, we have*

$$\chi^2(\rho)J(\rho,\square_m,p,D_m p) \leqslant C\left(3^{-\kappa m} + \sum_{n=0}^{m} 3^{-\kappa(m-n)}\tau_n\right). \tag{5.23}$$

Proof. We combine the estimates (5.6) and (5.12), which yield

$$\begin{aligned}&J(\rho,\square_m,p,D_m p) \\ &\leqslant C\tau_m + C\alpha(\rho)\left(\rho^{-2}\left(3^{-\beta m} + \sum_{n=0}^{m}3^{-\beta(m-n)}\tau_n\right) + \chi^{-2}(\rho)3^{(4d+6)m}e^{-\frac{\rho^2 3^{\frac{m}{4}}}{2}}\right).\end{aligned} \tag{5.24}$$

This estimate will explode when fixing m and sending ρ to 0 or 1, which is even worse than the trivial bound $0 \leqslant J(\rho,\square_m,p,D_m p) \leqslant 3\lambda$ for all $p \in B_1$; see (4.15) and (2) of Proposition 4.1. Therefore, we combine these two estimates by an interpolation. Let $s \in (1,\infty)$ be an exponent to be determined later, we have

$$\begin{aligned}\chi^2(\rho)J(\rho,\square_m,p,D_m p) &\leqslant \chi^2(\rho)J^{\frac{1}{s}}(\rho,\square_m,p,D_m p)(3\lambda)^{1-\frac{1}{s}} \\ &\leqslant 3\lambda\left(\chi^{2s}(\rho)J(\rho,\square_m,p,D_m p)\right)^{\frac{1}{s}}.\end{aligned} \tag{5.25}$$

Then we insert the expression (5.24). For the case $s \geqslant 2$, we have

$$\chi^{2s}(\rho)J(\rho,\square_m,p,D_m p)$$

$$\leqslant C\chi^{2s}(\rho)\alpha(\rho)\left(\rho^{-2}\left(3^{-\beta m}+\sum_{n=0}^{m}3^{-\beta(m-n)}\tau_n\right)+\chi^{-2}(\rho)3^{(4d+6)m}e^{-\frac{\rho^2 3^{\frac{m}{4}}}{2}}\right)$$
$$\leqslant C\chi^{2s-2}(\rho)\alpha(\rho)\left(\left(3^{-\beta m}+\sum_{n=0}^{m}3^{-\beta(m-n)}\tau_n\right)+3^{(4d+6)m}e^{-\frac{\rho^2 3^{\frac{m}{4}}}{2}}\right).$$

In the last line, we pay $\chi^2(\rho)$ to compensate the singularity from ρ^{-2} and $\chi^{-2}(\rho)$. Then we study the condition to balance the other term involving ρ.

- Concerning $\alpha(\rho)$, we pose a condition $s \geqslant 3$. Then the factor $(1-\rho)^{2s-2}$ from $\chi^{2s-2}(\rho)$ dominates $\alpha(\rho) = -\log(1-\rho)$ because
$$(1-\rho)^{2s-2}\alpha(\rho) \leqslant -(1-\rho)\log(1-\rho) \leqslant e^{-1},$$
and $-x\log x$ takes its maximum on $[0,1]$ at $x = e^{-1}$.
- Concerning $3^{(4d+6)m}e^{-\frac{\rho^2 3^{\frac{m}{4}}}{2}}$, we pose a condition $s \geqslant 16d+26$. Then the factor ρ^{2s-2} from $\chi^2(\rho)$ can improve the term of concentration estimate
$$\begin{aligned}\rho^{2s-2}3^{(4d+6)m}e^{-\frac{\rho^2 3^{\frac{m}{4}}}{2}} &\leqslant \rho^{32d+50}3^{(4d+6)m}e^{-\frac{\rho^2 3^{\frac{m}{4}}}{2}}\\ &\leqslant \rho^2\left(\rho^2 3^{\frac{m}{4}}\right)^{16d+24}e^{-\frac{\rho^2 3^{\frac{m}{4}}}{2}}.\end{aligned}$$
We use the fact that $x^{16d+24}e^{-\frac{x}{2}} \leqslant C(d)x^{-1}$ on $\mathbb{R}_+$ and the condition $\beta \in (0,\frac{1}{4})$ from Proposition 5.3, then we obtain
$$\rho^{2s-2}3^{(4d+6)m}e^{-\frac{\rho^2 3^{\frac{m}{4}}}{2}} \leqslant \rho^2\left(\rho^2 3^{\frac{m}{4}}\right)^{-1} \leqslant 3^{-\frac{m}{4}} \leqslant 3^{-\beta m}.$$

Thus we set $s = 16d+26$, and put these estimates back to (5.25), which yields

$$\begin{aligned}\chi^2(\rho)J(\rho,\square_m,p,D_m p) &\leqslant C\left(3^{-\beta m}+\sum_{n=0}^{m}3^{-\beta(m-n)}\tau_n\right)^{\frac{1}{16d+26}}\\ &\leqslant C\left(3^{-\frac{\beta m}{16d+26}}+\sum_{n=0}^{m}3^{-\frac{\beta m}{16d+26}}\tau_n\right).\end{aligned}$$

Here we use $(a+b)^{\frac{1}{n}} \leqslant a^{\frac{1}{n}}+b^{\frac{1}{n}}$ for $a,b>0$ and $n \in \mathbb{N}_+$. By setting $\kappa := \frac{\beta}{16d+26}$, we obtain the desired result (5.23). □

Proof of Proposition 5.1. The remaining part is similar to the proof of [11, Proposition 2.11] and we give its sketch. We define

$$F_m := \sum_{i=1}^{d}\chi^2(\rho)J(\rho,\square_m,e_i,D_m e_i), \tag{5.26}$$

and its weighted version with $\kappa > 0$ from Lemma 5.5.

$$\widetilde{F}_m := \sum_{n=0}^{m}3^{-\frac{\kappa}{2}(m-n)}F_n.$$

Since (5.5) and (5.23) all satisfy the estimate independent of ρ, it suffices to follow the iteration in [11, Proposition 2.11], and there exists a constant $C(d,\lambda,\mathbf{r}) < \infty$ such that

$$\widetilde{F}_{m+1} \leqslant C(\widetilde{F}_m-\widetilde{F}_{m+1})+C3^{-\frac{\kappa m}{2}}.$$

This implies a contraction

$$\widetilde{F}_{m+1} \leqslant \left(\frac{C}{1+C}\right)\widetilde{F}_m+\left(\frac{C}{1+C}\right)3^{-\frac{\kappa m}{2}}.$$

Following (5.22) and (5.26), there exists $\gamma_1(d,\lambda,\mathbf{r})>0$ and $C(d,\lambda,\mathbf{r})<\infty$ such that

$$\forall \rho \in [0,1], \qquad |\overline{\mathbf{c}}(\rho,\square_m)-\overline{\mathbf{c}}_*(\rho,\square_m)| \leqslant F_m \leqslant \widetilde{F}_m \leqslant C3^{-\gamma_1 m}. \tag{5.27}$$

Recall the monotone convergence in Corollary 4.3 for a fixed ρ, thus as $m\to\infty$, the sequence $\overline{\mathbf{c}}(\rho,\square_m)$ decreases and $\overline{\mathbf{c}}_*(\rho,\square_m)$ increases to the same limit $\overline{\mathbf{c}}(\rho)$. Then (5.27) gives the desired uniform convergence rate along the triadic cubes, and Lemma A.1 extends the result to a general cube, which concludes the proof. □

6. Density-free local corrector with uniform convergence

There still remains some difference between our homogenization result in Proposition 5.1 and Theorem 1.2. The optimizer F_L in Theorem 1.2 is a corrector-type function in homogenization, and we have already obtained a natural candidate from previous sections (see (4.1) and (1) of Proposition 4.1)

$$\phi_{\rho,\Lambda,\xi} := v(\rho,\Lambda,\xi)-\ell_{\Lambda^+,\xi}. \tag{6.1}$$

This local corrector is of $\mathcal{F}_0(\Lambda^-)$, but has the dependence on ρ. In this part, we explore various properties about this function at first in Sections 6.1, 6.2, then improve it by removing the dependence of density in Section 6.3, and finally prove the main theorem (Theorem 1.2) in Section 6.4 and 6.5.

6.1. **Regularity of local corrector.**

We resume at first some properties about our local corrector $\phi_{\rho,\Lambda,\xi}$ from previous sections.

Proposition 6.1. *The local corrector $\phi_{\rho,\Lambda,\xi}$ satisfies the following properties.*

(1) *(Elementary properties) $\phi_{\rho,\Lambda,\xi}$ is a $\mathcal{F}_0(\Lambda^-)$ function with $\langle\phi_{\rho,\Lambda,\xi}\rangle_\rho=0$, and $\xi\mapsto\phi_{\rho,\Lambda,\xi}$ is linear.*

(2) *(Approximation of conductivity) For every $L\in\mathbb{N}_+$, we have*

$$\sup_{\rho\in[0,1],\xi\in B_1}\left|\frac{1}{|\Lambda_L|}\left\langle\sum_{b\in\Lambda_L^*}\frac{1}{2}c_b(\pi_b(\ell_\xi+\phi_{\rho,\Lambda_L,\xi}))^2\right\rangle_\rho-\frac{1}{2}\xi\cdot\overline{\mathbf{c}}(\rho)\xi\right|\leqslant CL^{-\gamma_1}. \tag{6.2}$$

Here the exponent $\gamma_1(d,\lambda,\mathbf{r})$ and the constant $C(d,\lambda,\mathbf{r})$ are same as in Proposition 5.1.

(3) *(Sublinearity) There exists a constant $C(d,\lambda,\mathbf{r})<\infty$, such for every $L\in\mathbb{N}_+$, the following estimate holds for $\gamma_1(d,\lambda,\mathbf{r})$ from Proposition 5.1*

$$\sup_{\rho\in[0,1],\xi\in B_1}\left\langle\frac{1}{|\Lambda_L|}\phi^2_{\rho,\Lambda_L,\xi}\right\rangle_\rho\leqslant CL^{2-\gamma_1}. \tag{6.3}$$

Proof. (1) can be deduced from Proposition 4.1. (2) is the consequence from Proposition 5.1. (3) comes from Proposition 5.3, once we put the convergence rate (5.3) to (5.5). □

The main task of this part is to explore the regularity on ρ. We propose the following factor to measure the one-sided bias when change the probability

$$\forall \rho',\rho\in(0,1), \qquad \Theta_{\rho',\rho}:=\max\left\{\frac{\rho'}{\rho},\frac{1-\rho'}{1-\rho},\right\}, \tag{6.4}$$

This quantity is used to control the continuity with respect to the small change of density ρ. We remark that, the role of ρ and ρ' is not same here, since we should always view the second parameter as the targeted density and the first parameter as its perturbation. The following observation is obvious

$$|\rho' - \rho| < \varepsilon \min\{\rho, 1-\rho\} \Longrightarrow \Theta_{\rho',\rho} \in [1, 1+\varepsilon]. \tag{6.5}$$

Especially, we only require the targeted density ρ not to be degenerated to 0 or 1 to get a good estimate, but the perturbation ρ' can be degenerate. We also define the two-sided bias factor

$$\widetilde{\Theta}_{\rho',\rho} := \max\{\Theta_{\rho',\rho}, \Theta_{\rho,\rho'}\}, \tag{6.6}$$

where the role of ρ and ρ' is symmetric.

The following lemma is an example to apply the one-sided bias factor $\Theta_{\rho',\rho}$, and it will be useful throughout the section.

Lemma 6.2. *For $\Lambda \subseteq \mathbb{Z}^d$ and every local function $f \in \mathcal{F}_0(\Lambda)$, the following inequality holds for every $\rho', \rho \in (0,1)$*

$$\left|\langle f \rangle_{\rho'} - \langle f \rangle_{\rho}\right| \leqslant \left(\Theta_{\rho',\rho}^{|\Lambda|} - 1\right) \langle |f| \rangle_{\rho}\,. \tag{6.7}$$

Proof. We make the decomposition using the canonical ensemble. Denote by X the random variable $X := \sum_{x\in\Lambda} \eta_x$

$$\begin{aligned}
\left|\langle f \rangle_{\rho'} - \langle f \rangle_{\rho}\right| &= \left|\sum_{M=0}^{|\Lambda|} \left(\mathbb{P}_{\rho'}\left[X = M\right] - \mathbb{P}_{\rho}\left[X = M\right]\right) \langle f \rangle_{\Lambda,M}\right| \\
&\leqslant \left|\sum_{M=0}^{|\Lambda|} \left|\mathbb{P}_{\rho'}\left[X = M\right] - \mathbb{P}_{\rho}\left[X = M\right]\right| \langle |f| \rangle_{\Lambda,M}\right| \\
&= \left|\sum_{M=0}^{|\Lambda|} \left|\frac{\mathbb{P}_{\rho'}\left[X = M\right]}{\mathbb{P}_{\rho}\left[X = M\right]} - 1\right| \mathbb{P}_{\rho}\left[X = M\right] \langle |f| \rangle_{\Lambda,M}\right|.
\end{aligned}$$

Using the expression of Binomial distribution and the bound $M \leqslant |\Lambda|$, we obtain

$$\left|\frac{\mathbb{P}_{\rho'}\left[X = M\right]}{\mathbb{P}_{\rho}\left[X = M\right]} - 1\right| = \left|\frac{(\rho')^M (1-\rho')^{|\Lambda|-M}}{\rho^M (1-\rho)^{|\Lambda|-M}} - 1\right| \leqslant \Theta_{\rho',\rho}^{|\Lambda|} - 1.$$

This concludes the desired result. □

We now study the dependence of the density ρ in several quantities. Similar argument can be found in [36, Proposition 6.1].

Proposition 6.3 (Regularity on density)**.** *For every $L \in N_+$, every $\rho, \rho', \rho'' \in (0,1)$ and $\xi \in B_1$, we have the following estimates using the factors Θ and $\widetilde{\Theta}$ defined respectively in (6.4) and (6.6).*

(1) *Regularity of conductivity: we have*

$$\overline{\mathbf{c}}(\rho', \Lambda_L) \leqslant \Theta_{\rho',\rho}^{(L+2\mathbf{r})^d} \overline{\mathbf{c}}(\rho, \Lambda_L),. \tag{6.8}$$

and

$$|\overline{\mathbf{c}}(\rho, \Lambda_L) - \overline{\mathbf{c}}(\rho', \Lambda_L)| \leqslant \left(\widetilde{\Theta}_{\rho',\rho}^{(L+2\mathbf{r})^d} - 1\right) \max\left\{|\overline{\mathbf{c}}|(\rho, \Lambda_L), |\overline{\mathbf{c}}|(\rho', \Lambda_L)\right\}. \tag{6.9}$$

(2) *Regularity of mean:*

$$\frac{1}{|\Lambda_L|} \left\langle \phi_{\rho,\Lambda_L,\xi} \right\rangle_{\rho'}^2 \leqslant L^2 \left(\Theta_{\rho',\rho}^{(L+2\mathbf{r})^d} - 1\right)^2 |\overline{\mathbf{c}}|(\rho, \Lambda_L) \tag{6.10}$$

(3) *Regularity of Dirichlet energy and L^2: if $\widetilde{\Theta}_{\rho,\rho''}^{(L+2\mathbf{r})^d}, \widetilde{\Theta}_{\rho',\rho''}^{(L+2\mathbf{r})^d} \leqslant 2$, we have*

$$L^{-2}\left\langle \frac{1}{|\Lambda_L|}(\phi_{\rho',\Lambda_L,\xi}-\phi_{\rho,\Lambda_L,\xi})^2\right\rangle_{\rho''} + \frac{1}{|\Lambda_L|}\left\langle \sum_{b\in\overline{\Lambda_L^*}} c_b(\pi_b\phi_{\rho',\Lambda_L,\xi}-\pi_b\phi_{\rho,\Lambda_L,\xi})^2\right\rangle_{\rho''} \tag{6.11}$$
$$\leqslant 10\left(\max\left\{\widetilde{\Theta}_{\rho,\rho''}^{(L+2\mathbf{r})^d}, \widetilde{\Theta}_{\rho',\rho''}^{(L+2\mathbf{r})^d}\right\}-1\right)\max\left\{|\overline{\mathbf{c}}|(\rho,\Lambda_L), |\overline{\mathbf{c}}|(\rho',\Lambda_L), |\overline{\mathbf{c}}|(\rho'',\Lambda_L)\right\}.$$

Proof. (1) We test $\phi_{\rho',\Lambda_L,\xi}$ in the variational problem of $\overline{\mathbf{c}}(\rho,\Lambda)$, and apply (6.7) to change the parameter of the grand canonical ensemble

$$\begin{aligned}\frac{1}{2}\xi\cdot\overline{\mathbf{c}}(\rho,\Lambda_L)\xi &\leqslant \frac{1}{|\Lambda_L|}\left\langle \sum_{b\in\overline{\Lambda_L^*}}\frac{1}{2}c_b\left(\pi_b(\ell_\xi+\phi_{\rho',\Lambda_L,\xi})\right)^2\right\rangle_\rho \\ &\leqslant \Theta_{\rho',\rho}^{(L+2\mathbf{r})^d}\frac{1}{|\Lambda_L|}\left\langle \sum_{b\in\overline{\Lambda_L^*}}\frac{1}{2}c_b\left(\pi_b(\ell_\xi+\phi_{\rho',\Lambda_L,\xi})\right)^2\right\rangle_{\rho'} \\ &= \Theta_{\rho',\rho}^{(L+2\mathbf{r})^d}\frac{1}{2}p\cdot\overline{\mathbf{c}}(\rho',\Lambda_L)p.\end{aligned} \tag{6.12}$$

Here because of the correlation length $\mathbf{r}$, the integration $\sum_{b\in\overline{\Lambda_L^*}}\frac{1}{2}c_b\pi_b(\ell_\xi+\phi_{\rho',\Lambda_L,\xi})^2$ is in $\mathcal{F}_0(\Lambda_{L+2\mathbf{r}})$ and we need to enlarge the power for the factor $\Theta_{\rho',\rho}$. This proves (6.8). By exchanging the role of ρ and ρ', we can obtain the estimate on the other direction similarly, which concludes (6.9) using $\widetilde{\Theta}_{\rho',\rho}$ defined in (6.6).

(2) Recall that $\left\langle\phi_{\rho,\Lambda_L,\xi}\right\rangle_\rho = 0$ from (1) of Proposition 6.1, then we apply (6.7) to $\frac{1}{|\Lambda_L|}\left\langle\phi_{\rho,\Lambda_L,\xi}\right\rangle_{\rho'}^2$

$$\begin{aligned}\frac{1}{|\Lambda_L|}\left\langle\phi_{\rho,\Lambda_L,\xi}\right\rangle_{\rho'}^2 &= \frac{1}{|\Lambda_L|}\left|\left\langle\phi_{\rho,\Lambda_L,\xi}\right\rangle_{\rho'}-\left\langle\phi_{\rho,\Lambda_L,\xi}\right\rangle_{\rho}\right|^2 \\ &\leqslant \frac{1}{|\Lambda_L|}\left(\Theta_{\rho',\rho}^{(L+2\mathbf{r})^d}-1\right)^2\left\langle|\phi_{\rho,\Lambda_L,\xi}|\right\rangle_\rho^2 \\ &\leqslant \frac{1}{|\Lambda_L|}\left(\Theta_{\rho',\rho}^{(L+2\mathbf{r})^d}-1\right)^2\left\langle\phi_{\rho,\Lambda_L,\xi}^2\right\rangle_\rho \\ &\leqslant L^2\left(\Theta_{\rho',\rho}^{(L+2\mathbf{r})^d}-1\right)^2|\overline{\mathbf{c}}|(\rho,\Lambda_L).\end{aligned}$$

Here we apply Jensen's inequality from the second line to the third line, and the spectral inequality (2.5) from the third line to the forth line.

(3) To compare the Dirichlet energy, we add $\phi_{\rho'',\Lambda_L,\xi}$ as an intermediate term, which gives

$$\begin{aligned}&\frac{1}{|\Lambda_L|}\left\langle\sum_{b\in\overline{\Lambda_L^*}}c_b(\pi_b\phi_{\rho',\Lambda_L,\xi}-\pi_b\phi_{\rho,\Lambda_L,\xi})^2\right\rangle_{\rho''} \\ &\leqslant \frac{2}{|\Lambda_L|}\left\langle\sum_{b\in\overline{\Lambda_L^*}}c_b(\pi_b\phi_{\rho',\Lambda_L,\xi}-\pi_b\phi_{\rho'',\Lambda_L,\xi})^2\right\rangle_{\rho''}+\frac{2}{|\Lambda_L|}\left\langle\sum_{b\in\overline{\Lambda_L^*}}c_b(\pi_b\phi_{\rho,\Lambda_L,\xi}-\pi_b\phi_{\rho'',\Lambda_L,\xi})^2\right\rangle_{\rho''}.\end{aligned}$$

For each term, we can repeat the argument in (6.12), and conclude that

$$\frac{1}{|\Lambda_L|}\left\langle\sum_{b\in\overline{\Lambda_L^*}}c_b(\pi_b\phi_{\rho',\Lambda_L,\xi}-\pi_b\phi_{\rho,\Lambda_L,\xi})^2\right\rangle_{\rho''} \tag{6.13}$$

$$\leqslant \left(\widetilde{\Theta}^{(L+2\mathbf{r})^d}_{\rho,\rho''} + \widetilde{\Theta}^{(L+2\mathbf{r})^d}_{\rho',\rho''} - 2\right) \max\left\{|\overline{\mathbf{c}}|(\rho,\Lambda_L), |\overline{\mathbf{c}}|(\rho',\Lambda_L), |\overline{\mathbf{c}}|(\rho'',\Lambda_L)\right\}.$$

The L^2 term can be done similarly,

$$\left\langle \frac{1}{|\Lambda_L|}(\phi_{\rho',\Lambda_L,\xi} - \phi_{\rho,\Lambda_L,\xi})^2 \right\rangle_{\rho''} \leqslant \left\langle \frac{2}{|\Lambda_L|}(\phi_{\rho',\Lambda_L,\xi} - \phi_{\rho'',\Lambda_L,\xi})^2 \right\rangle_{\rho''} + \left\langle \frac{2}{|\Lambda_L|}(\phi_{\rho,\Lambda_L,\xi} - \phi_{\rho'',\Lambda_L,\xi})^2 \right\rangle_{\rho''}.$$

Here, we hope to use Poincaré inequality, but the constant part of the function should be truncated. We take the term involving ρ, ρ'' for example

$$\left\langle \frac{2}{|\Lambda_L|}(\phi_{\rho,\Lambda_L,\xi} - \phi_{\rho'',\Lambda_L,\xi})^2 \right\rangle_{\rho''} \leqslant \left\langle \frac{4}{|\Lambda_L|}\left(\phi_{\rho,\Lambda_L,\xi} - \left\langle \phi_{\rho,\Lambda_L,\xi}\right\rangle_{\rho''} - \phi_{\rho'',\Lambda_L,\xi}\right)^2 \right\rangle_{\rho''} + \frac{4}{|\Lambda_L|}\left\langle \phi_{\rho,\Lambda_L,\xi}\right\rangle_{\rho''}^2.$$

The Poincaré inequality (2.5) applies for the first term, and then we can use (6.13)

$$\begin{aligned}
&\left\langle \frac{4}{|\Lambda_L|}\left(\phi_{\rho,\Lambda_L,\xi} - \left\langle \phi_{\rho,\Lambda_L,\xi}\right\rangle_{\rho''} - \phi_{\rho'',\Lambda_L,\xi}\right)^2 \right\rangle_{\rho''} \\
&\leqslant \frac{4L^2}{|\Lambda_L|}\left\langle \sum_{b\in\overline{\Lambda_L^*}} c_b(\pi_b\phi_{\rho,\Lambda_L,\xi} - \pi_b\phi_{\rho'',\Lambda_L,\xi})^2 \right\rangle_{\rho''} \\
&\leqslant 4L^2\left(\widetilde{\Theta}^{(L+2\mathbf{r})^d}_{\rho,\rho''} - 1\right)\max\left\{|\overline{\mathbf{c}}|(\rho,\Lambda_L), |\overline{\mathbf{c}}|(\rho'',\Lambda_L)\right\}.
\end{aligned}$$

Concerning the term $\frac{4}{|\Lambda_L|}\left\langle \phi_{\rho,\Lambda_L,\xi}\right\rangle_{\rho''}^2$, we apply directly (6.10). Under the assumption $\widetilde{\Theta}^{(L+2\mathbf{r})^d}_{\rho,\rho''} \leqslant 2$, then we have $\left(\Theta^{(L+2\mathbf{r})^d}_{\rho,\rho''} - 1\right)^2 \leqslant \left(\Theta^{(L+2\mathbf{r})^d}_{\rho,\rho''} - 1\right)$ and the leading order should be $\left(\widetilde{\Theta}^{(L+2\mathbf{r})^d}_{\rho,\rho''} - 1\right)$. This concludes the proof. □

6.2. **Convergence rate under canonical ensemble.** In this part, we give the convergence rate of diffusion matrix under the canonical ensemble. Inspired by the subadditive quantities in (4.1), we can also define their counterparts under the canonical ensemble that

$$\begin{aligned}
&\frac{1}{2}p\cdot\widehat{D}(\Lambda,N)p := \inf_{v\in\ell_{p,\Lambda^+}+\mathcal{F}_0(\Lambda^-)}\left\{\frac{1}{2\chi(N/|\Lambda|)|\Lambda|}\sum_{b\in\overline{\Lambda^*}}\left\langle\frac{1}{2}c_b(\pi_b v)^2\right\rangle_{\Lambda,N}\right\}, \\
&\frac{1}{2}q\cdot\widehat{D}_*^{-1}(\Lambda,N)q \\
&\qquad := \sup_{v\in\mathcal{F}_0}\left\{\frac{1}{2\chi(N/|\Lambda|)|\Lambda|}\sum_{b\in\overline{\Lambda^*}}\left\langle(\pi_b\ell_{q,\Lambda})(\pi_b v) - \frac{1}{2}c_b(\pi_b v)^2\right\rangle_{\Lambda,N}\right\}.
\end{aligned} \tag{6.14}$$

Similarly to (5.1), we define the conductivity using the Einstein relation (1.10)

$$\hat{\mathbf{c}}(\Lambda,N) := 2\chi(N/|\Lambda|)\widehat{D}(\Lambda,N), \qquad \hat{\mathbf{c}}_*(\Lambda,N) := 2\chi(N/|\Lambda|)\widehat{D}_*(\Lambda,N). \tag{6.15}$$

Here $\hat{\mathbf{c}}(\Lambda,N)$ also coincides with the definition (1.23). Our main result in this subsection is the following proposition.

Proposition 6.4. *Under Hypothesis 1.1, there exists an exponent $\gamma_2(d,\lambda,\mathbf{r})>0$ and a positive constant $C(d,\lambda,\mathbf{r})<\infty$ such that for every $L,M\in\mathbb{N}_+$,*

$$|\hat{\mathbf{c}}(\Lambda_L,M)-\overline{\mathbf{c}}(M/|\Lambda_L|)|+|\hat{\mathbf{c}}_*(\Lambda_L,M)-\overline{\mathbf{c}}(M/|\Lambda_L|)|\leqslant CL^{-\gamma_2}, \tag{6.16}$$

where $\overline{\mathbf{c}}(\rho)$ is the same as that defined in (5.2).

We establish at first a result of the local equivalence of ensembles. Similar result can be found in [51, Appendix 2] and other references. In our setting, for any $\Lambda\subseteq\mathbb{Z}^d$ and $\varepsilon\in(0,1)$, we define the following set of integers such that the empirical density is not degenerate

$$\mathcal{M}_\varepsilon(\Lambda):=\{M\in\mathbb{N}_+:\varepsilon\leqslant M/|\Lambda|\leqslant 1-\varepsilon\}. \tag{6.17}$$

Lemma 6.5. *Let $L,\ell\in\mathbb{N}_+$ and $M\in\mathbb{N}$, we denote by $\hat{\rho}$ the empirical density $\hat{\rho}:=\frac{M}{|\Lambda_L|}$.*

(1) *If $10\ell^2\leqslant L$ and $0\leqslant M\leqslant L^d$, then for any local function $f\in\mathcal{F}_0(\Lambda_\ell)$ such that $\langle f\rangle_{\Lambda_\ell,N}\geqslant 0$ for any $N\in\mathbb{N}$, we have*

$$\langle f\rangle_{\Lambda_L,M}\leqslant\left(1+4\left(\frac{\ell^2}{L}\right)^d\right)\langle f\rangle_{\hat{\rho}}. \tag{6.18}$$

(2) *Given $\varepsilon\in(0,1)$, if $10\ell^{2d}\leqslant\varepsilon L^d$ and $M\in\mathcal{M}_\varepsilon(\Lambda_L)$, then for any local function $f\in\mathcal{F}_0(\Lambda_\ell)$, we have*

$$\left|\langle f\rangle_{\Lambda_L,M}-\langle f\rangle_{\hat{\rho}}\right|\leqslant\frac{1}{\varepsilon}\left(\frac{\ell^2}{L}\right)^d\langle|f|\rangle_{\hat{\rho}}. \tag{6.19}$$

Proof. The proof is similar to Lemma 6.2. For $f\in\mathcal{F}_0(\Lambda_\ell)$, we decompose the expectation as

$$\langle f\rangle_{\Lambda_L,M}-\langle f\rangle_{\hat{\rho}}=\sum_{M=0}^{|\Lambda_L|}\left(\frac{\mathbb{P}_{\Lambda_L,M}\left[\sum_{x\in\Lambda_\ell}\eta_x=N\right]}{\mathbb{P}_{\hat{\rho}}\left[\sum_{x\in\Lambda_\ell}\eta_x=N\right]}-1\right)\mathbb{P}_{\hat{\rho}}\left[\sum_{x\in\Lambda_\ell}\eta_x=N\right]\langle f\rangle_{\Lambda_\ell,N}. \tag{6.20}$$

It remains to analyze the Radon–Nikodym derivative under different setting.

(1) Under this setting, because $\langle f\rangle_{\Lambda_\ell,N}$ is positive, it suffice to have a upper bound for the following probability

$$\mathbb{P}_{\Lambda_L,M}\left[\sum_{x\in\Lambda_\ell}\eta_x=N\right]=\frac{\binom{|\Lambda_L\setminus\Lambda_\ell|}{M-N}}{\binom{|\Lambda_L|}{M}}=\frac{\left(\frac{M!}{(M-N)!}\right)\left(\frac{(|\Lambda_L|-M)!}{(|\Lambda_L\setminus\Lambda_\ell|-(M-N))!}\right)}{\frac{|\Lambda_L|!}{|\Lambda_L\setminus\Lambda_\ell|!}}. \tag{6.21}$$

Notice that

$$\begin{aligned}&\frac{|\Lambda_L|!}{|\Lambda_L\setminus\Lambda_\ell|!}\geqslant|\Lambda_L\setminus\Lambda_\ell|^{|\Lambda_\ell|},\qquad \frac{M!}{(M-N)!}\leqslant M^N,\\ &\frac{(|\Lambda_L|-M)!}{(|\Lambda_L\setminus\Lambda_\ell|-(M-N))!}\leqslant(|\Lambda_L|-M)^{|\Lambda_\ell|-N},\end{aligned} \tag{6.22}$$

thus we have an upper bound for (6.21)

$$\begin{aligned}\frac{\mathbb{P}_{\Lambda_L,M}\left[\sum_{x\in\Lambda_\ell}\eta_x=N\right]}{\mathbb{P}_{\hat{\rho}}\left[\sum_{x\in\Lambda_\ell}\eta_x=N\right]}&\leqslant\left(\frac{M^N(|\Lambda_L|-M)^{|\Lambda_\ell|-N}}{|\Lambda_L\setminus\Lambda_\ell|^{|\Lambda_\ell|}}\right)\Big/\left(\frac{M^N(|\Lambda_L|-M)^{|\Lambda_\ell|-N}}{|\Lambda_L|^{|\Lambda_\ell|}}\right)\\ &=\left(\frac{|\Lambda_L|}{|\Lambda_L\setminus\Lambda_\ell|}\right)^{|\Lambda_\ell|}.\end{aligned} \tag{6.23}$$

It suffice to give an estimate for the term $\left(\frac{|\Lambda_L|}{|\Lambda_L \smallsetminus \Lambda_\ell|}\right)^{|\Lambda_\ell|}$

$$\left(\frac{|\Lambda_L|}{|\Lambda_L \smallsetminus \Lambda_\ell|}\right)^{|\Lambda_\ell|} \leqslant \left(1 + 2\left(\frac{\ell}{L}\right)^d\right)^{\ell^d} \leqslant e^{\log\left(1+2\left(\frac{\ell}{L}\right)^d\right)\ell^d} \leqslant e^{2\left(\frac{\ell^2}{L}\right)^d} \leqslant 1 + 4\left(\frac{\ell^2}{L}\right)^d.$$

This estimate together with (6.21), (6.23) and (6.18) gives us the desired result for positive function f.

(2) Under this setting, $\langle f \rangle_{\Lambda_\ell, N}$ can be negative, so we need two-sided estimate for the Radon-Nikodym derivative. The upper bound is already proved in (6.23), and we only need the lower bound, which is as follows

$$\begin{aligned} &\frac{|\Lambda_L|!}{|\Lambda_L \smallsetminus \Lambda_\ell|!} \leqslant |\Lambda_L|^{|\Lambda_\ell|}, \qquad \frac{M!}{(M-N)!} \geqslant (M-N)^N, \\ &\frac{(|\Lambda_L| - M)!}{(|\Lambda_L \smallsetminus \Lambda_\ell| - (M-N))!} \geqslant ((|\Lambda_L| - M) - (|\Lambda_\ell| - N))^{|\Lambda_\ell| - N}. \end{aligned} \tag{6.24}$$

This gives us a lower bound of the proportion

$$\frac{\mathbb{P}_{\Lambda_L, M}\left[\sum_{x \in \Lambda_\ell} \eta_x = N\right]}{\mathbb{P}_{\hat{\rho}}\left[\sum_{x \in \Lambda_\ell} \eta_x = N\right]} \geqslant \left(1 - \frac{N}{M}\right)^N \left(1 - \frac{|\Lambda_\ell| - N}{|\Lambda_L| - M}\right)^{|\Lambda_\ell| - N}.$$

Because $M \in \mathcal{M}_\varepsilon(\Lambda_L)$, we have

$$0 \leqslant \max\left\{\frac{N}{M}, \frac{|\Lambda_\ell| - N}{|\Lambda_L| - M}\right\} \leqslant \frac{\ell^d}{\varepsilon L^d},$$

which results in

$$\frac{\mathbb{P}_{\Lambda_L, M}\left[\sum_{x \in \Lambda_\ell} \eta_x = N\right]}{\mathbb{P}_{\hat{\rho}}\left[\sum_{x \in \Lambda_\ell} \eta_x = N\right]} \geqslant \left(1 - \frac{\ell^d}{\varepsilon L^d}\right)^{\ell^d} \geqslant 1 - \frac{1}{\varepsilon}\left(\frac{\ell^2}{L}\right)^d. \tag{6.25}$$

Here we also make use of condition $10\ell^{2d} \leqslant \varepsilon L^d$. This estimate and (6.20) together conclude (6.19). □

With this local equivalence of ensembles result, we can obtain the convergence rate of diffusion matrix under the canonical ensemble.

Proof of Proposition 6.4. By similar analysis as Proposition 4.1 and Lemma 4.6 in previous sections, we obtain that

$$\mathsf{Id} \leqslant \widehat{D}_*(\Lambda, N) \leqslant \widehat{D}(\Lambda, N) \leqslant \lambda \mathsf{Id}. \tag{6.26}$$

The strategy is the following *sandwich argument*: we prove that in scale $1 \ll \ell \ll L$, we have

$$\begin{aligned} \overline{D}_*(M/|\Lambda_L|, \Lambda_\ell) - C\ell L^{-1}\mathsf{Id} &\leqslant \widehat{D}_*(\Lambda_L, M) \\ &\leqslant \widehat{D}(\Lambda_L, M) \leqslant \overline{D}(M/|\Lambda_L|, \Lambda_\ell) + C\ell L^{-1}\mathsf{Id}. \end{aligned} \tag{6.27}$$

Then the distance from $\overline{D}(M/|\Lambda_L|, \Lambda_\ell)$ (resp. $\overline{D}_*(M/|\Lambda_L|, \Lambda_\ell)$) to $\overline{D}(M/|\Lambda_L|)$ bounds that from $\widehat{D}(\Lambda_L, M)$ (resp. $\widehat{D}_*(\Lambda_L, M)$) to $\overline{D}(M/|\Lambda_L|)$.

In the following, for the convenience to implement the renormalization step, we justify the two sides in (6.27) with $L = 3^m, \ell = 3^n, m, n \in \mathbb{N}_+$, but one can easily adapt it to the general case. We always suppose that the parameters $L = 3^m, \ell = 3^n$ satisfy the conditions of Lemma 6.5. We also denote by $\hat{\rho} := \frac{M}{|\square_m|}$ to lighten the notation, which can be interpreted as the empirical density of particles under the canonical ensemble.

Step 1: comparison between $\widehat{D}(\square_m, M)$ *and* $\overline{D}(\hat{\rho}, \square_n)$. We propose a sub-minimizer $\ell_\xi + \widetilde{\phi}_{\hat{\rho},\square_m,\xi}$ for the optimization problem of $\widehat{D}(\square_m, M)$ that

$$\widetilde{\phi}_{\hat{\rho},\square_m,\xi} := \sum_{\substack{z\in\mathcal{Z}_{m,n},\\ \mathrm{dist}(z,\partial\square_m)>3^n}} \phi_{\hat{\rho},z+\square_n,\xi}.$$

Here we require the $\mathrm{dist}(z, \partial\square_m) > 3^n$ in order to cutoff the influence of particles from the domain outside $\square_m$. Since it is a sub-minimizer, we have

$$\begin{aligned}
&\frac{1}{2}\xi\cdot\widehat{D}(\square_m, M)\xi \\
&\leqslant \frac{1}{2\chi(\hat{\rho})|\square_m|}\left\langle \frac{1}{2}\sum_{b\in\square_m^*} c_b(\pi_b(\ell_\xi + \widetilde{\phi}_{\hat{\rho},\square_m,\xi}))^2 \right\rangle_{\square_m,M} \\
&\leqslant \frac{1}{|\mathcal{Z}_{m,n}|}\sum_{\substack{z\in\mathcal{Z}_{m,n},\\ \mathrm{dist}(z,\partial\square_m)>3^n}} \frac{1}{2\chi(\hat{\rho})|\square_n|}\left\langle \frac{1}{2}\sum_{b\in(z+\square_n)^*} c_b(\pi_b(\ell_\xi + \widetilde{\phi}_{\hat{\rho},z+\square_n,\xi}))^2 \right\rangle_{\square_m,M} \\
&\quad + \lambda 3^{-(m-n)}|p|^2.
\end{aligned} \tag{6.28}$$

Here the last error term $\lambda 3^{-(m-n)}|p|^2$ comes from the calculation of Dirichlet energy in the boundary layer of width 3^n, where we do not pose any corrector and the affine function can be calculated directly. For the Dirichlet energy in the interior and when $3^n > \mathbf{r}$, then for every $z \in \mathcal{Z}_{m,n}, \mathrm{dist}(z, \partial\square_m) > 3^n$, we have that

$$\sum_{b\in(z+\square_n)^*} c_b(\pi_b(\ell_\xi + \widetilde{\phi}_{\hat{\rho},z+\square_n,\xi}))^2 \in \mathcal{F}_0(z + \square_{n+1}),$$

and $z+\square_{n+1} \subseteq \square_m$. Then Lemma 6.5-(1) applies to this case ($z \in \mathcal{Z}_{m,n}, \mathrm{dist}(z, \partial\square_m) > 3^n > \mathbf{r}$) and we have

$$\begin{aligned}
&\frac{1}{2\chi(\hat{\rho})|\square_n|}\left\langle \frac{1}{2}\sum_{b\in(z+\square_n)^*} c_b(\pi_b(\ell_\xi + \widetilde{\phi}_{\hat{\rho},z+\square_n,\xi}))^2 \right\rangle_{\square_m,M} \\
&\leqslant \frac{(1+3^{d(2n-m)})}{2\chi(\hat{\rho})|\square_n|}\left\langle \frac{1}{2}\sum_{b\in(z+\square_n)^*} c_b(\pi_b(\ell_\xi + \widetilde{\phi}_{\hat{\rho},z+\square_n,\xi}))^2 \right\rangle_{\hat{\rho}} \\
&\leqslant (1+3^{d(2n-m)})\left(\frac{1}{2}\xi\cdot\overline{D}(\square_n, \hat{\rho})\xi\right).
\end{aligned}$$

We put this back to (6.28) and concludes that

$$\widehat{D}(\square_m, M) \leqslant \overline{D}(\square_n, \hat{\rho}) + C(\lambda)(3^{-(m-n)} + 3^{-d(m-2n)})\mathsf{Id}. \tag{6.29}$$

Step 2: comparison between $\widehat{D}_*(\square_m, M)$ *and* $\overline{D}_*(\hat{\rho}, \square_n)$. This part is quite close to that in Step 1. We make the decomposition that

$$\begin{aligned}
&\frac{1}{2\chi(\hat{\rho})|\square_m|}\sum_{b\in\square_m^*}\left\langle (\pi_b\ell_q)(\pi_b v) - \frac{1}{2}c_b(\pi_b v)^2 \right\rangle_{\square_m,M} \\
&= \frac{1}{|\mathcal{Z}_{m,n}|}\sum_{\substack{z\in\mathcal{Z}_{m,n},\\ \mathrm{dist}(z,\partial\square_m)>3^n}} \frac{1}{2\chi(\hat{\rho})|\square_n|}\sum_{b\in(z+\square_n)^*}\left\langle (\pi_b\ell_q)(\pi_b v) - \frac{1}{2}c_b(\pi_b v)^2 \right\rangle_{\square_m,M} \\
&\quad + \frac{1}{2\chi(\hat{\rho})|\square_m|}\sum_{\substack{b\in\square_m^*,\\ \text{other bonds}}}\left\langle (\pi_b\ell_q)(\pi_b v) - \frac{1}{2}c_b(\pi_b v)^2 \right\rangle_{\square_m,M}
\end{aligned} \tag{6.30}$$

Let $u_{z+\square_n,\hat{\rho},q}$ be the maximizer for the problem $\overline{\nu}_*(z+\square_n,\hat{\rho},q)$. Then for the case $z \in \mathcal{Z}_{m,n}, \operatorname{dist}(z,\partial\square_m) > 3^n > \mathbf{r}$ and by the definition of $\overline{\nu}_*(z+\square_n,\hat{\rho},q)$, we also have

$$
\begin{aligned}
&\frac{1}{2\chi(\hat{\rho})|\square_n|}\sum_{b\in(z+\square_n)^*}\left\langle(\pi_b\ell_q)(\pi_b v) - \frac{1}{2}c_b(\pi_b v)^2\right\rangle_{\square_m,M} \\
&\leqslant \frac{1}{2\chi(\hat{\rho})|\square_n|}\sum_{b\in(z+\square_n)^*}\left\langle(\pi_b\ell_q)(\pi_b u_{z+\square_n,\hat{\rho},q}) - \frac{1}{2}c_b(\pi_b u_{z+\square_n,\hat{\rho},q})^2\right\rangle_{\square_m,M} \\
&\leqslant (1+3^{d(2n-m)})\left(\frac{1}{2}q\cdot\overline{D}_*^{-1}(\square_n,\hat{\rho})q\right).
\end{aligned}
$$

Here from the first line to the second line, we use the fact that $u_{z+\square_n,\hat{\rho},q}$ is also the maixmiser under the canonical ensemble and gives positive functional; see (4.11) and Remark 4.2 for details. Then from the second line to the third line, note the positivity of the integral, Lemma 6.5-(1) applies. We only needs to treat the boundary layer and the layer between the small cubes appearing in the third line of (6.30), which can be solved by the following uniform point-wise estimate that

$$
(\pi_b\ell_q)(\pi_b v) - \frac{1}{2}c_b(\pi_b v)^2 \leqslant \frac{1}{2}(\pi_b\ell_q)^2 + \frac{1}{2}(\pi_b v)^2 - \frac{1}{2}(\pi_b v)^2 = \frac{1}{2}(\pi_b\ell_q)^2.
$$

Therefore, (6.30) can be reduced to

$$
\widehat{D}_*^{-1}(\square_m,M) \leqslant \overline{D}_*^{-1}(\square_n,\hat{\rho}) + C(\lambda)(3^{-(m-n)} + 3^{-d(m-2n)})\mathsf{Id},
$$

which implies, noting $\overline{D}_*(\square_n,\hat{\rho}) \leqslant \lambda\mathsf{Id}$, that

$$
\widehat{D}_*(\square_m,M) \geqslant \overline{D}_*(\square_n,\hat{\rho}) - C(\lambda)\lambda^2(3^{-(m-n)} + 3^{-d(m-2n)})\mathsf{Id}. \tag{6.31}
$$

Step 3: conclusion. Combining (6.26), (6.29) and (6.31), we obtain

$$
\begin{aligned}
\overline{D}_*(\square_n,\hat{\rho}) - C(\lambda)(3^{-(m-n)} + 3^{-d(m-2n)})\mathsf{Id} &\leqslant \widehat{D}_*(\square_m,M) \\
&\leqslant \widehat{D}(\square_m,M) \leqslant \overline{D}(\square_n,\hat{\rho}) + C(\lambda)(3^{-(m-n)} + 3^{-d(m-2n)})\mathsf{Id}.
\end{aligned}
$$

Applying (6.15) and (5.2), this is interpreted as

$$
\begin{aligned}
\overline{\mathbf{c}}_*(\square_n,\hat{\rho}) - C(\lambda)\chi(\hat{\rho})(3^{-(m-n)} + 3^{-d(m-2n)})\mathsf{Id} &\leqslant \hat{\mathbf{c}}_*(\square_m,M) \\
&\leqslant \hat{\mathbf{c}}(\square_m,M) \leqslant \overline{\mathbf{c}}(\square_n,\hat{\rho}) + C(\lambda)\chi(\hat{\rho})(3^{-(m-n)} + 3^{-d(m-2n)})\mathsf{Id}.
\end{aligned}
$$

Inserting the result of homogenization (5.3), and by a choice of n that

$$
d(m-2n) = \gamma_1 n \iff n = \frac{dm}{2d+\gamma_1},
$$

we obtain that

$$
|\hat{\mathbf{c}}(\square_m,M) - \overline{\mathbf{c}}(\hat{\rho})| + |\hat{\mathbf{c}}_*(\square_m,M) - \overline{\mathbf{c}}(\hat{\rho})| \leqslant C3^{-\left(\frac{d\gamma_1}{2d+\gamma_1}\right)m}.
$$

This is the desired result (6.16) by setting $\gamma_2 := \frac{d\gamma_1}{2d+\gamma_1}$. □

6.3. **Construction of density-free local corrector.** This subsection is devoted to remove the dependence of density in the local corrector. A first natural idea is to consider the following variational problem for general $\Lambda \subseteq \mathbb{Z}^d$ and $\xi \in \mathbb{R}^d$

$$
\mu(\Lambda,\xi) := \inf_{v\in\mathcal{F}_0(\Lambda^-)}\sup_{\rho\in[0,1]}\left\{\frac{1}{|\Lambda|}\sum_{b\in\Lambda^*}\left\langle\frac{1}{2}c_b(\pi_b(\ell_\xi + v))^2\right\rangle_\rho - \frac{1}{2}\xi\cdot\overline{\mathbf{c}}(\rho)\xi\right\}. \tag{6.32}
$$

As expected, the uniform convergence can be improved as follows.

Proposition 6.6. *There exists an exponent $\gamma_3(d,\lambda,\mathbf{r}) > 0$ and a constant $C(d,\lambda,\mathbf{r}) < \infty$ such that for every $L \in \mathbb{N}_+$ and $\xi \in B_1$, we have*

$$0 \leqslant \mu(\Lambda_L, \xi) \leqslant C L^{-\gamma_3}. \tag{6.33}$$

Compared with (4.1), the variational problem (6.32) is more complicated. Therefore, instead of attacking (6.32) directly, we turn to construct *one* sub-minimizer with a good uniform convergence for $\rho \in [0,1]$. This is the main task in the remaining part of this subsection. We are inspired by the function used to prove the qualitative version of (1.18) in the previous work [34, Lemma 2.1], where the main idea is to make the linear combination of the local corrector using the empirical density. Thus, we propose our density-free local corrector built on (6.1)

$$\hat{\phi}^{(\varepsilon)}_{m,n,\xi} := \sum_{\substack{z \in \mathcal{Z}_{m,n}, \\ \operatorname{dist}(z, \partial \square_m) > 3^n}} \phi_{\overline{\eta}_0 \vee \varepsilon \wedge (1-\varepsilon), z + \square_n, \xi}. \tag{6.34}$$

Recall $\mathcal{Z}_{m,n} = 3^n \mathbb{Z}^d \cap \square_m$. Here $\overline{\eta}_0$ is a (random) empirical density defined as $\overline{\eta}_0 := \frac{1}{|\square_m^-|} \sum_{x \in \square_m^-} \eta_x$. The mapping $\rho \mapsto \rho \vee \varepsilon \wedge (1-\varepsilon)$ restricts the density away from 0 and 1. That is, we make the truncation both for the spatial boundary layer and for the density, in order to avoid the perturbation from the rare but degenerate cases. More explicitly, recall the notation $\mathcal{M}_\varepsilon(\Lambda)$ defined in (6.17), and denote by M_*, M^*

$$M_* := \min \mathcal{M}_\varepsilon(\square_m^-), \qquad M^* := \max \mathcal{M}_\varepsilon(\square_m^-), \tag{6.35}$$

then we can rewrite this corrector as

$$\hat{\phi}^{(\varepsilon)}_{m,n,\xi} = \sum_{M=0}^{\infty} \left(\sum_{\substack{z \in \mathcal{Z}_{m,n}, \\ \operatorname{dist}(z, \partial \square_m) > 3^n}} \phi_{\frac{M \vee M_* \wedge M^*}{|\square_m^-|}, z + \square_n, \xi} \right) \mathbf{1}_{\left\{ \sum_{x \in \square_m^-} \eta_x = M \right\}}. \tag{6.36}$$

When n, ε is well-chosen with respect to m, this corrector will give us uniform convergence under all grand canonical ensembles.

Proposition 6.7. *There exist an exponent $\gamma_4(d,\lambda,\mathbf{r}) > 0$ and a constant $C(d,\lambda,\mathbf{r}) < \infty$ such that for every $m \in \mathbb{N}_+$, by setting $n := \lfloor \frac{m}{9d+3} \rfloor$ and $\varepsilon := 3^{-\frac{2dm}{9d+3}}$, we have*

$$\sup_{\rho \in [0,1], \xi \in B_1} \left| \frac{1}{|\square_m|} \left\langle \sum_{b \in \overline{\square_m^*}} \frac{1}{2} c_b \left(\pi_b (\ell_\xi + \hat{\phi}^{(\varepsilon)}_{m,n,\xi}) \right)^2 \right\rangle_\rho - \frac{1}{2} \xi \cdot \overline{\mathbf{c}}(\rho) \xi \right| \leqslant C 3^{-\gamma_4 m}. \tag{6.37}$$

Proof. We assume $\mathbf{r} \ll 3^n \ll 3^m$ and $0 < \varepsilon \ll 1$ with n, ε to be determined in the end. Since $\hat{\phi}^{(\varepsilon)}_{m,n,\xi} \in \mathcal{F}_0(\square_m^-)$, we compare it with the variational problem (4.1), which gives

$$\frac{1}{|\square_m|} \left\langle \sum_{b \in \overline{\square_m^*}} \frac{1}{2} c_b \left(\pi_b (\ell_\xi + \hat{\phi}^{(\varepsilon)}_{m,n,\xi}) \right)^2 \right\rangle_\rho \geqslant \frac{1}{2} \xi \cdot \overline{\mathbf{c}}(\rho, \square_m) \xi \geqslant \frac{1}{2} \xi \cdot \overline{\mathbf{c}}(\rho) \xi. \tag{6.38}$$

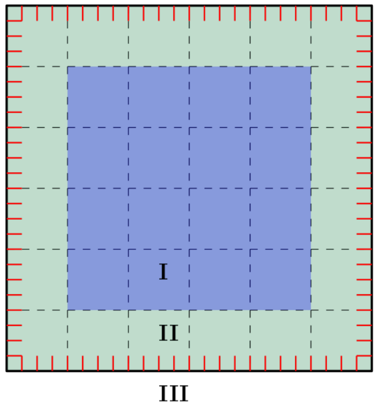

FIGURE 3. An illustration for the decomposition in (6.39): the bonds in the term **I**, **II** and **III** are respectively marked in blue, green and red.

Thus it suffices to give the upper bound, and we can decompose the sum as

$$\begin{aligned}
&\frac{1}{|\square_m|}\sum_{b\in\overline{\square_m^*}}\left\langle\frac{1}{2}c_b(\pi_b(\ell_\xi+\hat{\phi}^{(\varepsilon)}_{m,n,\xi}))^2\right\rangle_\rho = \mathbf{I}+\mathbf{II}+\mathbf{III},\\
&\mathbf{I} := \sum_{b\in\overline{\square_m^*}, \operatorname{dist}(b,\partial\square_m)>3^n}\frac{1}{|\square_m|}\left\langle\frac{1}{2}c_b(\pi_b(\ell_\xi+\hat{\phi}^{(\varepsilon)}_{m,n,\xi}))^2\right\rangle_\rho,\\
&\mathbf{II} := \sum_{\substack{b\in\overline{\square_m^*}\setminus(\square_m,\square_m^-)^*\\ \operatorname{dist}(b,\partial\square_m)\leqslant 3^n}}\frac{1}{|\square_m|}\left\langle\frac{1}{2}c_b(\pi_b(\ell_\xi+\hat{\phi}^{(\varepsilon)}_{m,n,\xi}))^2\right\rangle_\rho,\\
&\mathbf{III} := \sum_{b\in(\square_m,\square_m^-)^*}\frac{1}{|\square_m|}\left\langle\frac{1}{2}c_b(\pi_b(\ell_\xi+\hat{\phi}^{(\varepsilon)}_{m,n,\xi}))^2\right\rangle_\rho.
\end{aligned}\tag{6.39}$$

Roughly, **I** is the main contribution, while the terms **II** and **III** come from the boundary layer. Since the order of boundary layer is $3^{(d-1)m}$, they vanish after the normalization $\frac{1}{|\square_m|}$ when $m\nearrow\infty$. See Figure 3 for an illustration. We treat these three terms one by one in the remaining paragraphs.

Step 1: term **I**. This is the case that the bonds stay in the interior, and is the main contribution of the Dirichlet energy. This term can be further decomposed as

$$\mathbf{I}\leqslant\frac{1}{|\mathcal{Z}_{m,n}|}\sum_{\substack{z\in\mathcal{Z}_{m,n},\\ \operatorname{dist}(z,\partial\square_m)>3^n}}\frac{1}{|\square_n|}\sum_{b\in\overline{(z+\square_n)^*}}\left\langle\frac{1}{2}c_b(\pi_b(\ell_\xi+\hat{\phi}^{(\varepsilon)}_{m,n,\xi}))^2\right\rangle_\rho.$$

We focus on one term $(z+\square_n)$ above and $b\in\overline{(z+\square_n)^*}$, and have the following observation

$$\pi_b\hat{\phi}^{(\varepsilon)}_{m,n,\xi}(\eta)=\sum_{M=0}^{\infty}\mathbf{1}_{\left\{\sum_{x\in\square_m^-}\eta_x=M\right\}}\pi_b\phi_{\frac{M\vee M_*\wedge M^*}{|\square_m^-|},z+\square_n,\xi}(\eta),\tag{6.40}$$

because the Kawasaki operator π_b is conservative for the number of particles in $\square_m^-$ and only one local corrector is perturbed. Therefore, we apply the decomposition of canonical ensemble for the sum over $\overline{(z+\square_n)^*}$

$$\frac{1}{|\square_n|}\sum_{b\in\overline{(z+\square_n)^*}}\left\langle\frac{1}{2}c_b(\pi_b(\ell_\xi+\hat{\phi}^{(\varepsilon)}_{m,n,\xi}))^2\right\rangle_\rho$$

$$= \sum_{M=0}^{\infty} \mathbb{P}_\rho\left[\sum_{x\in\square_m^-} \eta_x = M\right] \frac{1}{|\square_n|} \sum_{b\in(z+\square_n)^*} \left\langle \frac{1}{2} c_b \left(\pi_b v \left(M \vee M_* \wedge M^*/|\square_m^-|, z+\square_n, \xi\right)\right)^2 \right\rangle_{M,\square_m^-}$$

$$\leqslant (1 + 4\cdot 3^{d(2n-m)}) \sum_{M=0}^{\infty} \mathbb{P}_\rho\left[\sum_{x\in\square_m^-} \eta_x = M\right]$$

$$\times \frac{1}{|\square_n|} \sum_{b\in(z+\square_n)^*} \left\langle \frac{1}{2} c_b \left(\pi_b v \left(M \vee M_* \wedge M^*/|\square_m^-|, z+\square_n, \xi\right)\right)^2 \right\rangle_{\frac{M}{|\square_m^-|}}.$$

We use the definition (6.1) that $\ell_\xi + \phi_{\frac{M}{|\square_m^-|}, z+\square_n, \xi} = v\left(M/|\square_m^-|, z+\square_n, \xi\right)$ from the first line to the second line. From the second line to the third line, we use the local equivalence of ensembles in (6.18) and need to assume

$$1 \leqslant n < \frac{m}{2}. \tag{6.41}$$

We then study the last line and distinguish it in 3 cases

Step 1.1: case $M \in \mathcal{M}_\varepsilon(\square_m^-)$. Then there is no bias between the probability space and the associated corrector, thus we have

(6.42)

$$\forall M \in \mathcal{M}_\varepsilon(\square_m^-), \quad \frac{1}{|\square_n|} \sum_{b\in(z+\square_n)^*} \left\langle \frac{1}{2} c_b \left(\pi_b v \left(M \vee M_* \wedge M^*/|\square_m^-|, z+\square_n, \xi\right)\right)^2 \right\rangle_{\frac{M}{|\square_m^-|}}$$
$$= \left(\frac{1}{2}\xi \cdot \bar{\mathbf{c}}(M/|\square_m^-|, z+\square_n)\xi\right).$$

Step 1.2: case $0 \leqslant M \leqslant M_*$. For this case, we have the following estimate

$$\frac{1}{|\square_n|} \sum_{b\in(z+\square_n)^*} \left\langle \frac{1}{2} c_b \left(\pi_b v \left(M \vee M_* \wedge M^*/|\square_m^-|, z+\square_n, \xi\right)\right)^2 \right\rangle_{\frac{M}{|\square_m^-|}}$$
$$= \frac{1}{|\square_n|} \sum_{b\in(z+\square_n)^*} \left\langle \frac{1}{2} c_b \left(\pi_b v \left(M_*/|\square_m^-|, z+\square_n, \xi\right)\right)^2 \right\rangle_{\frac{M}{|\square_m^-|}}$$
$$\leqslant \Theta^{(3^n+2\mathbf{r})^d}_{\frac{M}{|\square_m^-|}, \frac{M_*}{|\square_m^-|}} \frac{1}{|\square_n|} \sum_{b\in(z+\square_n)^*} \left\langle \frac{1}{2} c_b \left(\pi_b v \left(M_*/|\square_m^-|, z+\square_n, \xi\right)\right)^2 \right\rangle_{\frac{M_*}{|\square_m^-|}}.$$

Here from the second line to the third line, we use the regularity estimate (6.7), which gives an one-sided bias factor $\Theta^{(3^n+2\mathbf{r})^d}_{\frac{M}{|\square_m^-|}, \frac{M_*}{|\square_m^-|}}$. This factor will not explode because the targeted density $\frac{M_*}{|\square_m^-|}$ does not degenerate; recall the remark around (6.5). More precisely, using the definition of (6.4), $M \leqslant M_*$ and $\frac{M_*}{|\square_m^-|} \simeq \varepsilon \ll 1$, we have

$$\Theta_{\frac{M}{|\square_m^-|}, \frac{M_*}{|\square_m^-|}} = \max\left\{\frac{M}{M_*}, \frac{1 - \frac{M}{|\square_m^-|}}{1 - \frac{M_*}{|\square_m^-|}}\right\} = \frac{1 - \frac{M}{|\square_m^-|}}{1 - \frac{M_*}{|\square_m^-|}} \leqslant 1 + \frac{\frac{M_*}{|\square_m^-|}}{1 - \frac{M_*}{|\square_m^-|}} \leqslant 1 + 2\varepsilon,$$

which results in

$$\Theta^{(3^n+2\mathbf{r})^d}_{\frac{M}{|\square_m^-|}, \frac{M_*}{|\square_m^-|}} \leqslant (1+2\varepsilon)^{(3^n+2\mathbf{r})^d} \leqslant 1 + 10\cdot 3^{dn}\cdot\varepsilon.$$

Here we also need to assume

$$0 < \varepsilon \ll 3^{-dn}. \tag{6.43}$$

Using the stationarity, this concludes that

$$\forall 0 \leqslant M < M_*, \quad \frac{1}{|\square_n|} \sum_{b \in (z+\square_n)^*} \left\langle \frac{1}{2} c_b \left(\pi_b v \left(M \vee M_* \wedge M^* / |\square_m^-|, z + \square_n, \xi \right) \right)^2 \right\rangle_{\frac{M}{|\square_m^-|}} \leqslant \left(1 + 10 \cdot 3^{dn} \cdot \varepsilon \right) \left(\frac{1}{2} \xi \cdot \overline{\mathbf{c}}(M_*/|\square_m^-|, z + \square_n) \xi \right) \tag{6.44}$$

Step 1.3: case $M_* < M < |\square_m^-|$. This case is similar to Step 1.2, and we have

$$\forall M_* < M < |\square_m^-|, \quad \frac{1}{|\square_n|} \sum_{b \in (z+\square_n)^*} \left\langle \frac{1}{2} c_b \left(\pi_b v \left(M \vee M_* \wedge M^* / |\square_m^-|, z + \square_n, \xi \right) \right)^2 \right\rangle_{\frac{M}{|\square_m^-|}} \leqslant \left(1 + 10 \cdot 3^{dn} \cdot \varepsilon \right) \left(\frac{1}{2} \xi \cdot \overline{\mathbf{c}}(M^*/|\square_m^-|, z + \square_n) \xi \right) \tag{6.45}$$

Combing (6.42), (6.44), (6.45) and the spatial homogeneity from (2) of Hypothesis 1.1, we obtain an estimate of the term $\mathbf{I}$ under the assumption (6.41) and (6.43)

$$\mathbf{I} \leqslant \left(1 + 10 \cdot 3^{d(2n-m)} + 10 \cdot 3^{dn} \cdot \varepsilon \right) \sum_{M=0}^{\infty} \mathbb{P}_\rho \left[\sum_{x \in \square_m^-} \eta_x = M \right] \left(\frac{1}{2} \xi \cdot \overline{\mathbf{c}} \left(\frac{M \vee M_* \wedge M^*}{|\square_m^-|}, \square_n \right) \xi \right). \tag{6.46}$$

We aim to compare this quantity with the target $\frac{1}{2} \xi \cdot \overline{\mathbf{c}} \left(\rho, \square_n \right) \xi$, which relies on the concentration of measure and the regularity of $\overline{\mathbf{c}} \left(\rho, \square_n \right)$. We resume this as Step 1.4.

Step 1.4: regularity on density. Recall the Markov inequality of density

$$\mathbb{P}_\rho \left[\left| \frac{1}{|\square_m^-|} \sum_{x \in \square_m^-} \eta_x - \rho \right| > \delta \right] \leqslant \frac{\mathbb{V}\mathrm{ar}_\rho [\frac{1}{|\square_m^-|} \sum_{x \in \square_m^-} \eta_x]}{\delta^2} = \frac{\rho(1-\rho)}{|\square_m^-| \delta^2}. \tag{6.47}$$

We choose $\delta > 0$ in function of ρ and distinguish two cases. Viewing the assumption (6.43), here the threshold is tighter than ε.

Step 1.4.1: case $\rho \in [3^{-dn}, 1 - 3^{-dn}]$. We choose the window $\delta := 3^{-3dn}$ for such case, then treat the regime $\left| \frac{M}{|\square_m^-|} - \rho \right| \leqslant \delta$ at first in (6.46)

$$\begin{aligned}
&\sum_{M \in \mathbb{N}, \left| \frac{M}{|\square_m^-|} - \rho \right| \leqslant \delta} \mathbb{P}_\rho \left[\sum_{x \in \square_m^-} \eta_x = M \right] \left(\frac{1}{2} \xi \cdot \overline{\mathbf{c}} \left(\frac{M \vee M_* \wedge M^*}{|\square_m^-|}, \square_n \right) \xi \right) \\
&\leqslant \sum_{M \in \mathbb{N}, \left| \frac{M}{|\square_m^-|} - \rho \right| \leqslant \delta} \mathbb{P}_\rho \left[\sum_{x \in \square_m^-} \eta_x = M \right] \left(\frac{1}{2} \xi \cdot \overline{\mathbf{c}} \left(\rho, \square_n \right) \xi \right) \Theta^{(3^n + 2\mathbf{r})^d}_{\frac{M}{|\square_m^-|}, \rho} \\
&\leqslant (1 + 3^{dn} \delta)^{2 \cdot 3^{dn}} \left(\frac{1}{2} \xi \cdot \overline{\mathbf{c}} \left(\rho, \square_n \right) \xi \right) \\
&\leqslant (1 + 4 \cdot 3^{-dn}) \left(\frac{1}{2} \xi \cdot \overline{\mathbf{c}} \left(\rho, \square_n \right) \xi \right).
\end{aligned}$$

Here because of the assumption (6.43), we have $M \vee M_* \wedge M^* = M$ for all $\left| \frac{M}{|\square_m^-|} - \rho \right| \leqslant \delta$. Then we apply (6.8) from the first line to the second line. From the second line to the third line, we apply (6.5) to give an upper bound for $\Theta^{(3^n+2\mathbf{r})^d}_{\frac{M}{|\square^-_m|},\rho}$

$$\forall \rho \in [3^{-dn}, 1-3^{-dn}], \qquad \left|\frac{M}{|\square^-_m|} - \rho\right| \leqslant \delta \Longrightarrow \frac{\left|\frac{M}{|\square^-_m|} - \rho\right|}{\min\{\rho, 1-\rho\}} < 3^{dn}\delta$$
$$\overset{(6.5)}{\Longrightarrow} \Theta^{(3^n+2\mathbf{r})^d}_{\frac{M}{|\square^-_m|},\rho} \leqslant (1+3^{dn}\delta)^{2\cdot 3^{dn}}.$$

Finally, we insert the choice $\delta = 3^{-3dn}$ from the third line to the forth line.

For the regime $\left|\frac{M}{|\square^-_m|} - \rho\right| \geqslant \delta$, we just use the Markov inequality (6.47) and the trivial bound of $\overline{\mathbf{c}}$

$$\sum_{M\in\mathbb{N}, \left|\frac{M}{|\square^-_m|} - \rho\right|\geqslant\delta} \mathbb{P}_\rho\left[\sum_{x\in\square^-_m}\eta_x = M\right]\left(\frac{1}{2}\xi\cdot\overline{\mathbf{c}}\left(\frac{M\vee M_*\wedge M^*}{|\square^-_m|}, \square_n\right)\xi\right) \leqslant \lambda 3^{d(6n-m)}.$$

Combing the two estimates above and (6.46), we conclude the estimate for $\mathbf{I}$ under the condition $\rho \in [3^{-dn}, 1-3^{-dn}]$

$$\mathbf{I} \leqslant \left(\frac{1}{2}\xi\cdot\overline{\mathbf{c}}\,(\rho, \square_n)\,\xi\right) + C(\lambda)\left(3^{dn}\varepsilon + 3^{-dn} + 3^{d(6n-m)}\right). \tag{6.48}$$

Step 1.4.2: case $\rho \in [0, 3^{-dn}) \cup (1-3^{-dn}, 1]$. We choose the window $\delta := 3^{-\frac{dm}{2}}$, then for $\left|\frac{M}{|\square^-_m|} - \rho\right| \leqslant \delta$, we have $\frac{M}{|\square^-_m|} \in [0, 2\cdot 3^{-dn}] \cup [1-2\cdot 3^{-dn}, 1]$ since we assume (6.41). Then with the *a priori* bound $\left(\frac{1}{2}\xi\cdot\overline{\mathbf{c}}\left(\frac{M}{|\square^-_m|}, \square_n\right)\xi\right) \leqslant \chi(\frac{M}{|\square^-_m|})$, we have

$$\begin{aligned}
&\sum_{M\in\mathbb{N}, \left|\frac{M}{|\square^-_m|} - \rho\right|\leqslant\delta} \mathbb{P}_\rho\left[\sum_{x\in\square^-_m}\eta_x = M\right]\left(\frac{1}{2}\xi\cdot\overline{\mathbf{c}}\left(\frac{M\vee M_*\wedge M^*}{|\square^-_m|}, \square_n\right)\xi\right)\\
&\leqslant \sum_{M\in\mathbb{N}, \left|\frac{M}{|\square^-_m|} - \rho\right|\leqslant\delta} \mathbb{P}_\rho\left[\sum_{x\in\square^-_m}\eta_x = M\right]\chi\left(\frac{M}{|\square^-_m|}\right)\\
&\leqslant C(\lambda)3^{-dn}.
\end{aligned}$$

For $\left|\frac{M}{|\square^-_m|} - \rho\right| \geqslant \delta$, we use Markov inequality (6.47) to give

$$\sum_{M\in\mathbb{N}, \left|\frac{M}{|\square^-_m|} - \rho\right|\geqslant\delta} \mathbb{P}_\rho\left[\sum_{x\in\square^-_m}\eta_x = M\right]\left(\frac{1}{2}\xi\cdot\overline{\mathbf{c}}\left(\frac{M\vee M_*\wedge M^*}{|\square^-_m|}, \square_n\right)\xi\right) \leqslant C(\lambda)3^{-dn}.$$

Therefore, under (6.41) and (6.43), the estimate in (6.46) has an upper bound for $\rho \in [0, 3^{-dn}) \cup (1-3^{-dn}, 1]$

$$\mathbf{I} \leqslant C(\lambda)3^{-dn}. \tag{6.49}$$

Combing (6.48) and (6.49), we conclude the uniform estimate for Step 1 that

$$\mathbf{I} \leqslant \left(\frac{1}{2}\xi\cdot\overline{\mathbf{c}}\,(\rho, \square_n)\,\xi\right) + C(\lambda)\left(3^{dn}\varepsilon + 3^{-dn} + 3^{d(6n-m)}\right). \tag{6.50}$$

Step 2: term **II**. This term is the contribution of the boundary layer. For this case, because the local correctors and the indicator are both invariant when applying

π_b, we have $\pi_b \hat{\phi}^{(\varepsilon)}_{m,n,\xi}(\eta) = 0$. The only contribution comes from the affine function, which gives us

$$\mathbf{II} \leqslant \sum_{\substack{b \in \overline{\square^*_m} \smallsetminus (\square_m, \square^-_m)^* \\ \mathrm{dist}(b, \partial \square_m) \leqslant 3^n}} \frac{1}{|\square_m|} \left\langle \frac{1}{2} c_b (\pi_b \ell_\xi)^2 \right\rangle_\rho \leqslant \lambda \chi(\rho) 3^{(n-m)}. \tag{6.51}$$

This term is small under the assumption (6.41).

Step 3: term **III**. This term is also the contribution of the boundary layer, but its estimate is more delicate. We denote by $b = \{y_1, y_2\}$, where $y_1 \in \square^-_m$ and $y_2 \in \partial \square_m$. The operator π_{y_1,y_2} makes difference only when $\eta_{y_1} \neq \eta_{y_2}$. For example, for the case $\eta_{y_1} = 1$, $\eta_{y_2} = 0$, we have

$$\begin{aligned}
&(\pi_{y_1,y_2} \hat{\phi}^{(\varepsilon)}_{m,n,\xi})(\eta) \mathbf{1}_{\{\eta_{y_1}=1, \eta_{y_2}=0\}} \\
&= \left(\hat{\phi}^{(\varepsilon)}_{m,n,\xi}(\eta^{y_1,y_2}) - \hat{\phi}^{(\varepsilon)}_{m,n,\xi}(\eta) \right) \mathbf{1}_{\{\eta_{y_1}=1, \eta_{y_2}=0\}} \\
&= \sum_{M \in \mathcal{M}_\varepsilon(\square^-_m)} \left(\sum_{\substack{z \in \mathcal{Z}_{m,n}, \\ \mathrm{dist}(z, \partial \square_m) > 3^n}} \left(\phi_{\frac{M}{|\square^-_m|}, z + \square_n, \xi} - \phi_{\frac{M+1}{|\square^-_m|}, z + \square_n, \xi} \right) \right) \mathbf{1}_{\left\{ \sum_{x \in \square^-_m, x \neq y_1} \eta_x = M \right\}} \mathbf{1}_{\{\eta_{y_1}=1, \eta_{y_2}=0\}}.
\end{aligned}$$

Here we have $\phi_{\rho, z+\square_n, \xi}(\eta^{y_1,y_2}) = \phi_{\rho, z+\square_n,\xi}(\eta)$, because y_1, y_2 are both far from the support of the local function when $3^n \geqslant \mathbf{r}$. Meanwhile, the operator π_{y_1,y_2} will change the empirical density, that is the perturbation in the third line. We should also remark that such perturbation will vanish when $M \notin \mathcal{M}_\varepsilon(\square^-_m)$, since there is a regularization of density $\rho \mapsto \rho \vee \varepsilon \wedge (1-\varepsilon)$ in our definition (6.34).

The case $\eta_{y_1} = 0, \eta_{y_2} = 1$ is similar. As the indicator $\mathbf{1}_{\{\eta_{y_1}=1, \eta_{y_2}=0\}}$ is independent of the other terms in the last line under $\mathbb{P}_\rho$, we obtain

$$\begin{aligned}
&\left\langle c_{y_1,y_2} (\pi_{y_1,y_2}(\ell_\xi + \hat{\phi}^{(\varepsilon)}_{m,n,\xi}))^2 \right\rangle_\rho \\
&\leqslant 2\lambda\chi(\rho) \left(1 + \sum_{M \in \mathcal{M}_\varepsilon(\square^-_m)} \mathbb{P}_\rho \left[\sum_{x \in \square^-_m, x \neq y_1} \eta_x = M \right] \left\langle \left(\sum_{\substack{z \in \mathcal{Z}_{m,n}, \\ \mathrm{dist}(z, \partial\square_m) > 3^n}} \Delta_{M,z} \right)^2 \right\rangle_{\square^-_m \smallsetminus \{y_1\}, M} \right),
\end{aligned} \tag{6.52}$$

where we define $\Delta_{M,z}$ to simplify the notation

$$\Delta_{M,z} := \phi_{\frac{M+1}{|\square^-_m|}, z+\square_n, \xi} - \phi_{\frac{M}{|\square^-_m|}, z+\square_n,\xi}, \qquad M \in \mathcal{M}_\varepsilon(\square^-_m). \tag{6.53}$$

Because there are $3^{(d-1)m}$ terms like (6.52) from boundary layer $b \in (\square_m, \square^-_m)^*$ and the normalization factor is $\frac{1}{|\square_m|}$, our object is to show that each term above is of order $O(3^{Kn+sm})$ with some $s \in (0,1)$ and $K \in \mathbb{R}_+$. Then roughly we get an estimate

$$\mathbf{III} \leqslant \sum_{b \in (\square_m, \square^-_m)^*} \frac{1}{|\square_m|} \left\langle \frac{1}{2} c_b (\pi_b(\ell_\xi + \hat{\phi}^{(\varepsilon)}_{m,n,\xi}))^2 \right\rangle_\rho \leqslant O(3^{Kn + (s-1)m}). \tag{6.54}$$

By a careful choice of $n \ll m$, we can make the contribution from **III** small.

To obtain the estimate above, we also need to treat the spatial cancellation in $\left(\sum_{\substack{z \in \mathcal{Z}_{m,n}, \\ \mathrm{dist}(z, \partial\square_m) > 3^n}} \Delta_{M,z} \right)^2$. Thus, we develop it as

$$\left\langle \left(\sum_{\substack{z \in \mathcal{Z}_{m,n}, \\ \operatorname{dist}(z, \partial \square_m) > 3^n}} \Delta_{M,z} \right)^2 \right\rangle_{\square_m^- \smallsetminus \{y_1\}, M} = \sum_{\substack{z \in \mathcal{Z}_{m,n}, \\ \operatorname{dist}(z, \partial \square_m) > 3^n}} \left\langle (\Delta_{M,z})^2 \right\rangle_{\square_m^- \smallsetminus \{y_1\}, M} + \sum_{\substack{z, z' \in \mathcal{Z}_{m,n}, z \neq z' \\ \operatorname{dist}(z, \partial \square_m) > 3^n, \operatorname{dist}(z', \partial \square_m) > 3^n}} \left\langle \Delta_{M,z} \Delta_{M,z'} \right\rangle_{\square_m^- \smallsetminus \{y_1\}, M}.$$

We treat the diagonal terms and off-diagonal terms separately.

Step 3.1: the diagonal term $\left\langle (\Delta_{M,z})^2 \right\rangle_{\square_m^- \smallsetminus \{y_1\}, M}$. For this term, we use the local equivalence of ensembles from (6.18)

$$\left\langle (\Delta_{M,z})^2 \right\rangle_{\square_m^- \smallsetminus \{y_1\}, M} \leqslant (1 + 4 \cdot 3^{d(2n-m)}) \left\langle (\Delta_{M,z})^2 \right\rangle_{\frac{M}{|\square_m^-| - 1}}. \tag{6.55}$$

Assuming that $n < \frac{m}{2}$ as (6.41), the factor $(1 + 4 \cdot 3^{d(2n-m)})$ is smaller than 2. To simplify the notation, we introduce the following shorthand expression

$$\hat{\rho} := \frac{M}{|\square_m^-|}, \qquad \hat{\rho}' := \frac{M+1}{|\square_m^-|}, \qquad \hat{\rho}'' := \frac{M}{|\square_m^-| - 1}. \tag{6.56}$$

Then for $M \in \mathcal{M}_\varepsilon(\square_m^-)$, as the difference between $\hat{\rho}, \hat{\rho}', \hat{\rho}''$ is roughly 3^{-dm} and the three terms are not degenerate at 0 or 1, we have an estimate for the two-sided bias factor of (6.6)

$$\forall M \in \mathcal{M}_\varepsilon(\square_m^-), \qquad 1 \leqslant \widetilde{\Theta}_{\hat{\rho}, \hat{\rho}'}, \widetilde{\Theta}_{\hat{\rho}, \hat{\rho}''} \leqslant 1 + 3^{-dm} \varepsilon^{-1}. \tag{6.57}$$

Insert the notation (6.56) to (6.53), the quantity $\Delta_{M,z}$ has a clear expression

$$\Delta_{M,z} = \phi_{\hat{\rho}', z + \square_n, \xi} - \phi_{\hat{\rho}, z + \square_n, \xi}, \tag{6.58}$$

and $\left\langle (\Delta_{M,z})^2 \right\rangle_{\frac{M}{|\square_m^-| - 1}}$ is just the expectation under the grand canonical ensemble of density $\hat{\rho}''$ for difference between the correctors of densities $\hat{\rho}$ and $\hat{\rho}'$

$$\left\langle (\Delta_{M,z})^2 \right\rangle_{\frac{M}{|\square_m^-| - 1}} = \left\langle (\phi_{\hat{\rho}', z + \square_n, \xi} - \phi_{\hat{\rho}, z + \square_n, \xi})^2 \right\rangle_{\hat{\rho}''}. \tag{6.59}$$

Therefore, we use the continuity in density (6.11) and (6.57)

$$\begin{aligned} \left\langle (\Delta_{M,z})^2 \right\rangle_{\square_m^- \smallsetminus \{y_1\}, M} &\leqslant 2 \left\langle (\Delta_{M,z})^2 \right\rangle_{\hat{\rho}''} \\ &\leqslant 10 \lambda \cdot 3^{(d+2)n} \left(\widetilde{\Theta}_{\hat{\rho}, \hat{\rho}''}^{(3^n + 2\mathbf{r})^d} + \widetilde{\Theta}_{\hat{\rho}', \hat{\rho}''}^{(3^n + 2\mathbf{r})^d} - 2 \right) \\ &\leqslant 20 \lambda \cdot 3^{(d+2)n} \cdot \left(\left(1 + 3^{-dm} \varepsilon^{-1}\right)^{3^{dn}} - 1 \right) \\ &\leqslant 20 \lambda \cdot 3^{(2d+2)n} \cdot 3^{-dm} \cdot \varepsilon^{-1}. \end{aligned} \tag{6.60}$$

Here we also need to add an assumption

$$3^{dn} \ll 3^{dm} \cdot \varepsilon. \tag{6.61}$$

Under this assumption, the contribution of the diagonal terms are

$$\begin{aligned} \sum_{\substack{z \in \mathcal{Z}_{m,n}, \\ \operatorname{dist}(z, \partial \square_m) > 3^n}} \left\langle (\Delta_{M,z})^2 \right\rangle_{\square_m^- \smallsetminus \{y_1\}, M} &\leqslant 20 \lambda \cdot 3^{d(m-n)} \cdot 3^{(2d+2)n} \cdot 3^{-dm} \cdot \varepsilon^{-1} \\ &= 20 \lambda \cdot 3^{(d+2)n} \cdot \varepsilon^{-1}. \end{aligned} \tag{6.62}$$

Step 3.2: the off-diagonal term $\langle \Delta_{M,z} \Delta_{M,z'} \rangle_{\square_m^- \smallsetminus \{y_1\}, M}$. Before starting the proof, we remark that the error in each off-diagonal term should be smaller than that of the diagonal term. A naive application of Cauchy–Schwarz inequality and (6.60) does not work here, because the factor 3^{-dm} in (6.60) is not enough to dominate the total number of $3^{2d(m-n)}$ off-diagonal terms.

We observe that $\langle \Delta_{M,z}\Delta_{M,z'}\rangle_{\square_m^- \smallsetminus \{y_1\},M}$ is actually asymptotically independent, and we need a decoupling inequality to justify it. Thus we make the following decomposition

$$
\begin{aligned}
&\langle \Delta_{M,z}\Delta_{M,z'}\rangle_{\square_m^- \smallsetminus \{y_1\},M} \\
&= \langle \Delta_{M,z}\Delta_{M,z'}\rangle_{\frac{M}{|\square_m^-|-1}} + \left(\langle \Delta_{M,z}\Delta_{M,z'}\rangle_{\square_m^- \smallsetminus \{y_1\},M} - \langle \Delta_{M,z}\Delta_{M,z'}\rangle_{\frac{M}{|\square_m^-|-1}} \right) \\
&= \langle \Delta_{M,z}\rangle^2_{\frac{M}{|\square_m^-|-1}} + \left(\langle \Delta_{M,z}\Delta_{M,z'}\rangle_{\square_m^- \smallsetminus \{y_1\},M} - \langle \Delta_{M,z}\Delta_{M,z'}\rangle_{\frac{M}{|\square_m^-|-1}} \right).
\end{aligned}
\tag{6.63}
$$

From the second line to the third line, we use the fact that under $\mathbb{P}_\rho$, the variables $\Delta_{M,z}$ and $\Delta_{M,z'}$ are independent and of the same law. We continue to estimate the two terms respectively.

For the first term in (6.63), we make use of the definition (6.58). It transforms the estimate as the regularity of mean

$$
\langle \Delta_{M,z}\rangle^2_{\frac{M}{|\square_m^-|-1}} = \left\langle \phi_{\hat\rho',z+\square_n,\xi} - \phi_{\hat\rho,z+\square_n,\xi}\right\rangle^2_{\hat\rho''},
$$

so (6.10) applies

$$
\begin{aligned}
\langle \Delta_{M,z}\rangle^2_{\frac{M}{|\square_m^-|-1}} &\leqslant 2\left\langle \phi_{\hat\rho',z+\square_n,\xi}\right\rangle^2_{\hat\rho''} + 2\left\langle \phi_{\hat\rho,z+\square_n,\xi}\right\rangle^2_{\hat\rho''} \\
&\leqslant 2\lambda\cdot 3^{(d+2)n}\left((\widetilde\Theta^{(3^n+2\mathbf{r})^d}_{\hat\rho,\hat\rho''} - 1)^2 + (\widetilde\Theta^{(3^n+2\mathbf{r})^d}_{\hat\rho',\hat\rho''} - 1)^2 \right) \\
&\leqslant 4\lambda\cdot 3^{(d+2)n}\cdot\left(\left(1 + 3^{-dm}\varepsilon^{-1}\right)^{3^{dn}} - 1\right)^2 \\
&\leqslant 4\lambda\cdot 3^{(3d+2)n}\cdot 3^{-2dm}\cdot\varepsilon^{-2}.
\end{aligned}
\tag{6.64}
$$

Here from the second line to the third line, we use the estimate of the two-sided bias factor in (6.57) for $M\in\mathcal{M}_\varepsilon(\square_m^-)$. We also need to assume (6.61) from the third line to the forth line. Compare the regularity of mean (6.10) and of L^2 (6.11), we gain another factor of type $(\widetilde\Theta^{(3^n+2\mathbf{r})^d}_{\hat\rho,\hat\rho''} - 1)$, that is why we have 3^{-2dm} in the last line.

For the second term in (6.63), we use the local equivalence of ensembles. Especially, since $\Delta_{M,z}\Delta_{M,z'}$ is not necessarily positive, we need apply the version of (6.19), which requires some supplementary conditions. These conditions are satisfied as we recall that $M\in\mathcal{M}_\varepsilon(\square_m^-)$ and assume (6.61). Then we obtain

$$
\begin{aligned}
&\left| \langle \Delta_{M,z}\Delta_{M,z'}\rangle_{\square_m^- \smallsetminus \{y_1\},M} - \langle \Delta_{M,z}\Delta_{M,z'}\rangle_{\frac{M}{|\square_m^-|-1}} \right| \\
&\leqslant 3^{d(2n-m)}\cdot\varepsilon^{-1}\cdot\langle |\Delta_{M,z}\Delta_{M,z'}|\rangle_{\frac{M}{|\square_m^-|-1}} \\
&\leqslant 3^{d(2n-m)}\cdot\varepsilon^{-1}\cdot\left\langle (\Delta_{M,z})^2\right\rangle_{\frac{M}{|\square_m^-|-1}} \\
&\leqslant 20\lambda\cdot 3^{(4d+2)n}\cdot 3^{-2dm}\cdot\varepsilon^{-2}.
\end{aligned}
\tag{6.65}
$$

Here we use Cauchy–Schwarz inequality from the second line to the third line, and insert the estimate (6.60) in the last line. Finally, we also gain the factor 3^{-2dm}.

We combine the estimates (6.64) and (6.65) and obtain an upper bound for the off-diagonal terms

$$\begin{aligned}(6.66)\qquad &\left|\sum_{\substack{z,z'\in\mathcal{Z}_{m,n}, z\neq z'\\ \mathrm{dist}(z,\partial\square_m)>3^n, \mathrm{dist}(z',\partial\square_m)>3^n}} \langle \Delta_{M,z}\Delta_{M,z'}\rangle_{\square_m^-\smallsetminus\{y_1\},M}\right| \\ &\leqslant 20\lambda\cdot 3^{2d(m-n)}\cdot 3^{(4d+2)n}\cdot 3^{-2dm}\cdot\varepsilon^{-2} \\ &= 20\lambda\cdot 3^{(2d+2)n}\cdot\varepsilon^{-2}.\end{aligned}$$

The error (6.66) from the off-diagonal terms is the leading order compared to (6.62) from the diagonal terms. We put them back to (6.54) and (6.52), and obtain the estimates for the contribution from **III**

$$\mathbf{III}\leqslant 20\lambda\cdot 3^{-m}\cdot\left(3^{(2d+2)n}\cdot\varepsilon^{-2}+3^{(d+2)n}\cdot\varepsilon^{-1}\right). \tag{6.67}$$

Step 4: track of parameters. We collect all the estimates from (6.50), (6.51) and (6.67), and put them back to (6.39) to obtain

$$\begin{aligned}\frac{1}{|\square_m|}\sum_{b\in\overline{\square_m^*}}\left\langle\frac{1}{2}c_b(\pi_b(\ell_\xi+\hat{\phi}^{(\varepsilon)}_{m,n,\xi}))^2\right\rangle_\rho &\leqslant\left(\frac{1}{2}\xi\cdot\overline{\mathbf{c}}\left(\rho,\square_n\right)\xi\right)\\ &\quad+C(\lambda)\left(3^{dn}\varepsilon+3^{-dn}+3^{d(6n-m)}+3^{(n-m)}+3^{(2d+2)n-m}\varepsilon^{-2}\right)\mathsf{Id}.\end{aligned}$$

The choice of parameters should satisfy the assumptions (6.41), (6.43), (6.61), and also make the upper bound above be small. A possible choice is

$$\varepsilon=3^{-2dn},\qquad n=\left\lfloor\frac{m}{9d+3}\right\rfloor,$$

which gives

$$\frac{1}{|\square_m|}\sum_{b\in\overline{\square_m^*}}\left\langle\frac{1}{2}c_b(\pi_b(\ell_\xi+\hat{\phi}^{(\varepsilon)}_{m,n,\xi}))^2\right\rangle_\rho\leqslant\left(\frac{1}{2}\xi\cdot\overline{\mathbf{c}}\left(\rho,\square_n\right)\xi\right)+C(\lambda)3^{-\frac{m}{12}}.$$

Then we apply the uniform estimate from Proposition 5.3, which gives us

$$\frac{1}{|\square_m|}\sum_{b\in\overline{\square_m^*}}\left\langle\frac{1}{2}c_b(\pi_b(\ell_\xi+\hat{\phi}^{(\varepsilon)}_{m,n,\xi}))^2\right\rangle_\rho\leqslant\frac{1}{2}\xi\cdot\overline{\mathbf{c}}\left(\rho\right)\xi+C(d,\lambda)3^{-\frac{(1\wedge\gamma_1)m}{12}}.$$

Together with (6.38), this concludes (6.37) by setting $\gamma_4:=\frac{(1\wedge\gamma_1)}{12}$.

□

Proof of Proposition 6.6. Viewing Proposition 4.1 and Corollary 4.3, for any $v\in\mathcal{F}_0(\Lambda^-)$, we have

$$\frac{1}{|\Lambda|}\sum_{b\in\overline{\Lambda^*}}\left\langle\frac{1}{2}c_b(\pi_b(\ell_\xi+v))^2\right\rangle_\rho\geqslant\frac{1}{2}\xi\cdot\overline{\mathbf{c}}(\rho,\Lambda)\xi\geqslant\frac{1}{2}\xi\cdot\overline{\mathbf{c}}(\rho)\xi,$$

thus $\mu(\Lambda,\xi)\geqslant 0$. Moreover, Proposition 6.7 implies

$$0\leqslant\mu(\square_m,\xi)\leqslant\frac{1}{|\square_m|}\left\langle\sum_{b\in\overline{\square_m^*}}\frac{1}{2}c_b(\pi_b(\ell_\xi+\hat{\phi}^{(\varepsilon)}_{m,n,\xi}))^2\right\rangle_\rho-\frac{1}{2}\xi\cdot\overline{\mathbf{c}}(\rho)\xi\leqslant C3^{-\gamma_4 m},$$

then we prove (6.33) along a subsequence $\{\square_m\}_{m\in\mathbb{N}_+}$. Like (4) of Proposition 4.1, we can prove that $\mu(\square_m,\xi)$ is a subadditive quantity, so Lemma A.1 applies and generalizes the uniform convergence (6.33) to general cubes Λ_L with $\gamma_3:=\gamma_4$. □

6.4. **Stationary corrector.** Recall that we have two definitions of the conductivity, that is (1.13) for $\mathbf{c}(\rho)$ and (5.2) for $\overline{\mathbf{c}}(\rho)$. In this part, we will establish the identity $\mathbf{c}(\rho) = \overline{\mathbf{c}}(\rho)$, so the results proved in the previous sections are indeed for the convergence to $\mathbf{c}(\rho)$. Secondly, we will give the concrete construction of Φ_L valid for (1.18).

We first establish several auxiliary lemmas on the stationary function. Recall that a mapping $y \mapsto f(y,\eta)$ is stationary iff $f(y,\eta) = \tau_y f(0,\eta) = f(0,\tau_y\eta)$.

Lemma 6.8. *The following properties hold.*

(1) *For every $x,y,z \in \mathbb{Z}^d$, we have*

$$\pi_{x+z,y+z}\tau_z = \tau_z\pi_{x,y}. \tag{6.68}$$

(2) *For every integer $1 \leqslant i \leqslant d$, every $\xi \in \mathbb{R}^d$ and every local function $f \in \mathcal{F}_0$, the following mapping is stationary.*

$$y \mapsto c_{y,y+e_i}\pi_{y,y+e_i}\left(\ell_\xi + \sum_{x\in\mathbb{Z}^d}\tau_x f\right)^2. \tag{6.69}$$

(3) *For every $\xi \in \mathbb{R}^d, f \in \mathcal{F}_0(\Lambda_L^-)$, we have*

$$\sum_{i=1}^d \left\langle c_{0,e_i}\left(\pi_{0,e_i}\big(\ell_\xi + \frac{1}{|\Lambda_L|}\sum_{x\in\mathbb{Z}^d}\tau_x f\big)\right)^2\right\rangle_\rho \leqslant \frac{1}{|\Lambda_L|}\sum_{b\in\overline{\Lambda_L^*}}\left\langle c_b\left(\pi_b\left(\ell_\xi + f\right)\right)^2\right\rangle_\rho. \tag{6.70}$$

Proof. The identity (6.68) can be proved by a direct verification. To prove (6.69), it suffices to prove

$$c_{y,y+e_i}\pi_{y,y+e_i}\left(\ell_\xi + \sum_{x\in\mathbb{Z}^d}\tau_x f\right)^2 = \tau_y\left(c_{0,0+e_i}\pi_{0,e_i}\left(\ell_\xi + \sum_{x\in\mathbb{Z}^d}\tau_x f\right)^2\right). \tag{6.71}$$

We start from its right-hand side and evaluate the term one by one. From the definition of $c_{y,y+e_i}$, we have $\tau_y c_{0,0+e_i} = c_{y,y+e_i}$. For the term ℓ_ξ, we have

$$\tau_y\pi_{0,e_i}\ell_\xi = \tau_y\left(-\xi\cdot e_i(\eta_{e_i} - \eta_0)\right) = -\xi\cdot e_i(\eta_{y+e_i} - \eta_y) = \pi_{y,y+e_i}\ell_\xi.$$

For the term involving f, we apply (6.68)

$$\begin{aligned}
\tau_y\pi_{0,e_i}\left(\sum_{x\in\mathbb{Z}^d}\tau_x f\right) &= \pi_{y,y+e_i}\tau_y\left(\sum_{x\in\mathbb{Z}^d}\tau_x f\right)\\
&= \pi_{y,y+e_i}\left(\sum_{x\in\mathbb{Z}^d}\tau_{x+y} f\right)\\
&= \pi_{y,y+e_i}\left(\sum_{x\in\mathbb{Z}^d}\tau_x f\right).
\end{aligned}$$

Then we obtain (6.71) and establish the stationary property.

Finally, we verify the inequality (6.70). Notice that $\pi_{y,y+z}\tau_x\ell_\xi = \pi_{y,y+z}\ell_\xi$, we have

$$\begin{aligned}
&\sum_{i=1}^d \left\langle c_{0,e_i}\left(\pi_{0,e_i}\Big(\ell_\xi + \frac{1}{|\Lambda_L|}\sum_{x\in\mathbb{Z}^d}\tau_x f\Big)\right)^2\right\rangle_\rho \\
&= \sum_{i=1}^d \left\langle c_{0,e_i}\left(\pi_{0,e_i}\Big(\frac{1}{|\Lambda_L|}\sum_{x\in\Lambda_L}\tau_x(\ell_\xi + f)\Big)\right)^2\right\rangle_\rho \\
&\leqslant \sum_{i=1}^d \frac{1}{|\Lambda_L|}\sum_{x\in\Lambda_L}\left\langle c_{0,e_i}\left(\pi_{0,e_i}\left(\tau_x(\ell_\xi+f)\right)\right)^2\right\rangle_\rho .
\end{aligned} \tag{6.72}$$

From the first line to the second line, we also use the fact that $f\in\mathcal{F}_0(\Lambda_L^-)$, so the derivative vanish when translation is outside Λ_L. From the second line to the third line above, we apply Jensen's inequality. Then we simplify the result

$$\begin{aligned}
\left\langle c_{0,e_i}\left(\pi_{0,e_i}\left(\tau_x(\ell_\xi+f)\right)\right)^2\right\rangle_\rho &= \left\langle \tau_{-x}c_{0,e_i}\left(\pi_{0,e_i}\left(\tau_x(\ell_\xi+f)\right)\right)^2\right\rangle_\rho \\
&= \left\langle c_{-x,-x+e_i}\left(\pi_{-x,-x+e_i}\left(\ell_\xi+f\right)\right)^2\right\rangle_\rho .
\end{aligned}$$

The equality in the first line comes from the stationarity of $\mathbb{P}_\rho$, i.e. $\langle\tau_x F\rangle_\rho = \langle F\rangle_\rho$ for all $x\in\mathbb{Z}^d$, while the equality from the second line comes from (6.68). We put this identity back to (6.72) and conclude the desired result

$$\begin{aligned}
&\sum_{i=1}^d \left\langle c_{0,e_i}\left(\pi_{0,e_i}\Big(\ell_\xi + \frac{1}{|\Lambda_L|}\sum_{x\in\Lambda_L}\tau_x f\Big)\right)^2\right\rangle_\rho \\
&\leqslant \sum_{i=1}^d \frac{1}{|\Lambda_L|}\sum_{x\in\Lambda_L}\left\langle c_{-x,-x+e_i}\left(\pi_{-x,-x+e_i}\left(\ell_\xi+f\right)\right)^2\right\rangle_\rho \\
&\leqslant \frac{1}{|\Lambda_L|}\sum_{b\in\Lambda_L^*}\left\langle c_b\left(\pi_b\left(\ell_\xi+f\right)\right)^2\right\rangle_\rho .
\end{aligned}$$

□

With the preparation of Lemma 6.8 and Proposition 6.6, we can now prove the main result in this subsection.

Proposition 6.9. *For every $\rho\in[0,1]$, the quantities defined in (1.13) and (5.2) coincide*

$$\mathbf{c}(\rho) = \overline{\mathbf{c}}(\rho).$$

Moreover, let $\phi_{\Lambda,\xi}$ be an optimizer for $\mu(\Lambda,\xi)$ defined in (6.32) and $\Phi_L\in\mathcal{F}_0^d(\Lambda_L)$ be defined as

$$\Phi_L := \left(\frac{1}{|\Lambda_L|}\phi_{\Lambda_L,e_1}, \frac{1}{|\Lambda_L|}\phi_{\Lambda_L,e_2}, \cdots, \frac{1}{|\Lambda_L|}\phi_{\Lambda_L,e_d}\right), \tag{6.73}$$

then there exists an exponent $\gamma(d,\lambda,\mathbf{r})>0$ and a positive constant $C(d,\lambda,\mathbf{r})<\infty$, such that

$$\sup_{\rho\in[0,1]} |\mathbf{c}(\rho;\Phi_L) - \mathbf{c}(\rho)| \leqslant CL^{-\gamma}. \tag{6.74}$$

Proof. The proof is similar to the sandwich argument in (6.27) and can be divided the into three steps. Similar arguments can be found in [35, Theorem B.1].

Step 1: $\mathbf{c}(\rho) \leqslant \overline{\mathbf{c}}(\rho)$. We recall (1.12) that

$$\xi \cdot \mathbf{c}(\rho; \Phi_L)\xi = \sum_{i=1}^{d} \left\langle c_{0,e_i} \left(\xi \cdot \left\{ e_i(\eta_{e_i} - \eta_0) - \pi_{0,e_i}\Big(\sum_{y \in \mathbb{Z}^d} \tau_y \Phi_L \Big) \right\} \right)^2 \right\rangle_\rho .$$

and notice the fact that

$$\begin{aligned} \xi \cdot x(\eta_x - \eta_0) &= -\pi_{0,x}\ell_\xi, \\ \xi \cdot \Phi_L &= \frac{1}{|\Lambda_L|}\phi_{\Lambda_L,\xi}, \end{aligned}$$

where the second identity comes from the linear map $\xi \mapsto \phi_{\Lambda_L,\xi}$. Therefore, we have

$$\begin{aligned} \xi \cdot \mathbf{c}(\rho; \Phi_L)\xi &= \sum_{i=1}^{d} \left\langle c_{0,e_i} \left(\pi_{0,e_i} \left(\ell_\xi + \frac{1}{|\Lambda_L|} \sum_{y \in \mathbb{Z}^d} \tau_y \phi_{\Lambda_L,\xi} \right) \right)^2 \right\rangle_\rho \\ &\leqslant \frac{1}{|\Lambda_L|} \sum_{b \in \Lambda_L^*} \left\langle c_b \left(\pi_b \left(\ell_\xi + \phi_{\Lambda_L,\xi} \right) \right)^2 \right\rangle_\rho \\ &\leqslant \xi \cdot \overline{\mathbf{c}}(\rho)\xi + \mu(\Lambda_L, \xi), \end{aligned}$$

where we apply (6.70) from the first line to the second line and (6.32) from the second line to the third line. Using the estimate (6.33) about $\mu(\Lambda_L, \xi)$, we obtain an important inequality chain with C, γ_3 independent of $\rho \in [0,1]$

$$\xi \cdot \mathbf{c}(\rho)\xi \leqslant \xi \cdot \mathbf{c}(\rho; \Phi_L)\xi \leqslant \xi \cdot \overline{\mathbf{c}}(\rho)\xi + CL^{-\gamma_3}|\xi|^2. \tag{6.75}$$

We take the first one and the third one, and let $L \nearrow \infty$, and obtain the desired result

$$\mathbf{c}(\rho) \leqslant \lim_{L \to \infty} \overline{\mathbf{c}}(\rho, \Lambda_L) = \overline{\mathbf{c}}(\rho).$$

Step 2: $\mathbf{c}(\rho) \geqslant \overline{\mathbf{c}}(\rho)$. We pick F_K as a sequence of local functions to approximate $\mathbf{c}(\rho)$, and compare $v_{p,K} = \ell_p + \sum_{x \in \mathbb{Z}^d} p \cdot \tau_x F_K$ in the functional of $\overline{\nu}_*(\rho, \Lambda_L, q)$

$$\begin{aligned} \overline{\nu}_*(\rho, \Lambda_L, q) &= \frac{1}{2} q \cdot \overline{D}_*^{-1}(\rho, \Lambda_L) \cdot q \\ &\geqslant \frac{1}{2\chi(\rho)|\Lambda|} \sum_{b \in \overline{\Lambda^*}} \left\langle (\pi_b \ell_q)(\pi_b v_{p,K}) - \frac{1}{2} c_b (\pi_b v_{p,K})^2 \right\rangle_\rho . \end{aligned}$$

Applying the stationarity of mapping in (6.69), we obtain that

$$\frac{1}{2} q \cdot \overline{D}_*^{-1}(\rho, \Lambda_L) q \geqslant p \cdot q - \frac{1}{4\chi(\rho)} p \cdot \mathbf{c}(\rho; F_K) p.$$

Letting $L, K \to \infty$, this yields

$$\frac{1}{2} q \cdot \overline{D}^{-1}(\rho) q \geqslant p \cdot q - \frac{1}{4\chi(\rho)} p \cdot \mathbf{c}(\rho) p.$$

By taking $q = \overline{D}(\rho)p$ and recall $2\chi(\rho)\overline{D}(\rho) = \overline{\mathbf{c}}(\rho)$ from (5.2), we conclude $\mathbf{c}(\rho) \geqslant \overline{\mathbf{c}}(\rho)$.

Step 3: Error estimate. Once we identify that $\mathbf{c}(\rho) = \overline{\mathbf{c}}(\rho)$, we go back to (6.75) to obtain

$$\xi \cdot \mathbf{c}(\rho)\xi \leqslant \xi \cdot \mathbf{c}(\rho; \Phi_L)\xi \leqslant \xi \cdot \mathbf{c}(\rho)\xi + CL^{-\gamma_3}|\xi|^2.$$

Since C, γ_3 come from Proposition 6.6 and are independent of ρ, this concludes the estimate (6.74) by setting $\gamma := \gamma_3$. $\square$

6.5. **Proof of Theorem 1.2 and Theorem 1.4.** We resume the proof of the main theorems in this paper.

Proof of Theorem 1.2 and Theorem 1.4. The finite-volume approximation defined in (1.20) gives a subadditive quantity $\overline{\mathbf{c}}(\rho, \Lambda_L)$, which defines a limit $\overline{\mathbf{c}}(\rho)$ in (5.2). Using the dual quantity, we prove the convergence rate of this approximation in Proposition 5.1 and Proposition 6.4. In Lemma 6.8, we identify that $\overline{\mathbf{c}}(\rho)$ coincides with $\mathbf{c}(\rho)$ defined in (1.13), therefore Proposition 5.1 together with Proposition 6.4 proves Theorem 1.4. Finally, Proposition 6.6 removes the dependence of density of the local corrector and Lemma 6.8 proves Theorem 1.2, where (6.73) gives a concrete construction of the density-free local corrector. □

7. Quantitative hydrodynamic limit

In this part, we reveal the connection between homogenization and hydrodynamic limit. The main task for establishing the hydrodynamic limit for non-gradient models lies in showing the gradient replacement, that is, a replacement of a diverging term, appearing in the scaling and caused by the non-gradient nature of the microscopic current, by a well-behaving gradient term under a large scale space-time sample average. For this, one needs Varadhan's lemma, that is, the characterization of the closed forms. So far, this method is the only mathematically rigorous way to deal with the non-gradient models, however it is usually hard to prove. In this section, we propose a new approach with two advantages: it avoids to show Varadhan's lemma, and it gives the convergence rate for the hydrodynamic limit.

As briefly explained in Section 1.1, the hydrodynamic limit was studied in [34] for the non-gradient Kawasaki dynamics on the lattice torus $\mathbb{T}_N^d$. It is described as a Markov process $\eta^N(t) = \{\eta_x^N(t), x \in \mathbb{T}_N^d\}$ on the configuration space $\mathcal{X}_N = \{0,1\}^{\mathbb{T}_N^d}$ governed by the infinitesimal generator $\mathcal{L}_N = N^2\mathcal{L}$, where $\mathcal{L}$ is the operator defined by (1.6) replacing $\mathbb{Z}^d$ with $\mathbb{T}_N^d$. The hydrodynamic limit is the problem to study the asymptotic behavior as $N \to \infty$ of the macroscopic empirical mass distribution $\rho^N(t, dv)$ of $\eta^N(t)$ defined by (1.8) and the nonlinear diffusion equation (1.9) was derived in the limit; see also the discussions around (1.14).

7.1. **Convergence rate in gradient replacement.** As we pointed, one main step in non-gradient hydrodynamic limit was to show the gradient replacement. In the previous work [34], it is Theorem 3.2 together with Lemma 3.4. The formula in Theorem 3.2 still contains diverging factor N in N^{1-d} but it can be absorbed as in Lemma 3.4 due to the gradient property of the function A defined by (7.3) below. Theorem 3.2 follows from Corollary 5.1 or especially from Theorem 5.1 for which Varadhan's lemma formulated in Theorem 4.1 was used. Varadhan's lemma roughly claims that the closed forms are determined from and spanned by the functions v in the variational formula (1.20) or (1.23) for the conductivity $\mathbf{c}(\rho)$. Physically, the linear term appears as the first order term in the Taylor expansion of the local equilibrium state of second order approximation (see (7.14) below), which corresponds to the corrector in the homogenization theory.

Let us start with a refinement of Theorem 5.1 and Corollary 5.1 in [34]. For any $\Lambda \subseteq \mathbb{Z}^d$, we denote by $\mathcal{X}_\Lambda = \{0,1\}^\Lambda$. Using the notation (1.42), every configuration $\eta \in \mathcal{X}$ can be decomposed into two parts

$$\eta = \xi \cdot \zeta, \qquad \xi := \eta_{|\Lambda} \in \mathcal{X}_\Lambda, \qquad \zeta := \eta_{|\Lambda^c} \in \mathcal{X}_{\Lambda^c}. \tag{7.1}$$

We let $\mathcal{L}_{\Lambda,\zeta}$ be a restriction of the generator $\mathcal{L}$ defined in (1.6) on Λ with the exterior condition ζ

$$\mathcal{L}_{\Lambda,\zeta} v := \sum_{b \in \Lambda^*} c_b \pi_b v. \tag{7.2}$$

Using the decomposition $\eta = \xi \cdot \zeta$ from (7.1) on the domain $\Lambda_L = (-L/2, L/2) \cap \mathbb{Z}^d$, we define the quantity $A_L, B_{L,\zeta}$ and $H_{L,\zeta,F}$ as

$$\begin{aligned} A_L(\xi) &:= \frac{1}{2} \sum_{x,y \in \Lambda_L : |x-y|=1} (\xi_y - \xi_x)(y - x), \\ B_{L,\zeta}(\xi) &:= \frac{1}{2} \sum_{x,y \in \Lambda_L : |x-y|=1} W_{x,y}(\eta)(y - x) = -\mathcal{L}_{\Lambda_L,\zeta}\Big(\sum_{x \in \Lambda_L} x \xi_x\Big), \\ H_{L,\zeta,F}(\xi) &:= \sum_{x \in \Lambda_{L-r(F)-1}} \tau_x(\mathcal{L}_N F)(\eta) = \mathcal{L}_{\Lambda_L,\zeta}\left(\sum_{x \in \Lambda_{L-r(F)-1}} \tau_x F\right)(\xi). \end{aligned} \tag{7.3}$$

Note that $W_{x,y}(\eta) := c_{x,y}(\eta)(\eta_y - \eta_x)$ in $B_{L,\zeta}$ denotes the microscopic current, while $F = (F_i)_{i=1}^d \in \mathcal{F}_0^d$ in $H_{L,\zeta,F}$ will be taken as the function appearing in the local equilibrium state of second order approximation ψ_t in (7.14) defined in next section . The correction function F should be local compared to Λ_L, so we require $r(F) \leqslant L - 1$, where $r(F)$ measures the diameter of the support of F

$$r(F) := \min \left\{ r \in \mathbb{N}_+ : F \in \mathcal{F}_0^d(\Lambda_r) \right\}. \tag{7.4}$$

For a function $f \in \mathcal{F}_0$ such that $\langle f \rangle_{\Lambda_L,M} = 0$, its CLT variance is defined by

$$\Delta_{L,M,\zeta}[f] := \Delta_{L,M,\zeta}[f, f],$$

for $0 \leqslant M \leqslant |\Lambda_L|$ and $\zeta \in \mathcal{X}_{\Lambda_L^c} = \{0,1\}^{\Lambda_L^c}$, where

$$\Delta_{L,M,\zeta}[f, g] := \left\langle f (-\mathcal{L}_{\Lambda_L,\zeta})^{-1} g \right\rangle_{\Lambda_L,M},$$

The CLT variance for the gradient replacement is defined by

$$\begin{aligned} Q_L(F; q, M, \zeta) := |\Lambda_L|^{-1} \Delta_{L,M,\zeta} &\left[q \cdot \{ D(M/|\Lambda_L|) A_L - (B_{L,\zeta} - H_{L,\zeta,F}) \} \right] \\ &- \frac{1}{2} q \cdot R(M/|\Lambda_L|; F) q, \end{aligned} \tag{7.5}$$

where $R(\rho; F) = \mathbf{c}(\rho; F) - \mathbf{c}(\rho)$ is defined in (1.16).

The following proposition states that, when taking F such that $R(\rho; F)$ is very small, the non-gradient term $B_{L,\zeta}$ can be replaced by a gradient term A_L plus a term $H_{L,\zeta,F}$ vanishing under the time integral and scaling. Proposition 7.1 provides the convergence rate in the gradient replacement, which is a quantitative refinement of the result shown in Corollary 5.1 of [34] and is a key for proving the hydrodynamic limit in non-gradient model.

Proposition 7.1. *There exists a positive constant $C(d, \lambda, \mathbf{r})$ such that the quantity Q_L defined in (7.5) satisfies the following estimate for all $L \in \mathbb{N}_+$ and $F \in \mathcal{F}_0^d$*

$$\begin{aligned} &\sup_{q,M,\zeta} |Q_L(F; q, M, \zeta)| \\ &\qquad \leqslant C \left(L^{-\alpha_1} + r(F)^d (1 + r(F)^{2d} \|F\|_\infty^2) L^{-1} + r(F)^{2d} \|F\|_\infty L^{-d} + C L^{-1} \right). \end{aligned} \tag{7.6}$$

Here the supreme is taken over all $q \in B_1$, $0 \leqslant M \leqslant |\Lambda_L|$ and $\zeta \in \mathcal{X}_{\Lambda_L^c}$.

Proof. We denote by $\hat{\rho} := \frac{M}{|\Lambda_L|}$ as the empirical density to simplify the notation. As proposed in [34, Proposition 5.1 and Theorem 5.1], the estimate of $Q_L(F; q, M, \zeta)$ can be reduced to the following three terms

$$
\begin{aligned}
Q_L^{(4)}(F; q, M, \zeta) &:= |\Lambda_L|^{-1} \Delta_{L,M,\zeta}\left[q \cdot (B_{L,\zeta} - H_{L,\zeta,F})\right] - \frac{1}{2} q \cdot \mathbf{c}(\hat{\rho}; F) q, \\
Q_L^{(5)}(F; q, \widetilde{q}, M, \zeta) &:= |\Lambda_L|^{-1} \Delta_{L,M,\zeta}\left[\widetilde{q} \cdot A_L, q \cdot (B_{L,\zeta} - H_{L,\zeta,F})\right] - (q \cdot \widetilde{q}) \chi(\hat{\rho}), \\
Q_L^{(6)}(\widetilde{q}, M, \zeta) &:= |\Lambda_L|^{-1} \Delta_{L,M,\zeta}\left[\widetilde{q} \cdot A_L\right] - 2(\widetilde{q} \cdot \mathbf{c}^{-1}(\hat{\rho}) \widetilde{q}) \chi^2(\hat{\rho}).
\end{aligned}
\tag{7.7}
$$

By choosing $\widetilde{q} = D(\hat{\rho}) q$ in (7.3), we obtain immediately an identity of Q_L

$$
\begin{aligned}
&Q_L(F; q, M, \zeta) \\
&\qquad = Q_L^{(4)}(F; q, M, \zeta) - 2 Q_L^{(5)}(F; q, D(\hat{\rho}) q, M, \zeta) + Q_L^{(6)}(D(\hat{\rho}) q, M, \zeta).
\end{aligned}
\tag{7.8}
$$

To show (7.6), we need the error estimate for each term. Among them, the quantitative estimates of $Q_L^{(4)}$ and $Q_L^{(5)}$ are given in Proposition 4.1 of [32]: for $q, \widetilde{q} \in \mathbb{R}^d$ such that $|q| = |\widetilde{q}| = 1$,

$$
|Q_L^{(4)}(F; q, M, \zeta)| \leqslant C r(F)^d (1 + r(F)^{2d} \|F\|_\infty^2) L^{-1}, \tag{7.9}
$$

$$
|Q_L^{(5)}(F; q, \widetilde{q}, M, \zeta)| \leqslant C r(F)^{2d} \|F\|_\infty L^{-d} + C L^{-1}. \tag{7.10}
$$

The computation of $Q_L^{(6)}$, as the error for the CLT variance of A_L, is indeed deeply related to the dual quantity studied in (6.14), and this makes Varadhan's lemma avoidable. It is, in a sense, hidden in the variational formula for $\widehat{D}_*(\Lambda_L, M)$. In other words, our dual computation well fits to computing the CLT variance of A_L. This matches with the dual computation in (5.7) in the proof of Theorem 5.1 in [34], but our computation is presented at higher level on the configuration space. We can write the error $Q_L^{(6)}$ of the CLT variance of A_L defined by (7.7) concretely in terms of the dual quantity as in (7.11) below. It is important that we can give such an exact formula before taking the limit so that one can directly apply our estimate to obtain its convergence rate. Note that we need the uniformity in the density $\rho \in [0,1]$ as in Corollary 5.1 in [34].

Let us implement the discussion above to $Q_L^{(6)}$. From (7.3) and (1.44), the quantity $q \cdot A_L$ satisfies the identity

$$
q \cdot A_L = - \sum_{b = \{x,y\} \in (\Lambda_L)^*} \pi_b \ell_q.
$$

Then by variational formula of the dual quantity $\widehat{D}_*(\Lambda_L, M)$ defined in (6.14), the optimiser $u(\Lambda_L, q)$ satisfies that

$$
\sum_{b \in (\Lambda_L)^*} c_b \pi_b u(\Lambda_L, q) = \sum_{b \in (\Lambda_L)^*} \pi_b \ell_q.
$$

Recall the generator $\mathcal{L}_{\Lambda_L, \zeta}$ defined in (7.2), thus the quantity $(-\mathcal{L}_{\Lambda_L,\zeta})^{-1}(q \cdot A_L)$ is nothing but $u(\Lambda_L, q)$ defined in (1) of Proposition 4.1 (see also Remark 4.2) because of the following identity

$$
(-\mathcal{L}_{\Lambda_L,\zeta})^{-1}(q \cdot A_L) = (\mathcal{L}_{\Lambda_L,\zeta})^{-1}\left(\sum_{b \in (\Lambda_L)^*} \pi_b \ell_q\right) = u(\Lambda_L, q).
$$

Here we abuse the notation because the generator defined in (7.2) from [34] is slightly different from (1.21). Nevertheless, this tiny difference does no harm; see Remark 1.5.

Therefore, we put this result back to $\Delta_{L,M,\zeta}[q \cdot A_L]$ and obtain

$$\begin{aligned}\Delta_{L,M,\zeta}[q \cdot A_L] &= \left\langle (q \cdot A_L)(-\mathcal{L}_{\Lambda_L,\zeta})^{-1}(q \cdot A_L)\right\rangle_{\Lambda_L,M} \\ &= \langle (q \cdot A_L) u(\Lambda_L, q)\rangle_{\Lambda_L,M} \\ &= -\sum_{b \in (\Lambda_L)^*} \langle (\pi_b \ell_q) u(\Lambda_L, q)\rangle_{\Lambda_L,M} \\ &= \frac{1}{2}\sum_{b \in (\Lambda_L)^*} \langle (\pi_b \ell_q)(\pi_b u(\Lambda_L, q))\rangle_{\Lambda_L,M} \\ &= \frac{1}{2}\sum_{b \in (\Lambda_L)^*} \left\langle c_b (\pi_b u(\Lambda_L, q))^2\right\rangle_{\Lambda_L,M}.\end{aligned}$$

Here from the third line to the forth line, we use the identity $\pi_b\pi_b = -2\pi_b$ and integration by part. From the forth line to the fifth line, we use the variational formula of $\widehat{D}_*(\Lambda_L, M)$ once again rewritten similar to (4.3). This concludes that

$$|\Lambda_L|^{-1}\Delta_{L,M,\zeta}[q \cdot A_L] = (2\chi(\hat{\rho}))\left(\frac{1}{2} q \cdot \widehat{D}_*^{-1}(\Lambda_L, M) q\right) = \chi(\hat{\rho}) q \cdot \widehat{D}_*^{-1}(\Lambda_L, M) q.$$

We put this result back to (7.7), and obtain that

$$\begin{aligned}Q_L^{(6)}(q, M, \zeta) &= \chi(\hat{\rho}) q \cdot \left(\widehat{D}_*^{-1}(\Lambda_L, M) - D^{-1}(\hat{\rho})\right) q \\ &= \chi(\hat{\rho}) q \cdot \widehat{D}_*^{-1}(\Lambda_L, M)\left(D(\hat{\rho}) - \widehat{D}_*(\Lambda_L, M)\right) D^{-1}(\hat{\rho}) q \\ &= \frac{1}{2} q \cdot \widehat{D}_*^{-1}(\Lambda_L, M)\left(\mathbf{c}(\hat{\rho}) - \hat{\mathbf{c}}_*(\Lambda_L, M)\right) D^{-1}(\hat{\rho}) q.\end{aligned} \tag{7.11}$$

Apply the quantitative homogenization of canonical ensemble in (1.25) (see also Proposition 6.4 and $\overline{\mathbf{c}}(\rho) = \mathbf{c}(\rho)$ proved in Proposition 6.9), noting the uniform positivity of matrices $\widehat{D}_*(\Lambda_L, M)$ and $D(\hat{\rho})$ in (2) of Proposition 4.1, we obtain that

$$|Q_L^{(6)}(q, M, \zeta)| \leqslant C L^{-\alpha_1}, \tag{7.12}$$

for some $C, \alpha_1 > 0$. The estimates (7.9), (7.10) and (7.12) together apply to (7.8) and yield (7.6).

□

7.2. **Application to the quantitative hydrodynamic limit.** The proof of the hydrodynamic limit in [34] relies on the relative entropy method in which we compare the distribution of our system $\eta^N(t)$ with the local equilibrium state of second order approximation.

Let $h_N(f|\psi)$ be the relative entropy per volume for two probability densities f and ψ with respect to $\nu_{1/2}^N$ defined as

$$h_N(f|\psi) := N^{-d}\int_{\mathcal{X}_N} f \log(f/\psi)\, \mathrm{d}\nu_{1/2}^N. \tag{7.13}$$

Here $\nu_{1/2}^N$ is the Bernoulli product measure on $\mathcal{X}_N$ with mean 1/2. The local equilibrium state of second order approximation $\psi_t \equiv \psi_t^N$ was defined by

$$\psi_t^N(\eta) := Z_t^{-1}\exp\left\{\sum_{x \in \mathbb{T}_N^d} \lambda(t, x/N)\eta_x + \frac{1}{N}\sum_{x \in \mathbb{T}_N^d} \partial\lambda(t, x/N) \cdot \tau_x F_N(\eta)\right\}, \tag{7.14}$$

where Z_t is a normalization constant with respect to the Bernoulli product measure $\nu_{1/2}^N$ and $\partial\lambda\,(\equiv \nabla\lambda) = (\partial_i\lambda)_{i=1}^d$; see p.5 of [34] or (2.1) in [32]. Here, we determine $\lambda(t, v)$ in ψ_t^N as $\lambda(t, v) = \overline{\lambda}(\rho(t, v))$ from the solution $\rho(t, v)$ of the hydrodynamic equation (1.9) with $\overline{\lambda}(\rho) = \log\{\rho/(1-\rho)\}$ and $F \equiv F_N := \Phi_{n(N)}$ with Φ_n defined in (6.73) with $n = n(N)$ to be determined later.

The choice of $F \equiv F_N := \Phi_{n(N)}$ is the main result obtained in this paper, as it gives the decay rate for $R(\rho; F) = \mathbf{c}(\rho; F) - \mathbf{c}(\rho)$. Following Proposition 6.9 and Lemma 4.4, one can find a sequence of functions Φ_n such that

$$r(\Phi_n) \leqslant n, \qquad \|\Phi_n\|_\infty \leqslant C_2 n^2 \log n, \qquad \sup_{\rho\in[0,1]} |R(\rho;\Phi_n)| \leqslant C_2 n^{-\alpha_2}. \tag{7.15}$$

for some $C_2, \alpha_2 > 0$ independent of ρ.

Then, we are now ready to give the proof of Theorem 1.3.

Proof of Theorem 1.3. We consider the relative entropy $h_N(t) := h_N(f_t^N|\psi_t^N)$, where we denote by f_t^N the density of the distribution of $\eta^N(t)$ on $\mathcal{X}_N$ with respect to $\nu_{1/2}^N$. The proof is given by a combination of a bound for $h_N(t)$, the entropy inequality and a large deviation estimate for ψ_t^N. Besides the previous work [34], we also borrow some estimates from [32], where the hydrodynamic limit for a non-gradient Glauber–Kawasaki dynamics is studied.

Step 1: relative entropy estimate. A bound for $h_N(t)$ was obtained for Glauber–Kawasaki dynamics with a strength K in the Glauber part in (3.64) of [32]. Our Kawasaki dynamics corresponds to the case of $K = 0$ and, picking the contribution from Kawasaki part only, we obtain the estimate

$$\begin{aligned} 0 \leqslant h_N(t) \leqslant e^{t/\delta}\Big(& h_N(0) + Q^{\Omega_1}_{N,L,\beta}(\lambda, F_N) + Q^{LD}_{N,L,\delta}(\lambda, F_N) \\ & + C(\beta+1)\, \|\partial\lambda\|_\infty^2 \sup_{\rho\in[0,1]} |R(\rho;F_N)| \\ & + Q^{En}_N(\lambda, F_N) + Q^{\Omega_2}_{N,L}(\lambda, F_N)\Big), \end{aligned} \tag{7.16}$$

for every $\beta > 0$, $1 \leqslant L \leqslant N/2$, $t \in [0,T]$ and $\delta > 0$ sufficiently small. Note that a bound for $h(t) = \overline{\lim}_{N\to\infty} h_N(t)$ is given in Theorem 2.1 in [34], but we actually have an error term $o(1)$ for $h_N(t)$ as in Lemma 3.2 in [34]. In the above estimate (7.16) (which is obtained after applying Gronwall's inequality), we give an estimate for the term $o(1)$ clarifying its decay rate.

Term 1: $Q^{\Omega_1}_{N,L,\beta}(\lambda, F)$. In (7.16), the first error $Q^{\Omega_1}_{N,L,\beta}(\lambda, F)$ for the microscopic current of non-gradient type adjusting with the corrector F consists of four terms as

$$Q^{\Omega_1}_{N,L,\beta}(\lambda, F) = \frac{C}{\beta} + \beta^2 Q^{(1)}_{N,L}(\lambda, F) + \beta Q^{(2)}_L(\lambda, F) + Q^{(3)}_{N,L}(\lambda),$$

with $F = F_N$. Here, the term $Q^{(1)}_{N,L}$ has a bound

$$|Q^{(1)}_{N,L}| \leqslant C N^{-1} L^{2d+4} (1 + r(F)^{3d} \|F\|_\infty^3), \tag{7.17}$$

see Theorem 3.5 of [32]. Note that λ does not depend on K in the present setting so that $\|\partial\lambda\|_\infty$ is bounded by C. Note also that $\frac{C}{\beta}K$ in Lemma 3.6 of [32] may be replaced by $\frac{C}{\beta}$ which is needed as a cost for the entropy bound, though we have no Glauber part. The term $Q^{(2)}_L \equiv Q^{(2)}_L(\lambda, F)$ is estimated by [32, (3.31)] as

$$|Q^{(2)}_L| \leqslant C\, \|\partial\lambda\|_\infty^2 \sup_{q,M,\zeta} |Q_L(F; q, M, \zeta)|, \tag{7.18}$$

so we can cite the estimate of the gradient replacement (7.6) in Section 7.1 directly.

The term $Q^{(3)}_{N,L} \equiv Q^{(3)}_{N,L}(\lambda)$ does not depend on F and it is estimated from above as

$$Q^{(3)}_{N,L} \leqslant C(L^{-1}+N^{-1}) + C\beta L^{-1} + \frac{C}{\beta} + C\beta^2 N^{-1} L^{2d-2}, \tag{7.19}$$

see Lemma 3.7 in [32].

Term 2: $Q^{LD}_{N,L,\delta}(\lambda,F)$. The error $Q^{LD}_{N,L,\delta}(\lambda,F)$ appearing in a large deviation bound is estimated as

$$|Q^{LD}_{N,L,\delta}(\lambda,F)| \leqslant CN^{-1}(L+\|F\|_\infty) + CL^{-d}\log L. \tag{7.20}$$

This estimate is shown in Theorem 3.9 of [32] for Glauber–Kawasaki dynamics, and for a function $G = G_1 + \frac{1}{K}G_2$ with G_1 and G_2 coming respectively from the Glauber and Kawasaki part, as in [32, (3.59)]. In the present setting, since $G_1 \equiv 0$ in this theorem and also in Section 3.6 of [32], one can apply it by taking $K = 1$ so that $G = G_2$. Note that $K \geqslant 1$ in Theorem 3.9 of [32] was just a parameter which can be different from $K(N)$ in [32, (1.1)], and a uniform estimate in K was provided. Note also that, by the same reason, we may take $K = 1$ in Theorem 3.11 of [32].

Term 3: $Q^{En}_N(\lambda,F)$. The error $Q^{En}_N(\lambda,F)$ in the entropy calculation, especially in the time derivative of $h_N(t)$, has a bound

$$|Q^{En}_N(\lambda,F)| \leqslant CN^{-1}(1 + r(F)^{d+2}\|F\|_\infty)^3, \tag{7.21}$$

if the condition $N^{-1}r(F)^d\|F\|_\infty \leqslant 1$ is satisfied; see Lemma 3.1 in [32].

Term 4: $Q^{\Omega_2}_{N,L}(\lambda,F)$. The error $Q^{\Omega_2}_{N,L}(\lambda,F)$ for the gradient term is bounded as

$$\begin{aligned}|Q^{\Omega_2}_{N,L}(\lambda,F)| \leqslant & C\big(N^{-1}L^{(d+2)/2} + L^{-d}(1+r(F)^d)\big) \\ & \times (1 + r(F)^{2d}\|F\|^2_\infty + r(F)^{d+1}\|F\|_\infty) + CN^{-1}.\end{aligned} \tag{7.22}$$

This estimate is shown in Lemma 3.8 of [32] relying on Lemma 3.4 of [32]. To apply the latter, since K in the estimate covers both Glauber and Kawasaki effects, we need to take $K = 1$ rather than $K = 0$ by a similar reason to Term 2 above.

Step 2: choice of parameters. Now, we choose $1 \ll n \ll L \ll N$ and $F \equiv F_N = \Phi_{n(N)}$ as stated at the beginning of this section, and insert the estimate (7.15). More precisely, we set n, L, β as mesoscopical scales

$$n := N^{s_1} < L := N^{s_2} < N, \quad 0 < s_1 < s_2 < 1, \qquad \beta := n^{s_3}, \quad s_3 > 0, \tag{7.23}$$

with s_1, s_2, s_3 to be determined. Then, from (7.16) and several estimates stated above, we obtain for some $\mu > 0$

$$\begin{aligned}h_N(t) \leqslant C\Big(& h_N(0) + n^{-s_3} + n^{\mu}N^{-1}L^{2d+4} + n^{s_3}L^{-\alpha_1} \\ & + n^{\mu}L^{-1} + L^{-d}\log L + n^{s_3}n^{-\alpha_2}\Big).\end{aligned}$$

Since $h_N(0) \leqslant CN^{-\alpha}$, choosing $s_3 \in (0,\alpha_2)$, $s_2 = 1/(2d+5)$ and then $s_1 \in (0,s_2)$ small enough, one can derive

$$h_N(t) \leqslant CN^{-\kappa},$$

for some $C, \kappa > 0$.

Step 3: entropy inequality. To show the estimate (1.19) in Theorem 1.3, consider the event

$$\mathcal{A} = \left\{\eta \in \mathcal{X}_N : \left|\int_{\mathbb{T}^d}\phi(v)\rho^N(t,\mathrm{d}v) - \int_{\mathbb{T}^d}\phi(v)\rho(t,\mathrm{d}v)\right| > \varepsilon\right\}.$$

Then, by the large deviation estimate with respect to $P^{\psi_t} = \psi_t^N \, d\nu_{1/2}^N$, we have

$$P^{\psi_t}(\mathcal{A}) \leqslant e^{-C(\varepsilon)N^d} \tag{7.24}$$

for some $C(\varepsilon) > 0$. Indeed, (7.24) is shown since the contribution of F_N is negligible due to the factor $1/N$ in (7.14), the bound on $\|\Phi_n\|_\infty$ and the choice of $n = n(N)$; see (2.6) and Corollary 3.10 in [32]. Applying the entropy inequality for $P^{f_t} = f_t^N \, d\nu_{1/2}^N$, we obtain

$$P^{f_t}(\mathcal{A}) \leqslant \frac{\log 2 + N^d h_N(f_t^N|\psi_t^N)}{\log\{1 + 1/P^{\psi_t}(\mathcal{A})\}} \leqslant \frac{\log 2 + CN^{d-\kappa}}{\log\{1 + e^{C(\varepsilon)N^d}\}} \leqslant C'N^{-\kappa}.$$

This completes the proof of Theorem 1.3. □

Our method and estimates apply also for non-gradient Glauber–Kawasaki dynamics. In particular, one can give the upper bound for the strength $K(N)$ of the Glauber part for which one can prove the hydrodynamic limit. This is discussed in [32].

8. Extension to disordered lattice gas

We have worked with the non-gradient model where the jump rate is non-constant but a deterministic function on the configuration space. One may wonder whether the method is effective when the external randomness comes into the jump rate function. These models are known as *the disordered lattice gas* or *the exclusion process in random/inhomogeneous environment*. In the literature, [70, 30] considered the lattice gas with disorder on site; [42, 49, 50] studied the cases where the jump rate dependents on the random environment and combines the homogenization theory to obtain the diffusion matrix; [28, 29] further relaxed the underlying graph to the supercritical percolation and other stationary random graphs. It is not immediate to cover the quantitative homogenization results in all these models, because there are various ways to pose the disorder and sometimes the jump rate also degenerates. Here we just give one typical example to illustrate that our proof still works in the presence of external disorder. Our model can be seen as a lattice gas with disorder on bonds, where randomness is introduced without breaking the spatial homogeneous property, the uniform ellipticity and the product Bernoulli measure is still an invariant measure. The argument will give a quantitative convergence rate (with a random fluctuation) of finite-volume conductivity.

8.1. Model and hypothesis.

The notations in this new model is almost the same of the original one. However, instead of the genrator (1.6) defined by a collection of functions $\{c_b(\eta)\}_{b\in(\mathbb{Z}^d)^*}$, we now consider a collection of random functions:

$$\begin{aligned} c:\ \Omega &\to \mathcal{F}_0^{(\mathbb{Z}^d)^*}, \\ \omega &\mapsto \{c_b^\omega(\cdot)\}_{b\in(\mathbb{Z}^d)^*}, \end{aligned}$$

on some probability space $(\Omega, \mathscr{G}, \mathbf{P})$ satisfying the following hypothesis.

Hypothesis 8.1. The following conditions are supposed for $c : \Omega \to \mathcal{F}_0^{(\mathbb{Z}^d)^*}$.

(1) Non-degenerate and local: there exists a positive integer $\mathbf{r}$ and a positive number $\lambda > 1$ such that for every ω, the function $c_{x,y}^\omega(\eta)$ depends only on $\{\eta_z : |z - x| \leqslant \mathbf{r}\}$, and is bounded on two sides $1 \leqslant c_{x,y}^\omega(\eta) \leqslant \lambda$.

(2) Detailed balance under Bernoulli measures: for every ω, the function $c_{x,y}^\omega(\eta)$ does not depend on $\{\eta_x, \eta_y\}$.

(3) Spatially homogeneous: the joint distribution of $(c^{\omega}_{x,x+e})_{x\in\mathbb{Z}^d,e\in U}$ is the same as that of $(\tau_x c^{\omega}_{0,e})_{x\in\mathbb{Z}^d,e\in U}$, where both of them are viewed as a family of random functions taking value in $\mathcal{F}_0$.
(4) Unit range dependence: for any two edge sets $E,F\subseteq(\mathbb{Z}^d)^*$ such that if there is no adjacent pair $b\in E$, $b'\in F$ sharing a common vertex, then the functions $\{c^{\cdot}_b\}_{b\in E}$ and $\{c^{\cdot}_b\}_{b\in F}$ are independent.

We specially mention that given each sample $\omega\in\Omega$, the functions $\{c^{\omega}_b\}_{b\in(\mathbb{Z}^d)^*}$ satisfy the same properties as that in Hypothesis 1.1 except the spatial homogeneous property, which is replaced by the equality in distribution. Meanwhile, the detailed balance condition ensures that the product Bernoulli measure $\mathrm{Ber}(\rho)^{\otimes\mathbb{Z}^d}$ is still an invariant measure for the Kawasaki dynamics of jump rate $\{c^{\omega}_b\}_{b\in(\mathbb{Z}^d)^*}$ for every $\rho\in[0,1]$.

8.2. **Quantitative stochastic homogenization.** We use the renormalization approach to establish the quantitative homogenization result. Given $\omega\in\Omega$, we define the *quenched* subadditive quantities $\overline{\nu}(\omega,\rho,\Lambda,p)$ and $\overline{\nu}_*(\omega,\rho,\Lambda,q)$ as in (4.1)

$$
\begin{aligned}
\overline{\nu}(\omega,\rho,\Lambda,p) &:= \inf_{v\in\ell_{p,\Lambda^+}+\mathcal{F}_0(\Lambda^-)}\left\{\frac{1}{2\chi(\rho)|\Lambda|}\sum_{b\in\overline{\Lambda^*}}\left\langle\frac{1}{2}c^{\omega}_b(\pi_b v)^2\right\rangle_\rho\right\},\\
\overline{\nu}_*(\omega,\rho,\Lambda,q) &:= \sup_{v\in\mathcal{F}_0}\left\{\frac{1}{2\chi(\rho)|\Lambda|}\sum_{b\in\overline{\Lambda^*}}\left\langle(\pi_b\ell_q)(\pi_b v)-\frac{1}{2}c^{\omega}_b(\pi_b v)^2\right\rangle_\rho\right\}.
\end{aligned}
\tag{8.1}
$$

Then for each ω, (1)-(4) in Proposition 4.1 still hold except (4.6) and (4.7). This gives the definition of the quenched diffusion matrix and conductivity

$$
\begin{aligned}
\overline{\nu}(\omega,\rho,\Lambda,p) &= \frac{1}{2}p\cdot\overline{D}(\omega,\rho,\Lambda)p, \qquad \overline{\mathbf{c}}(\omega,\rho,\Lambda)=2\chi(\rho)\overline{D}(\omega,\rho,\Lambda),\\
\overline{\nu}_*(\omega,\rho,\Lambda,q) &= \frac{1}{2}q\cdot\overline{D}_*^{-1}(\omega,\rho,\Lambda)q, \qquad \overline{\mathbf{c}}_*(\omega,\rho,\Lambda)=2\chi(\rho)\overline{D}_*(\omega,\rho,\Lambda).
\end{aligned}
$$

The results (4.6) and (4.7) are missing, because now for every $\Lambda\subseteq\mathbb{Z}^d$ and $z\in\mathbb{Z}^d$ we only have $\overline{\nu}(\omega,\rho,z+\Lambda,p)\overset{(d)}{=}\overline{\nu}(\omega,\rho,\Lambda,p)$ (from (3) of Hypothesis 8.1) instead of $\overline{\nu}(\rho,z+\Lambda,p)=\overline{\nu}(\rho,\Lambda,p)$ in Proposition 4.1. Nevertheless, denote by $\mathbf{E}$ the expectation associated to $(\Omega,\mathscr{G},\mathbf{P})$, we recover

$$
\begin{aligned}
\mathbf{E}[\overline{\nu}(\cdot,\rho,\square_{m+1},p)] &\leqslant \mathbf{E}[\overline{\nu}(\cdot,\rho,\square_m,p)],\\
\mathbf{E}[\overline{\nu}_*(\cdot,\rho,\square_{m+1},q)] &\leqslant \mathbf{E}[\overline{\nu}_*(\cdot,\rho,\square_m,q)].
\end{aligned}
\tag{8.2}
$$

The monotone property allows us to define the limit

$$
\overline{D}(\rho):=\lim_{m\to\infty}\mathbf{E}[\overline{D}(\cdot,\rho,\square_m)], \qquad \overline{\mathbf{c}}(\rho):=2\chi(\rho)\overline{D}(\rho).
$$

Our goal is to prove the quantitative convergence rate of the finite-volume quenched conductivity matrix to this limit. Because $\overline{\mathbf{c}}(\omega,\rho,\Lambda)$ depends on the random environment ω, we need to measure the random fluctuation and we use the following notation introduced in [11, Appendix A]: for a random variable X in $(\Omega,\mathscr{G},\mathbf{P})$ and exponents $s,\theta\in(0,\infty)$, we define

$$
X\leqslant\mathcal{O}_s(\theta)\iff\mathbf{E}[\exp((\theta^{-1}X_+)^s)]\leqslant 2,
\tag{8.3}
$$

where $X_+:=\max\{X,0\}$. Roughly, (8.3) tells that X is concentrated with a stretched exponential tail. More properties of $\mathcal{O}_s$ can be found in [11, Appendix A]. Using this notation, we state our result as follows.

Theorem 8.2. *Fix $t \in (0,d)$, there exists $\kappa(d,\lambda) > 0$, $C(t,d,\lambda) < \infty$ such that for any $\rho \in (0,1)$ and any $m \in \mathbb{N}_+$:*

$$|\overline{\mathbf{c}}(\rho) - \overline{\mathbf{c}}(\omega,\rho,\square_m)| + |\overline{\mathbf{c}}(\rho) - \overline{\mathbf{c}}_*(\omega,\rho,\square_m)| \leqslant \left(C3^{-m\kappa(d-t)} + \mathcal{O}_1(C3^{-mt})\right). \tag{8.4}$$

Proof. We follow [11, Theorem 2.4] to handle the fluctuation part. We combine the quenched subadditive quantity defined in (8.1) to obtain the quenched master quantity

$$J(\omega,\rho,\Lambda,p,q) := \overline{\nu}(\omega,\rho,\Lambda,p) + \overline{\nu}_*(\omega,\rho,\Lambda,q) - p\cdot q. \tag{8.5}$$

Then for each ω fixed, Lemmas 4.5 and 4.6 remain valid and especially we have

$$|\overline{\mathbf{c}}(\omega,\rho,\square_m) - \overline{\mathbf{c}}_*(\omega,\rho,\square_m)| \leqslant C(d,\lambda)\chi(\rho)\left(\sup_{|p|=1} J(\omega,\rho,\square_m,p,\overline{D}(\rho)p)^{\frac{1}{2}}\right).$$

One can verify that $J(\omega,\rho,\Lambda,p,q)$ is a subadditive quantity. The key step of the proof relies on the observation in [11, Lemma A.7], which breaks the control of subadditive quantities with mixing condition into its mean and the random fluctuation. It also applies to the mapping $\Lambda \mapsto J(\omega,\rho,\Lambda,p,q)$ in our setting: for every $p \in B_1$, there exists a constant C independent of ρ such that

$$J(\omega,\rho,\square_m,p,\overline{D}(\rho)p) \leqslant 2\mathbf{E}[J(\cdot,\rho,\square_n,p,\overline{D}(\rho)p)] + \mathcal{O}_1(C\lambda 3^{-(m-n)d}). \tag{8.6}$$

Here we have the freedom to choose the parameter $n \in (0,m) \cap \mathbb{N}_+$, which will determine t in (8.4) by setting $t := d\left(1 - \frac{n}{m}\right) \in (0,d)$.

Therefore, it suffices to study of the decay of the mean part $\mathbf{E}[J(\cdot,\rho,\square_n,p,\overline{D}(\rho)p)]$ and the proof is quite similar to what we have done in Section 5, where the main steps are Lemmas 5.2 and 5.4, and Proposition 5.3. They can be carried to the disordered setting for the following reasons.

(1) Lemma 5.2 depends on the modified Caccioppoli inequality (2.10), which is valid for each $\omega \in \Omega$ because $\{c_b^\omega\}_{b\in(\mathbb{Z}^d)^*}$ satisfies the uniform ellipticy by (1) of Hypothesis 8.1, and the underlying invariant measure is still Bernoulli measure by (2) of Hypothesis 8.1.

(2) The variance decay in Lemma 5.4 uses the spatial independence, which is ensured by (3) and (4) of Hypothesis 8.1. More precisely, let $v(\omega,\rho,\square_n,p,q)$ be the optimiser of $J(\omega,\rho,\square_n,p,q)$ like (4.16), we aim to estimate

$$\begin{aligned}&\frac{1}{2\chi(\rho)}\mathbf{E}\left[\left\langle\left|\frac{1}{|\square_n|}\sum_{x\in\square_n}\nabla_x\left(v(\omega,\rho,\square_n,p,q) - \ell_{D_n^{-1}q-p}\right)\right|^2\right\rangle_\rho\right]\\ &\leqslant C3^{-\beta n} + C\sum_{k=0}^{n-1}3^{-\beta(n-k)}\tau_k,\end{aligned}$$

where the quantity $D_n(\rho)$ is defined as

$$D_n(\rho) := \mathbf{E}[\overline{D}_*(\cdot,\rho,\square_n)^{-1}]^{-1},$$

and τ_n is defined as

$$\tau_n = \sup_{p,q\in B_1}\mathbf{E}[J(\cdot,\rho,\square_n,p,q) - J(\cdot,\rho,\square_{n+1},p,q)].$$

We write v_n as a shorthand for $v(\omega, \rho, \square_n, p, q)$ and $v_{n-1,z}$ as a shorthand of $v(\omega, \rho, z + \square_{n-1}, p, q)$, then we have the decomposition

$$\begin{aligned}
&\mathbf{E}\left[\left\langle \frac{1}{2\chi(\rho)} \left| \frac{1}{|\square_n|} \sum_{x \in \square_n} \nabla_x \left(v_n - \ell_{D_n^{-1}q - p}\right)\right|^2 \right\rangle_\rho\right]^{\frac{1}{2}} \\
&\leqslant \mathbf{E}\left[\left\langle \frac{1}{2\chi(\rho)} \left| \frac{1}{|\square_n|} \sum_{z \in \mathcal{Z}_{n,n-1}} \sum_{x \in z + \square_{n-1}} \nabla_x \left(v_{n-1,z} - \ell_{D_{n-1}^{-1}q - p}\right)\right|^2 \right\rangle_\rho\right]^{\frac{1}{2}} \\
&\quad + \mathbf{E}\left[\left\langle \frac{1}{2\chi(\rho)} \left| \frac{1}{|\square_n|} \sum_{z \in \mathcal{Z}_{n,n-1}} \sum_{x \in z + \square_{n-1}} \nabla_x \left(v_{n-1,z} - v_n\right)\right|^2 \right\rangle_\rho\right]^{\frac{1}{2}} \\
&\quad + \lambda |D_n^{-1} q - D_{n-1}^{-1} q|.
\end{aligned}$$

The key is the cancellation in the first term under $\mathbf{E}[\langle \cdots \rangle_\rho]$. After expanding the sum, we not only use the finite-range dependence of $\eta \mapsto c_b^\omega(\eta)$ to obtain the independence over $\mathbb{P}_\rho$, but we also use the unit-range dependence of $\mathbf{P}$.

(3) The L^2-flatness estimate in Proposition 5.3 should also be carried under the expectation $\mathbf{E}$. It relies on the weighted multiscale Poincaré inequality (3.15), which does not involve the jump rate. Then we further develop it, and the variance decay in Lemma 5.4 applies to the typical case (5.14), which has been discussed as above. For the atypical case (5.18), the L^∞ estimate also applies since it only requires the log-Sobolev inequality (see the proof in Appendix B), which remains valid thanks to the uniform ellipticity of the jump rate by (1) of Hypothesis 8.1 .

□

One can deduce further results built on Theorem 8.2, and we leave them to the future work.

Appendix A. Subadditivity and Whitney inequality

A lot of results on quantitative homogenization are stated for the triadic cubes, but they actually hold for the general domain with reasonable boundary regularity, and such generalization only relies on the subadditivity and a nice decomposition. We state this observation under $\mathbb{R}^d$ setting, and one can adapt it easily in lattice models. In the following statement, a triadic cube Q in $\mathbb{R}^d$ is an open set of type $z + (-\frac{3^m}{2}, \frac{3^m}{2})^d$, where $z \in \mathbb{R}^d$ is called its *center* and 3^m with $m \in \mathbb{Z}$ is called its *size* and is denoted by $\operatorname{size}(Q)$. Especially, we denote by $\square_m = (-\frac{3^m}{2}, \frac{3^m}{2})^d$ as the cube centered at 0 and of size 3^m in this section. We also denote by $|U|$ the $\mathbb{R}^d$-Lebesgue measure for Borel set U, and by $\sigma(\mathcal{C})$ the area of the $(d-1)$-dimensional surface $\mathcal{C}$.

Lemma A.1. *Let the quantity ν be defined on the bounded open sets of $\mathbb{R}^d$ with Lipschitz boundary and satisfy the following properties.*

(1) *(Spatial homogeneous) For any $z \in \mathbb{R}^d$ and open set $U \subseteq \mathbb{R}^d$, $\nu(z + U) = \nu(U)$.*
(2) *(Subadditivity) Given disjoint open sets $\{U_i\}_{1 \leqslant i \leqslant n}$ such that they partition U in the sense that $U_1, \cdots, U_n \subseteq U$ and $|U \setminus (\bigcup_{i=1}^n U_i)| = 0$, then we have*

$$\nu(U) \leqslant \sum_{i=1}^n \frac{|U_i|}{|U|} \nu(U_i). \tag{A.1}$$

(3) *(Decay for triadic cubes) There exist two finite positive constants C, α, such that for any triadic cube $\square_m$, we have*

$$\nu(\square_m) \leqslant C(3^{-\alpha m} \wedge 1). \tag{A.2}$$

Then there exists a positive constant C', such that for any open set U circumscribed by piecewise C^1 surfaces, we have

$$\nu(U) \leqslant C' \left(\frac{\sigma(\partial U) \operatorname{diam}(U)^{(1-\alpha)\vee 0}}{|U|} \wedge 1 \right). \tag{A.3}$$

Remark A.2. As an application, for the cube $\Lambda_L = (-\frac{L}{2}, \frac{L}{2})^d$ with $L \geqslant 1$, we have $\nu(\Lambda_L) \leqslant C' L^{-(\alpha\wedge 1)}$ and it gives a polynomial decay in function of the diameter.

Proof. The case for $\operatorname{diam}(U) \leqslant 1$ is trivial and we focus on the case that U of large diameter. The proof relies on the standard Whitney decomposition, which will give a family of open triadic cubes $\{Q_j\}_{j\geqslant 0}$ of nice properties; see [44, Appendix J] for its construction and proof. In our proof, the useful properties from such decomposition are

- $\{Q_j\}_{j\geqslant 0}$ are disjoint and they partition U in the sense of (2) in Lemma A.1;
- $\sqrt{d}\operatorname{size}(Q_j) \leqslant \operatorname{dist}(Q_j, \partial U) \leqslant 4\sqrt{d}\operatorname{size}(Q_j)$.

Apply the subadditivity to $\{Q_j\}_{j\geqslant 0}$ and by passing to the limit, we have

$$\nu(U) \leqslant \sum_{i=1}^{\infty} \frac{|Q_i|}{|U|} \nu(Q_i). \tag{A.4}$$

Let m be the positive integer such that

$$3^{m-1} \leqslant \operatorname{diam}(U) \leqslant 3^m, \tag{A.5}$$

then we classify the cubes by their sizes

$$I_k := \{i \in \mathbb{N}_+ : \operatorname{size}(Q_i) = 3^k\}.$$

Using the second properties listed above about the decomposition, all the cubes in I_k stay in distance $4\sqrt{d}3^k$ from the boundary, then we have

$$\bigcup_{i\in I_k} Q_i \subseteq \{x \in U : \operatorname{dist}(x, \partial U) \leqslant 5\sqrt{d}3^k\}. \tag{A.6}$$

Since U admits piecewise C^1 boundary, we denote by $\partial U = \bigcup_{j=1}^m \mathcal{C}_j$, where different pieces $\mathcal{C}_j$ only have intersection of null set. Then we have the following volume estimate

$$\begin{aligned}\sum_{i\in I_k} |Q_i| \leqslant \left|\{x \in U : \operatorname{dist}(x, \partial U) \leqslant 5\sqrt{d}3^k\}\right| &\leqslant \sum_{j=1}^m \left|\{x \in U : \operatorname{dist}(x, \partial \mathcal{C}_j) \leqslant 5\sqrt{d}3^k\}\right| \\ &\leqslant 10\sqrt{d}3^k \sum_{j=1}^m \sigma(\mathcal{C}_j) = 10\sqrt{d}3^k \sigma(\partial U).\end{aligned} \tag{A.7}$$

Then we put these estimates back to the cubes from Whitney decomposition (A.4) to obtain

$$\begin{aligned}\nu(U) &\leqslant \sum_{k=-\infty}^{0} \frac{1}{|U|}\left(\sum_{i\in I_k}|Q_i|\right)\nu(\square_k) + \sum_{k=1}^{m} \frac{1}{|U|}\left(\sum_{i\in I_k}|Q_i|\right)\nu(\square_k) \\ &\leqslant \frac{C\sigma(\partial U)}{|U|} \sum_{k=-\infty}^{0} 3^k + \frac{C\sigma(\partial U)}{|U|} \sum_{k=1}^{m} 3^{(1-\alpha)k}\end{aligned}$$

$$\leqslant \frac{C\sigma(\partial U)}{|U|} 3^{m(1-\alpha)\vee 0}.$$

Here apply (A.2) and (A.7) from the first line to the second line, then we use (A.5) to conclude the proof. ☐

Appendix B. Proof of L^∞ norm using mixing time

In this part, we prove Lemma 4.14. Such estimate should be generally valid for Markov chain, and we state the case of Kawasaki dynamics.

Lemma B.1. *There exists a constant $C(\lambda, d) < \infty$, such that for any connected domain $\Lambda \subseteq \mathbb{Z}^d$ of diameter L and any two functions $u, f : \mathcal{X} \to \mathbb{R}$ satisfying*

$$\mathcal{L}_\Lambda u = f,$$

then for any $N \in \mathbb{N}_+$ we have

$$\left\| u - \langle u \rangle_{\Lambda,N} \right\|_\infty \leqslant CL^2 \log L \, \|f\|_\infty . \tag{B.1}$$

Proof. Let $P_t(\cdot,\cdot) : \mathcal{X} \times \mathcal{X} \to \mathbb{R}_+$ be the transition probability of the continuous-time Kawasaki dynamics generated by the generator $\mathcal{L}_\Lambda$, which has $\mu_{\Lambda,N} = \mathbb{P}_{\Lambda,N}$ as its stationary measure for any $N \in \mathbb{N}_+$. Then we use the Duhamel's formula

$$u(\eta) - \langle u \rangle_{\Lambda,N} = \int_0^\infty (P_t f)(\eta) - \langle f \rangle_{\Lambda,N} \, \mathrm{d}t, \tag{B.2}$$

which is well-defined thanks to the spectral gap and exponential decay of the mapping $t \mapsto \mathbb{V}\mathrm{ar}_{\Lambda,N}[P_t f]$.

Then we estimate the L^∞ norm. We notice that

$$\begin{aligned} \left\| (P_t f)(\eta) - \langle f \rangle_{\Lambda,N} \right\|_\infty &= \sup_{\eta \in \mathcal{X}} \left(\sum_{\eta' \in \mathcal{X}} P_t(\eta, \eta') f(\eta') - \sum_{\eta' \in \mathcal{X}} \mu_{\Lambda,N}(\eta') f(\eta') \right) \\ &\leqslant 2 \sup_{\eta \in \mathcal{X}} \| P_t(\eta, \cdot) - \mu_{\Lambda,N} \|_{TV} \, \|f\|_\infty . \end{aligned} \tag{B.3}$$

Here we denote by $\|\cdot\|_{TV}$ the total variation distance and make use of its definition [56, (4.7)]. By convention in [56, (4.22),(4.30),(4.31)], we also define

$$d(t) := \sup_{\eta \in \mathcal{X}} \| P_t(\eta, \cdot) - \mu_{\Lambda,N} \|_{TV} ,$$

$$t_{\mathrm{mix}} := \inf \left\{ t \in \mathbb{R}_+ : d(t) < \frac{1}{4} \right\} .$$

Then $t \mapsto d(t)$ is decreasing and satisfies (see [56, (4.33)])

$$\forall n \in \mathbb{N}, \qquad d(n t_{\mathrm{mix}}) \leqslant 2^{-n}.$$

We put this observation and (B.3) back to (B.2)

$$\begin{aligned} \left\| (P_t f)(\eta) - \langle f \rangle_{\Lambda,N} \right\|_\infty &\leqslant \sum_{n=0}^\infty \int_{n t_{\mathrm{mix}}}^{(n+1) t_{\mathrm{mix}}} \left\| (P_t f)(\eta) - \langle f \rangle_{\Lambda,N} \right\|_\infty \mathrm{d}t \\ &\leqslant \sum_{n=0}^\infty 2 d(n t_{\mathrm{mix}}) \, \|f\|_\infty \, t_{\mathrm{mix}} \\ &\leqslant 2 \sum_{n=0}^\infty 2^{-n} \, \|f\|_\infty \, t_{\mathrm{mix}} \\ &= 2 \, \|f\|_\infty \, t_{\mathrm{mix}}. \end{aligned}$$

One classical method to obtain the mixing time of Markov process is the log-Sobolev inequality; see [25, (1.8)]. In our case, we use the log-Sobolev inequality on general

Kawasaki dynamics developed in [58, Theorem 3]; see also [61, (2)] and discussion there. This gives us

$$t_{\mathrm{mix}} \leqslant C(d,\lambda)L^2 \log\log\binom{L^d}{N} \leqslant \widetilde{C}(d,\lambda)L^2\log L, \tag{B.4}$$

which yields (B.1). □

Remark B.2. The steps before (B.4) is standard for all the reversible Markov processes. We apply the log-Sobolev inequality for the mixing time, because the jump rate in this paper depends on the local configuration. We remark some other recent progresses [68, 53, 61] on the mixing time of the exclusion processes (i.e. the constant-speed Kawasaki dynamics). Their generalization on the non-gradient models is an interesting and challenging question.

Proof of Lemma 4.4. We take $u(\Lambda, q)$ for example, which is the solution of

$$\mathcal{L}_\Lambda u(\Lambda, q) = \sum_{b\in\overline{\Lambda_L^*}} \pi_b \ell_q.$$

Therefore, we apply (B.1) with $\left\|\sum_{b\in\overline{\Lambda_L^*}} \pi_b\ell_q\right\|_\infty \leqslant C(\lambda,d)L^d$. The case for $v(\rho,\Lambda,\xi)$ can be done similarly. □

Remark B.3. For $2 \leqslant p < \infty$, we have $\|u - \langle u\rangle_{\Lambda,N}\|_p \leqslant CL^2\|f\|_p$ without $\log L$. Indeed, we may apply Riesz–Thorin interpolation theorem for two inequalities $\|P_t f - \langle f\rangle_{\Lambda,N}\|_\infty \leqslant 2\|f\|_\infty$ and $\|P_t f - \langle f\rangle_{\Lambda,N}\|_2 \leqslant e^{-cL^{-2}t}\|f\|_2$ for some $c > 0$, which follows from the spectral gap estimate for $\mathcal{L}_\Lambda$.

Acknowledgements

The research of CG is supported in part by the National Key R&D Program of China (No. 2023YFA1010400) and NSFC (No. 12301166). The research of TF is supported in part by International Scientists Project of BJNSF (No. IS23007). Part of this project was developed when CG and HW attended the workshop “Probability and Statistical Physics” at TSIMF. CG thanks Jean-Christophe Mourrat to mention this problem during his thesis. We would like to thank Yuval Peres and Shangjie Yang for helpful discussions, and thank Scott Armstrong to share the recent progress in high contrast homogenization.

References

[1] S. Armstrong, A. Bordas, and J.-C. Mourrat. Quantitative stochastic homogenization and regularity theory of parabolic equations. *Analysis & PDE*, 11(8):1945–2014, 2018.
[2] S. Armstrong, A. Bou-Rabee, and T. Kuusi. Superdiffusive central limit theorem for a brownian particle in a critically-correlated incompressible random drift. *arXiv preprint arXiv:2404.01115*, 2024.
[3] S. Armstrong and P. Dario. Elliptic regularity and quantitative homogenization on percolation clusters. *Commun. Pure Appl. Math.*, 71(9):1717–1849, 2018.
[4] S. Armstrong and P. Dario. Quantitative hydrodynamic limits of the Langevin dynamics for gradient interface models. *Electron. J. Probab.*, 29:Paper No. 9, 93, 2024.
[5] S. Armstrong, S. J. Ferguson, and T. Kuusi. Higher-order linearization and regularity in nonlinear homogenization. *Arch. Ration. Mech. Anal.*, 237(2):631–741, 2020.
[6] S. Armstrong, S. J. Ferguson, and T. Kuusi. Homogenization, linearization, and large-scale regularity for nonlinear elliptic equations. *Comm. Pure Appl. Math.*, 74(2):286–365, 2021.
[7] S. Armstrong and T. Kuusi. Elliptic homogenization from qualitative to quantitative. *arXiv preprint arXiv:2210.06488*, 2022.
[8] S. Armstrong and T. Kuusi. Renormalization group and elliptic homogenization in high contrast. *preprint*, 2024.

[9] S. Armstrong, T. Kuusi, and J.-C. Mourrat. Mesoscopic higher regularity and subadditivity in elliptic homogenization. *Comm. Math. Phys.*, 347(2):315–361, 2016.
[10] S. Armstrong, T. Kuusi, and J.-C. Mourrat. The additive structure of elliptic homogenization. *Invent. Math.*, 208(3):999–1154, 2017.
[11] S. Armstrong, T. Kuusi, and J.-C. Mourrat. *Quantitative stochastic homogenization and large-scale regularity*, volume 352 of *Grundlehren der mathematischen Wissenschaften*. Springer Nature, 2019.
[12] S. Armstrong and W. Wu. C^2 regularity of the surface tension for the $\nabla\phi$ interface model. *Comm. Pure Appl. Math.*, 75(2):349–421, 2022.
[13] S. Armstrong and W. Wu. The scaling limit of the continuous solid-on-solid model. *arXiv preprint arXiv:2310.13630*, 2023.
[14] S. N. Armstrong and J.-C. Mourrat. Lipschitz regularity for elliptic equations with random coefficients. *Arch. Ration. Mech. Anal.*, 219(1):255–348, 2016.
[15] S. N. Armstrong and C. K. Smart. Quantitative stochastic homogenization of convex integral functionals. *Ann. Sci. Éc. Norm. Supér. (4)*, 49(2):423–481, 2016.
[16] C. Bernardin. Regularity of the diffusion coefficient for lattice gas reversible under Bernoulli measures. *Stochastic Process. Appl.*, 101(1):43–68, 2002.
[17] A. Bordas. *Homogénéisation stochastique quantitative.* PhD thesis, Université de Lyon, 2018.
[18] S. Boucheron, G. Lugosi, and P. Massart. *Concentration inequalities.* Oxford University Press, Oxford, 2013.
[19] C.-C. Chang. Equilibrium fluctuations of nongradient reversible particle systems. In *Nonlinear stochastic PDEs (Minneapolis, MN, 1994)*, volume 77 of *IMA Vol. Math. Appl.*, pages 41–51. Springer, New York, 1996.
[20] P. Dario. Quantitative homogenization of the disordered $\nabla\phi$ model. *Electron. J. Probab.*, 24:Paper No. 90, 99p., 2019.
[21] P. Dario. Optimal corrector estimates on percolation cluster. *Ann. Appl. Probab.*, 31(1):377–431, 2021.
[22] P. Dario. Quantitative homogenization of differential forms. *Ann. Inst. Henri Poincaré Probab. Stat.*, 57(2):1157–1202, 2021.
[23] P. Dario and C. Gu. Quantitative homogenization of the parabolic and elliptic Green's functions on percolation clusters. *Ann. Probab.*, 49(2):556–636, 2021.
[24] P. Dario and W. Wu. Massless phases for the Villain model in $d \geqslant 3$. Preprint, arXiv:2002.02946.
[25] P. Diaconis and L. Saloff-Coste. Logarithmic Sobolev inequalities for finite Markov chains. *Ann. Appl. Probab.*, 6(3):695–750, 1996.
[26] D. Dizdar, G. Menz, F. Otto, and T. Wu. The quantitative hydrodynamic limit of the Kawasaki dynamics. *arXiv preprint arXiv:1807.09850*, 2018.
[27] D. Dizdar, G. Menz, F. Otto, and T. Wu. Toward a quantitative theory of the hydrodynamic limit. *arXiv preprint arXiv:1807.09857*, 2018.
[28] A. Faggionato. Random walks and exclusion processes among random conductances on random infinite clusters: homogenization and hydrodynamic limit. *Electron. J. Probab.*, 13:no. 73, 2217–2247, 2008.
[29] A. Faggionato. Hydrodynamic limit of simple exclusion processes in symmetric random environments via duality and homogenization. *Probab. Theory Related Fields*, 184(3-4):1093–1137, 2022.
[30] A. Faggionato and F. Martinelli. Hydrodynamic limit of a disordered lattice gas. *Probab. Theory Related Fields*, 127(4):535–608, 2003.
[31] T. Funaki. Equilibrium fluctuations for lattice gas. In *Itô's stochastic calculus and probability theory*, pages 63–72. Springer, Tokyo, 1996.
[32] T. Funaki. Interface motion from Glauber-Kawasaki dynamics of non-gradient type. *preprint*, 2024.
[33] T. Funaki and H. Spohn. Motion by mean curvature from the Ginzburg-Landau $\nabla\phi$ interface model. *Comm. Math. Phys.*, 185(1):1–36, 1997.
[34] T. Funaki, K. Uchiyama, and H. T. Yau. Hydrodynamic limit for lattice gas reversible under Bernoulli measures. In *Nonlinear stochastic PDEs (Minneapolis, MN, 1994)*, volume 77 of *IMA Vol. Math. Appl.*, pages 1–40. Springer, New York, 1996.
[35] A. Giunti, C. Gu, and J.-C. Mourrat. Quantitative homogenization of interacting particle systems. *Ann. Probab.*, 50(5):1885–1946, 2022.
[36] A. Giunti, C. Gu, J.-C. Mourrat, and M. Nitzschner. Smoothness of the diffusion coefficients for particle systems in continuous space. *Commun. Contemp. Math.*, 25(03):2250027, 2023.

[37] A. Gloria, S. Neukamm, and F. Otto. Quantification of ergodicity in stochastic homogenization: optimal bounds via spectral gap on Glauber dynamics. *Invent. Math.*, 199(2):455–515, 2015.
[38] A. Gloria, S. Neukamm, and F. Otto. A regularity theory for random elliptic operators. *Milan J. Math.*, 88(1):99–170, 2020.
[39] A. Gloria and F. Otto. An optimal variance estimate in stochastic homogenization of discrete elliptic equations. *Ann. Probab.*, 39(3):779–856, 2011.
[40] A. Gloria and F. Otto. An optimal error estimate in stochastic homogenization of discrete elliptic equations. *Ann. Appl. Probab.*, 22(1):1–28, 2012.
[41] A. Gloria and F. Otto. Quantitative results on the corrector equation in stochastic homogenization. *J. Eur. Math. Soc. (JEMS)*, 19(11):3489–3548, 2017.
[42] P. Goncalves and M. Jara. Scaling limits for gradient systems in random environment. *J. Stat. Phys.*, 131(4):691–716, 2008.
[43] P. Goncalves and M. Jara. Density fluctuations for a zero-range process on the percolation cluster. *Electron. Commun. Probab.*, 14:382–395, 2009.
[44] L. Grafakos. *Classical and modern Fourier analysis*. Pearson Education, Inc., Upper Saddle River, NJ, 2004.
[45] N. Grunewald, F. Otto, C. Villani, and M. G. Westdickenberg. A two-scale approach to logarithmic Sobolev inequalities and the hydrodynamic limit. *Ann. Inst. Henri Poincaré Probab. Stat.*, 45(2):302–351, 2009.
[46] C. Gu. An efficient algorithm for solving elliptic problems on percolation clusters. *Ann. Appl. Probab.*, 32(4):2755–2810, 2022.
[47] C. Gu, J.-C. Mourrat, and M. Nitzschner. Quantitative equilibrium fluctuations for interacting particle systems. *arXiv preprint arXiv:2401.10080*, 2024.
[48] E. Janvresse, C. Landim, J. Quastel, and H. T. Yau. Relaxation to equilibrium of conservative dynamics. I. Zero-range processes. *Ann. Probab.*, 27(1):325–360, 1999.
[49] M. Jara. Hydrodynamic limit of the exclusion process in inhomogeneous media. In *Dynamics, games and science. II*, volume 2 of *Springer Proc. Math.*, pages 449–465. Springer, Heidelberg, 2011.
[50] M. D. Jara and C. Landim. Quenched non-equilibrium central limit theorem for a tagged particle in the exclusion process with bond disorder. *Ann. Inst. Henri Poincaré Probab. Stat.*, 44(2):341–361, 2008.
[51] C. Kipnis and C. Landim. *Scaling limits of interacting particle systems*, volume 320 of *Grundlehren der Mathematischen Wissenschaften*. Springer-Verlag, Berlin, 1999.
[52] C. Kipnis, C. Landim, and S. Olla. Hydrodynamical limit for a nongradient system: the generalized symmetric exclusion process. *Comm. Pure Appl. Math.*, 47(11):1475–1545, 1994.
[53] H. Lacoin. Mixing time and cutoff for the adjacent transposition shuffle and the simple exclusion. *Ann. Probab.*, 44(2):1426–1487, 2016.
[54] C. Landim, S. Olla, and S. R. S. Varadhan. Symmetric simple exclusion process: regularity of the self-diffusion coefficient. *Comm. Math. Phys.*, 224(1):307–321, 2001.
[55] G. Last and M. Penrose. *Lectures on the Poisson process*, volume 7 of *Institute of Mathematical Statistics Textbooks*. Cambridge University Press, Cambridge, 2018.
[56] D. A. Levin and Y. Peres. *Markov chains and mixing times*. American Mathematical Society, Providence, RI, second edition, 2017. With contributions by Elizabeth L. Wilmer, With a chapter on “Coupling from the past” by James G. Propp and David B. Wilson.
[57] S. L. Lu. Equilibrium fluctuations of a one-dimensional nongradient Ginzburg-Landau model. *Ann. Probab.*, 22(3):1252–1272, 1994.
[58] S. L. Lu and H.-T. Yau. Spectral gap and logarithmic Sobolev inequality for Kawasaki and Glauber dynamics. *Comm. Math. Phys.*, 156(2):399–433, 1993.
[59] A. Menegaki. Quantitative studies and hydrodynamical limits for interacting particle systems. *PhD thesis*, 2021.
[60] A. Menegaki and C. Mouhot. A consistence-stability approach to hydrodynamic limit of interacting particle systems on lattices. In *Séminaire Laurent Schwartz—Équations aux dérivées partielles et applications. Année 2021–2022*, pages Exp. No. VII, 15. Inst. Hautes Études Sci., Bures-sur-Yvette, [2021–2022] ©2021–2022.
[61] B. Morris. The mixing time for simple exclusion. *Ann. Appl. Probab.*, 16(2):615–635, 2006.
[62] J.-C. Mourrat. Variance decay for functionals of the environment viewed by the particle. *Ann. Inst. Henri Poincaré Probab. Stat.*, 47(1):294–327, 2011.
[63] J.-C. Mourrat. An informal introduction to quantitative stochastic homogenization. *J. Math. Phys.*, 60(3):031506, 11, 2019.

[64] A. Naddaf and T. Spencer. Estimates on the variance of some homogenization problems, 1998, unpublished preprint.
[65] Y. Nagahata. Regularity of the diffusion coefficient matrix for the lattice gas with energy. *Ann. Inst. H. Poincaré Probab. Statist.*, 41(1):45–67, 2005.
[66] Y. Nagahata. Regularity of the diffusion coefficient matrix for generalized exclusion process. *Stochastic Process. Appl.*, 116(6):957–982, 2006.
[67] Y. Nagahata. Regularity of the diffusion coefficient matrix for lattice gas reversible under Gibbs measures with mixing condition. *Comm. Math. Phys.*, 273(3):637–650, 2007.
[68] R. I. Oliveira. Mixing of the symmetric exclusion processes in terms of the corresponding single-particle random walk. *Ann. Probab.*, 41(2):871–913, 2013.
[69] J. Quastel. Diffusion of color in the simple exclusion process. *Comm. Pure Appl. Math.*, 45(6):623–679, 1992.
[70] J. Quastel. Bulk diffusion in a system with site disorder. *Ann. Probab.*, 34(5):1990–2036, 2006.
[71] M. Sasada. On the Green-Kubo formula and the gradient condition on currents. *Ann. Appl. Probab.*, 28(5):2727–2739, 2018.
[72] H. Spohn. *Large scale dynamics of interacting particles*. Texts and Monographs in Physics. Springer-Verlag, Berlin, 1991.
[73] Sudeep Raja Putta. A family of efron-stein like lower bounds for variance, 2017-09-21, https://sudeepraja.github.io/EfronStein2/.
[74] S. R. S. Varadhan. Nonlinear diffusion limit for a system with nearest neighbor interactions. II. In *Asymptotic problems in probability theory: stochastic models and diffusions on fractals (Sanda/Kyoto, 1990)*, volume 283 of *Pitman Res. Notes Math. Ser.*, pages 75–128. Longman Sci. Tech., Harlow, 1993.
[75] S. R. S. Varadhan and H.-T. Yau. Diffusive limit of lattice gas with mixing conditions. *Asian J. Math.*, 1(4):623–678, 1997.

(Tadahisa Funaki) Beijing Institute of Mathematical Sciences and Applications, Beijing, China

(Chenlin Gu) Yau Mathematical Sciences Center, Tsinghua University, Beijing, China

(Han Wang) Qiuzhen College, Tsinghua University, Beijing, China