**Communications in Mathematical Physics**

# Kardar–Parisi–Zhang Equation from Long-Range Exclusion Processes

**Kevin Yang**[1,2]

1 Department of Mathematics, Stanford University, Stanford, USA.
  E-mail: kyang95@stanford.edu
2 Department of Mathematics, University of California, Berkeley, Berkeley, USA.
  E-mail: kevin_yang95@berkeley.edu

**Abstract:** We prove here that the height function associated to non-simple exclusion processes with arbitrary jump-length converges to the solution of the Kardar–Parisi–Zhang SPDE under suitable scaling and renormalization. This extends the work of Dembo and Tsai (Commun Math Phys 341(1):219–261, 2016) for arbitrary jump-length and Goncalves and Jara (Stoch Process Appl 127(12):4029–4052, 2017) for the non-stationary regime. Thus we answer a "Big Picture Question" from the AIM workshop on KPZ and also expand on the almost empty set of non-integrable and non-stationary particle systems for which weak KPZ universality is proven. We use an approximate microscopic Cole-Hopf transform like in Dembo and Tsai (2016) but we develop tools to analyze local statistics of the particle system via local equilibrium and work of Goncalves and Jara (2017). Local equilibrium is done via the one-block step in Guo et al. (Commun Math Phys 118:31, 1988) for path-space/*dynamic* statistics.

## Contents

## 1. Introduction

The Kardar–Parisi–Zhang SPDE, which we call the KPZ equation, is an SPDE whose statistics are conjectured to be universal among a large class of rough dynamic interfaces including fluctuations of burning fronts, bacterial growth colonies, and crack formation. In the physics literature this was unrigorously shown in [19] but mathematical proof is a

difficult open problem. To begin a precise discussion, we write down the KPZ equation with an effective diffusivity $\alpha \in \mathbb{R}_{>0}$/ asymmetry $\alpha' \in \mathbb{R}$:

$$\partial_T \mathfrak{h} \; = \; \frac{\alpha}{2} \partial_X^2 \mathfrak{h} \; - \; \frac{\alpha'}{2} |\partial_X \mathfrak{h}|^2 \; + \; \alpha^{\frac{1}{2}} \xi. \tag{1.1}$$

The $\xi$-term is a Gaussian space-time white noise of delta-covariance $\mathbf{E}\xi_{T,X}\xi_{S,Y} = \delta_{T-S}\delta_{X-Y}$. The KPZ equation is a *singular* SPDE because of both the roughness of the $\xi$-term and the nonlinear dependence on the slope $\partial_X \mathfrak{h}$. To motivate the universality problem of interest here, both from a mathematical and physics perspective, we give the following interpretation of "singular".

- Consider a growth model $\widetilde{\mathfrak{h}}$ given by the solution of a version of (1.1) but where the quadratic function of the slope is instead an arbitrary nonlinear function. To get the KPZ equation from the $\widetilde{\mathfrak{h}}$-equation, we may Taylor expand the nonlinear function there as a function of the slope $\partial_X \widetilde{\mathfrak{h}}$. The first two terms, which are constant and linear in $\partial_X \widetilde{\mathfrak{h}}$, can be removed by elementary transformations, and the leading-order term that remains is the quadratic in (1.1). Though this argument may hold water at a heuristic level, it ultimately gives an incorrect effective asymmetry $\alpha'$ because of the singular nature of (1.1). Dealing with this singular equation rigorously is a major goal of [15] which was later generalized into a theory of regularity structures in [16]. Rigorous application of regularity structures to make the current bullet point on universality *correct* is done in [18].
- We emphasize that although regularity structures were successfully implemented to confirm universality of the KPZ equation for SPDE growth models in [18], it depends on the noise in these SPDEs to be space-time white noise. This allows one to build solutions of said SPDEs using explicit Gaussian-based objects. Work of [17] extends universality to more general continuum noises. Extensions of universality via regularity structures to a few types of semi-discrete noise are done in [11,23]. However, this does not include noises that are relevant for non-simple exclusions. In particular, using regularity structures for universality in the context of general particle systems with a genuinely discrete flavor is open.

The previous bullet point illustrates difficulties/interesting aspects of the proposed universality of the KPZ equation. The rest of this introduction before we introduce the particle system of interest in this paper is organized as follows. First we introduce a solution to KPZ which will allow us to avoid dealing with the singular features of the KPZ Eq. (1.1). Then we record a list of results concerning convergence to KPZ for special *integrable* or *solvable* particle systems. We conclude with progress beyond solvable models and the contributions of this paper. Later in this introduction section we will discuss additional background.

To deal with the singular nature of the KPZ equation, instead of using the regularity structures in [16] we will instead employ the following Cole-Hopf transform/solution from [2]. First, for $\alpha \in \mathbb{R}_{>0}$ and $\lambda \in \mathbb{R}$ we introduce the *stochastic heat equation*:

$$\partial_T \mathfrak{Z} \; = \; \frac{\alpha}{2} \partial_X^2 \mathfrak{Z} \; - \; \lambda \alpha^{\frac{1}{2}} \mathfrak{Z} \xi. \tag{1.2}$$

- We specialize to $\lambda = \alpha' \alpha^{-1}$ and *define* the solution of the KPZ Eq. (1.1) as $\mathfrak{h} \overset{\bullet}{=} -\lambda^{-1} \log \mathfrak{Z}$. We clarify this log-transform is well-defined because (1.2) admits a continuous solution via Ito calculus for its Duhamel form, and for positive initial data it remains positive with probability 1; see [24]. We will also call the stochastic heat Eq. (1.2) the "SHE".

- We must also include an infinite renormalization/counter-term in $\mathfrak{h}$ to handle singular aspects of KPZ; see Remark 1.4.

In [2] the authors show that the height function associated to a nearest-neighbor ASEP/"asymmetric simple exclusion process" model converges to the Cole-Hopf solution of KPZ, and since then a number of similar results were obtained in [6], [8], and [9], for example. The key input for these papers is that the height function for the microscopic particle system exhibits an *algebraic duality*; the exponential of the height function satisfies an *exact* microscopic version of the stochastic heat equation. Duality of such particle systems therefore provides a direction towards establishing the KPZ equation scaling limit for microscopic height functions while only encountering stochastic analysis of the SHE and thus *without* directly addressing the singular features of the KPZ equation; recall from earlier that addressing these singular aspects for SPDEs in the context of particle systems is open. However duality is an indication of integrability or solvability, and the "set of integrable models" is very sparse and rare among all particle systems, so universality of the KPZ equation is unlikely to be solved with ideas based solely on integrability.

In a step outside the set of integrable models, the authors in [10] confirm the universality of the KPZ equation for the height functions associated to a non-simple generalization of the ASEP model in [2]. Roughly speaking the exponential of the height function solves a microscopic version of SHE as in [2] but with additional error terms. In [10] the authors assume the maximal jump-length in the non-simple ASEP-generalization is at most 3 in which case the aforementioned error terms can be addressed via standard ideas of hydrodynamic limits. In a nutshell the contribution of this paper is dealing with these errors for arbitrary jump-lengths. The approach that we will take to doing this is analyzing these errors using a homogenization strategy based on ideas from the first step in a general "Boltzmann-Gibbs principle" originally introduced in [4]; we cite [12] for a refined version of this Boltzmann-Gibbs principle. The Boltzmann-Gibbs principle is generally accessible for systems with invariant measure initial conditions while for general non-stationary initial measures this becomes a difficult problem. Actually one version of the Boltzmann-Gibbs principle was established in [5] for non-equilibrium initial measures. However, the proofs within [5] are not applicable in this article because of a few stochastic analytic difficulties for the SHE that come from the singular features of the KPZ equation. Thus, we introduce another mechanism to access parts of a non-equilibrium Boltzmann-Gibbs principle based on adapting the classical one-block scheme within [14] to path-space/dynamic statistics of the particle system. This additionally gives another *probabilistic* and *general* homogenization tool to establishing KPZ equation scaling limits for particle systems.

To reiterate, our first main goal is to extend [10] to arbitrary maximal jump-length, answering, in large part, one "Big Picture Problem" from the AIM workshop on KPZ. The second is to develop general tools for non-equilibrium particle systems.

*1.1. The model.* The particle system we focus on here is a generalization of the non-simple exclusion process studied in [10] but for arbitrary-length steps in the particle random walk; we cite [10] for this entire subsection. We give its generator below.

- Using spin-notation adopted previously in [10], given any sub-lattice $\mathfrak{I} \subseteq \mathbb{Z}$ define $\Omega_\mathfrak{I} = \{\pm 1\}^\mathfrak{I}$. Observe such an association prescribes a mapping $\mathfrak{I} \mapsto \Omega_\mathfrak{I}$ for which any containment $\mathfrak{I} \subseteq \mathfrak{I}'$ induces the canonical projection $\Omega_{\mathfrak{I}'} \to \Omega_\mathfrak{I}$ given by

$$\Pi_{\mathfrak{I}' \to \mathfrak{I}} : \Omega_{\mathfrak{I}'} \to \Omega_\mathfrak{I}, \quad (\eta_x)_{x \in \mathfrak{I}'} \mapsto (\eta_x)_{x \in \mathfrak{I}}. \tag{1.3}$$

We abuse notation and denote $\Pi_{\mathfrak{I}} \overset{\bullet}{=} \Pi_{\mathfrak{I}' \to \mathfrak{I}}$ for any $\mathfrak{I} \subseteq \mathfrak{I}' \subseteq \mathbb{Z}$. We adopt the physical interpretation that for any $\eta \in \Omega_{\mathbb{Z}}$, the value $\eta_x = 1$ indicates the presence of a particle located at $x \in \mathbb{Z}$ and that $\eta_x = -1$ indicates the absence of a particle.

- We introduce a maximal jump-length $\mathfrak{m} \in \mathbb{Z}_{>0} \cup \{+\infty\}$ and two sets of coefficients/speeds; note ellipticity $\alpha_1 > 0$ below:

$$A = \left\{ \alpha_1, \ldots, \alpha_{\mathfrak{m}} \in \mathbb{R}_{\geqslant 0} : \alpha_1 > 0, \ \sum_{k=1}^{\mathfrak{m}} \alpha_k = 1 \right\}, \quad \Gamma = \{\gamma_k \in \mathbb{R}\}_{k=1,\ldots,\mathfrak{m}} \, . \tag{1.4}$$

For any pair of sites $x, y \in \mathbb{Z}$, we denote by $\mathfrak{S}_{x,y}$ the generator for a speed-1 symmetric exclusion process on the bond $\{x, y\}$. We specify the generator $\mathfrak{S}^{N,!!}$ of our dynamic for $N \in \mathbb{Z}_{>0}$ large via its action on a generic functional $\varphi : \Omega_{\mathbb{Z}} \to \mathbb{R}$:

$$\mathfrak{S}^{N,!!}\varphi(\eta) \overset{\bullet}{=} N^2 \sum_{k=1}^{\mathfrak{m}} \alpha_k \sum_{x\in\mathbb{Z}} \left( 1 + \frac{1}{2}N^{-\frac{1}{2}}\gamma_k \mathbf{1}_{\eta_x=-1}\mathbf{1}_{\eta_{x+k}=1} \right.$$
$$\left. - \frac{1}{2}N^{-\frac{1}{2}}\gamma_k \mathbf{1}_{\eta_x=1}\mathbf{1}_{\eta_{x+k}=-1} \right) \mathfrak{S}_{x,x+k}\varphi(\eta). \tag{1.5}$$

We will denote by $\eta_T$ the particle configuration observed after time-$T$ evolution under the $\mathfrak{S}^{N,!!}$ dynamic. To be clear, every superscript ! denotes another scaling factor of $N$ for any operator for the entirety of this paper.

**Definition 1.1.** Provided any time $T \in \mathbb{R}_{\geqslant 0}$, let us define $\mathfrak{h}_{T,0}^N$ to be 2 times the net flux of particles across the bond $\{0, 1\}$, with leftward traveling particles counting as positive flux. We also define the following *height function* from [10]:

$$\mathfrak{h}_{T,x}^N \overset{\bullet}{=} \mathfrak{h}_{T,0}^N + \mathbf{1}_{x\geqslant 1}\sum_{y=1,\ldots,x}\eta_{T,y} - \mathbf{1}_{x<1}\sum_{y=x,\ldots,0}\eta_{T,y}. \tag{1.6a}$$

The height function $\mathfrak{h}^N$ becomes the solution of the KPZ equation in the large-$N$ limit under appropriate renormalization. This is the main theorem of the current paper. Looking more closely at what showing this scaling limit entails, we address first what the effective diffusivity $\alpha \in \mathbb{R}_{>0}$ and the effective asymmetry $\alpha' \in \mathbb{R}$ should be. This tells us what the limit KPZ stochastic PDE for $\mathfrak{h}^N$ should be, and this also tells us how to define the corresponding microscopic Cole-Hopf transform. Indeed, let us remark from our discussion of SHE above that defining the Cole-Hopf transform for the limit KPZ equation requires knowing the ratio $\alpha'\alpha^{-1} \in \mathbb{R}$. We will not perform the calculation in here, although to this end we reference the *KPZ scaling theory* calculation in [10] immediately after (1.8) therein. Ultimately, the effective diffusivity and asymmetry are those of the particle random walk.

**Definition 1.2.** We define $\alpha \overset{\bullet}{=} \sum_{k=1}^{\mathfrak{m}} k^2\alpha_k$ and $\alpha' \overset{\bullet}{=} \sum_{k=1}^{\mathfrak{m}} k\alpha_k\gamma_k$. We also define $\lambda \overset{\bullet}{=} \alpha'\alpha^{-1}$ following notation in [10].

**Definition 1.3.** Define the *microscopic Cole-Hopf transform* denoted by $\mathfrak{Z}^N$ by the following analog of the Cole-Hopf transform of the continuum SHE/KPZ equation. In the following, the growth speed $\mathfrak{v}_N \in \mathbb{R}$ is defined in (1.29) in [10] for arbitrary $\mathfrak{m}$:

$$\mathfrak{Z}_{T,x}^N \overset{\bullet}{=} \mathbf{Exp}\left( -\lambda N^{-\frac{1}{2}}\mathfrak{h}_{T,x}^N + N\mathfrak{v}_N T \right). \tag{1.7}$$

We will realize $\mathfrak{h}^N$ and $\mathfrak{Z}^N$ as functions on $\mathbb{R}_{\geqslant 0} \times \mathbb{R}$ via piecewise linear interpolation of their values on $\mathbb{R}_{\geqslant 0} \times \mathbb{Z}$.

*Remark 1.4.* Technically $\mathfrak{Z}^N$ is the microscopic Cole-Hopf transform of the *renormalized* height function $\mathfrak{h}^N$ with counter-term of speed $N\mathfrak{v}_N \gg 1$. The renormalization/counter-term is a microscopic indication of singular features of the KPZ equation.

*1.2. Main theorem.* The primary result of the paper is a scaling limit for $\mathfrak{Z}^N$ under a large class of initial probability measures defined as follows. We emphasize this class of initial measures is also of interest in [2] and [10] for example.

**Definition 1.5.** We say a probability measure $\mu_{0,N}$ on $\Omega_\mathbb{Z}$ is *near stationary*, if the following moment bounds hold with respect to $\mu_{0,N}$ *uniformly* in $N \in \mathbb{Z}_{>0}$ provided any $p \in \mathbb{Z}_{\geqslant 1}$ and $0 \leqslant \vartheta < \frac{1}{2}$:

$$\sup_{x\in\mathbb{Z}}\mathbf{E}|\mathfrak{Z}_{0,x}^N|^{2p} + \sup_{x,y\in\mathbb{Z}}N^{2p\vartheta}\left(|x-y|^{-2p\vartheta}\mathbf{E}|\mathfrak{Z}_{0,x}^N - \mathfrak{Z}_{0,y}^N|^{2p}\right) \lesssim_{p,\vartheta} 1. \quad (1.8)$$

Moreover, we require $\mathfrak{Z}_{0,NX}^N \to_{N\to\infty} \mathfrak{Z}_{0,X}$ locally uniformly for some continuous initial data $\mathfrak{Z}_{0,\bullet} : \mathbb{R} \to \mathbb{R}_{\geqslant 0}$.

Before we can present the main result, we must first introduce a few assumptions which we package as one. The first part of the following set of assumptions is a *finite* maximal jump-length that does not depend on $N \in \mathbb{Z}_{>0}$. We comment on what may be done to relax this constraint though this just amounts to technical adjustments throughout a few different parts of the paper. We will take the assumption of finite maximal jump-length to make this paper more readable. The second part of the following set of assumptions is more serious. It asserts that the speed of the asymmetric jumps in the particle random walk approximately satisfies a linear constraint from [10]. We actually borrow and improve on such constraint for asymmetric jumps in the particle system from [10]. In spirit of universality, the constraint should not necessarily be there. It would be interesting to remove it.

**Assumption 1.6.** The maximal jump-length $\mathfrak{m}$ is uniformly bounded and independent of $N \in \mathbb{Z}_{>0}$. Moreover, we have the a priori estimate $\sup_{k=1,\ldots,\mathfrak{m}} \alpha_k|\gamma_k - \bar{\gamma}_k| \lesssim N^{-1/2}$, where the "specialized" speeds of asymmetric jumps are given by the following formula:

$$\alpha_k \bar{\gamma}_k \doteq 2\lambda\sum_{\ell=k}^{\mathfrak{m}}\alpha_\ell\left(\ell k^{-1} - 1\right) + \lambda\alpha_k. \quad (1.9)$$

In the following statement of our main result, we will employ the Skorokhod space of $\mathbf{D}_1 \doteq \mathbf{D}([0,1], \mathbf{C}(\mathbb{K}))$ of cadlag paths valued in the Banach space of continuous functions $\mathbf{C}(\mathbb{K})$ on an arbitrary but fixed compact set $\mathbb{K} \subseteq \mathbb{R}$. We cite [3] for details.

**Theorem 1.7.** *Under any near-stationary initial data, the space-time process $\mathfrak{Z}_{T,NX}^N$ is tight in the large-$N$ limit with respect to the Skorokhod topology on $\mathbf{D}_1$. All limit points are the solution to* SHE *with parameters $\alpha, \lambda \in \mathbb{R}$ defined earlier and initial data $\mathfrak{Z}_{0,\bullet}$.*

Before we proceed, we make a few brief remarks.

- There is nothing special about time 1. It may be replaced by any fixed positive time. We also allow $\mathbb{K}$ to be any compact set.

- In [10], and even in [2] and related works, the class of near-stationary initial probability measures considered allow for the a priori estimates given in Definition 1.5 to grow in some exponential-linear fashion. We did not allow for that here. However, our methods will still hold for this larger set of initial data. The differences are almost cosmetic and depend only on diffusive tails of a discretization of the classical Gaussian heat kernel as in [10]. We focus on the above class of initial data introduced in Definition 1.5 without the exponential growth just to make this paper more readable.
- We have assumed finite maximal jump-length for the particle random walk in our paper. We can actually remove this with a strategy based on the following outline *if* we assume the sequences $\{\alpha_k\}_{k \in \mathbb{Z}_{>0}}$ and $\{\alpha_k \gamma_k\}_{k \in \mathbb{Z}_{>0}}$ have all moments as measures on $\mathbb{Z}_{>0}$. Take any $\varepsilon \in \mathbb{R}_{>0}$ as small as we want but independent of $N \in \mathbb{Z}_{>0}$. Because both $\{\alpha_k\}_{k \in \mathbb{Z}_{>0}}$ and $\{\alpha_k \gamma_k\}_{k \in \mathbb{Z}_{>0}}$ admit all moments, the speed of jump of length more than $N^\varepsilon$ is at most $\kappa_C N^{-C}$ for any $C \in \mathbb{R}_{>0}$, so roughly speaking we can forget about all jumps of length more than $N^\varepsilon$, and $\mathfrak{m} = N^\varepsilon$. We are almost in a situation of finite maximal jump-length, but $\mathfrak{m} = N^\varepsilon$ still grows in $N \in \mathbb{Z}_{>0}$ even if slowly. It turns out that the presence of $\mathfrak{m} = N^\varepsilon$ only affects estimates which are power-savings in $N \in \mathbb{Z}_{>0}$ and their dependence on $\mathfrak{m} = N^\varepsilon$ is polynomial. Thus these extra $N^{10\varepsilon}$ factors, for example, are negligible.

*1.3. Narrow-wedge initial measure.* The analysis in this article can be adjusted to treat the exclusion processes considered herein where the initial probability measure on the set of particle configurations is *not* near-stationary or anything nearby, but rather the *narrow-wedge* initial measure/configuration. This initial measure for the set of particle configurations gives rise to a microscopic Cole-Hopf transform with large-$N$ limit the Dirac point mass; for a detailed discussion, see [1] and [7].

Because of the distributional nature of Dirac point masses the adjustments that we need in order to treat exclusion processes here but with narrow-wedge initial measure are nontrivial and technical since they require adjustments of the hydrodynamic-limit-input of this paper and not just the stochastic analytic inputs as was the case for [10]. Adding these would require several detailed technical arguments. For this reason we will defer the extension to narrow-wedge initial measure to a separate paper.

*1.4. Background.* We have mentioned already the open problem of implementing the theory of regularity structures outside stochastic PDEs in [16] and [18] and to height functions of interacting particle systems. We now discuss a different approach to weak KPZ universality known as a *theory of energy solutions*. This approach was thoroughly explored in [12] by Goncalves and Jara. The approach via energy solutions is designed around a nonlinear martingale problem for the KPZ equation or its avatar in the stochastic Burgers equation. As much of previous literature on hydrodynamic limits and their fluctuations for interacting particle systems was also based on martingale problems, such nonlinear martingale problem was engineered to fit in the "same framework" to confirm the universality of (1.1) for many non-integrable interacting particle systems. In particular, it does not apply the Cole-Hopf transform and instead directly makes sense of the singular features of KPZ. The theory of energy solutions, however, depends on the model being at/*very* close to some invariant measure to make sense of these singular features of KPZ. In particular, the theory of energy solutions applies only to *stationary* interacting particle systems while much of the interest in the current work is proving a KPZ scaling limit for *non-stationary* interacting particle systems.

*1.5. Organization.* We will not be able to discuss the actual content of this paper until the end of Sect. 2 at which point we will have set up an *approximate* microscopic version of SHE for $\mathfrak{Z}^N$. We instead give the following high-level outline for now.

- In Sect. 2, we re-develop the framework in Sect. 2 of [10]. The strategy for the proof of Theorem 1.7 is also given.
- In Sect. 3, we establish local equilibrium via entropy production in infinite-volume along with equilibrium calculations.
- In Sect. 4, we introduce compactification of the $\mathfrak{Z}^N$-dynamics. This is effectively done by heat kernel estimates.
- In Sect. 5, we develop our main technical contribution which we call a *dynamic variation* of the one-block strategy.
- In Sect. 6, we establish preliminary time-regularity estimates for the microscopic Cole-Hopf transform $\mathfrak{Z}^N$.
- In Sect. 7, we use the dynamical one-block strategy and time-regularity in a multi-scale analysis to get a "key" estimate.
- In Sect. 8, we prove Theorem 1.7 for near-stationary data using the above "key" estimate.

In the "Appendix" we record auxiliary heat kernel and martingale estimates which are based on Proposition A.1/Corollary A.2 in [10] and Lemma 3.1 in [10]. We also include a list of notation in the "Appendix" for consult while reading this paper. Lastly, this paper is long consequence of trying to provide enough explanation for ideas in here and clarify each of the many big points. In particular, we try to leave no seemingly abstract estimate/construction without a word about why it is helpful/what it is doing. We also often explain proofs of various results to clarify the technical details here which adds to the length of the paper as well.

*1.6. Comments on notation.* We have an index for notation we use often in this paper in the "Appendix" section of this paper. We give here a few pieces of notation from the aforementioned "Appendix" section that are more commonly used.

- A "universal" constant is one that depends on nothing beyond possibly fixed data of the particle system, for example speeds $\{\alpha_k\}_{k=1}^{\mathfrak{m}}$. When we refer to a constant as "arbitrarily/sufficiently small but universal", we mean arbitrarily/sufficiently small depending only on a uniformly bounded number of universal constants. The same is true for "arbitrarily/sufficiently large but universal" constants except the word "small" is replaced by "large". The reader is invited to take "arbitrarily/sufficiently small but universal" constants to be $999^{-999}$ and to take "arbitrarily/sufficiently large but universal" constants to be $999^{999}$.
- Provided any $a, b \in \mathbb{R}$, we define the discretized interval $[\![a, b]\!] = [a, b] \cap \mathbb{Z}$.
- For any finite set $I$, let $\kappa_I \in \mathbb{R}$ be a constant depending *only* on $I$. We also define a *normalized* sum $\widetilde{\sum}_{i \in I} \overset{\bullet}{=} |I|^{-1} \sum_{i \in I}$.
- The "microscopic time-scale" is order $N^{-2}$ and the "macroscopic time-scale" is order 1. The "mesoscopic time-scales" are any time-scales that are between these two time-scales. Similarly the "microscopic length-scale" is order 1 and the "macroscopic length-scale" is order $N$. The "mesoscopic length-scales" are between these two length-scales.
- Script font will be used for operators. Fraktur font will be used for anything related to PDEs, the particle system, or anything related to details/analysis of the particle system (such as time-shifts, time-scales for averages, functionals, and subsets in $\mathbb{Z}$).

- Finally, starting in Sect. 6, the subscript "st" for random times $\mathfrak{t}_{st} \in \mathbb{R}_{\geqslant 0}$ stands for "stopping time".

## 2. Approximate Microscopic Stochastic Heat Equation

We recap framework developed in [10] with adjustments catered to our analysis in this paper. We then provide an outline of overcoming obstacles discussed in [10] that limit the maximal jump-length in [10]. We will employ here invariant probability measures for relevant exclusion processes from Definition 3.1. These are canonical and grand-canonical ensembles.

*2.1. Quantities of interest.* The goal of the current section is to write down an approximate microscopic SHE as a stochastic equation for the dynamics of the microscopic Cole-Hopf transform. To do so we must first identify functionals/coefficients for which we can conduct some probabilistic analysis. First we introduce some notation.

**Definition 2.1.** For any $\mathfrak{f} : \Omega_{\mathbb{Z}} \to \mathbb{R}$, define its *support* as the smallest subset $\mathfrak{I} \subseteq \mathbb{Z}$ so that $\mathfrak{f}$ depends only on $\eta_x$ for $x \in \mathfrak{I}$.

**Definition 2.2.** For $X \in \mathbb{Z}$, define $\tau_X : \Omega_{\mathbb{Z}} \to \Omega_{\mathbb{Z}}$ to shift a configuration so that $(\tau_X \eta)_Z = \eta_{Z+X}$ for all $(\eta, Z) \in \Omega_{\mathbb{Z}} \times \mathbb{Z}$.

We define three classes of functionals below. The first is a generalization of "weakly vanishing" terms in [10].

**Definition 2.3.** A functional $\mathfrak{w} : \mathbb{R}_{\geqslant 0} \times \mathbb{Z} \times \Omega_{\mathbb{Z}} \to \mathbb{R}$ is *weakly vanishing* if $|\mathfrak{w}| \lesssim N^{-\beta}$ for $\beta \in \mathbb{R}_{>0}$ universal, or if:

- For all $(T, X, \eta) \in \mathbb{R}_{\geqslant 0} \times \mathbb{Z} \times \Omega_{\mathbb{Z}}$, we have $\mathfrak{w}_{T,X}(\eta) = \mathfrak{w}_{0,0}(\tau_{-X}\eta_T)$ for a "reference functional" $\mathfrak{w}_{0,0} : \Omega_{\mathbb{Z}} \to \mathbb{R}$.
- We have $\mathbf{E}^{\mu_{0,\mathbb{Z}}} \mathfrak{w}_{0,0} = 0$, where $\mu_{0,\mathbb{Z}}$ is the product Bernoulli measure on $\Omega_{\mathbb{Z}}$ defined by $\mathbf{E}^{\mu_{0,\mathbb{Z}}} \eta_x = 0$ for all $x \in \mathbb{Z}$.
- We have the deterministic bound $|\mathfrak{w}_{0,0}| \lesssim 1$ uniformly in $N \in \mathbb{Z}_{>0}$ and $\eta \in \Omega_{\mathbb{Z}}$ with a universal implied constant.
- The support of $\mathfrak{w}_{0,0}$ is uniformly bounded, so that it is contained in an interval of length independent of $N \in \mathbb{Z}_{>0}$.

*At the level of hydrodynamic limits*, weakly vanishing terms are negligible as near-stationary initial measures imply that the global $\eta$-density is roughly 0. This is the defining property of weakly vanishing terms in [10]. The one-block and two-blocks steps in the proof of Lemma 2.5 in [10] will also apply for weakly vanishing terms defined above. We return to this in the proof for Theorem 1.7, but we note here weakly vanishing terms will not give any difficulties that were not already treated in [10].

**Definition 2.4.** A functional $\mathfrak{g} : \mathbb{R}_{\geqslant 0} \times \mathbb{Z} \times \Omega_{\mathbb{Z}} \to \mathbb{R}$ is a *pseudo-gradient* if the following conditions are satisfied:

- For all $(T, X, \eta) \in \mathbb{R}_{\geqslant 0} \times \mathbb{Z} \times \Omega_{\mathbb{Z}}$, we have $\mathfrak{g}_{T,X}(\eta) = \mathfrak{g}_{0,0}(\tau_{-X}\eta_T)$ with reference functional $\mathfrak{g}_{0,0} : \Omega_{\mathbb{Z}} \to \mathbb{R}$.
- We have the deterministic bound $|\mathfrak{g}_{0,0}| \lesssim 1$ uniformly in $N \in \mathbb{Z}_{>0}$ and $\eta \in \Omega_{\mathbb{Z}}$. The support of $\mathfrak{g}_{0,0}$ is uniformly bounded.

- We have $\mathbf{E}^{\mu_{\varrho,\mathfrak{I}}^{\mathrm{can}}}\mathfrak{g}_{0,0} = 0$ for any canonical ensemble parameter $\varrho \in \mathbb{R}$ and set $\mathfrak{I} \subseteq \mathbb{Z}$ containing the support of $\mathfrak{g}_{0,0}$.

If $\mathfrak{g}_{0,0}$ is a discrete gradient, so that $\mathfrak{g}_{0,0} = \tau_{-\mathfrak{j}}\mathfrak{f}_{0,0} - \mathfrak{f}_{0,0}$ for a local functional $\mathfrak{f}_{0,0} : \Omega_{\mathbb{Z}} \to \mathbb{R}$ satisfying the required uniform boundedness and support condition and $\mathfrak{j} \in \mathbb{Z}$ is uniformly bounded, then $\mathfrak{g}_{0,0}$ is the reference functional for a pseudo-gradient. To show this, uniform boundedness of $\mathfrak{j} \in \mathbb{Z}$, of $\mathfrak{f}_{0,0}$, and of the support of $\mathfrak{f}_{0,0}$ guarantee the second bullet point for $\mathfrak{g}_{0,0}$. To see the last bullet point for $\mathfrak{g}_{0,0}$, it suffices to see that the canonical ensembles in that third bullet point are "invariant under shifts"; on any subset containing supports of $\tau_{-\mathfrak{j}}\mathfrak{f}_{0,0}$ and $\mathfrak{f}_{0,0}$, with respect to any canonical ensemble $\tau_{-\mathfrak{j}}\mathfrak{f}_{0,0}$ and $\mathfrak{f}_{0,0}$ are equal in law.

For a non-gradient example of a pseudo-gradient we give the cubic nonlinearity in Proposition 2.3 in [10] in the case where the maximal jump-length is $\mathfrak{m} \geqslant 4$. This cubic functional evaluated at the particle system at $(0, 0) \in \mathbb{R}_{\geqslant 0} \times \mathbb{Z}$ satisfies required estimates for the reference functional in the second bullet point in the definition of pseudo-gradients that we gave above. This can be checked directly because this cubic nonlinearity is the sum of a uniformly bounded $\mathfrak{m}$-dependent number of difference of cubic monomials in spins contained in some neighborhood of length-scale at most $10\mathfrak{m}$. To justify the interesting vanishing-in-expectation requirement in the third bullet point in the definition above, we observe that canonical ensembles are invariant under swapping spins at deterministic points. Thus, if we take $\eta_1\eta_2\eta_3 - \eta_{-1}\eta_{-2}\eta_{-3}$, for example, in expectation with respect to any canonical ensemble containing $\{\pm 1, \pm 2, \pm 3\}$ we may replace $\eta_{-1}\eta_{-2}\eta_{-3} \mapsto \eta_1\eta_2\eta_3$ at the level of expectations.

**Definition 2.5.** A functional $\widetilde{\mathfrak{g}} : \mathbb{R}_{\geqslant 0} \times \mathbb{Z} \times \Omega_{\mathbb{Z}} \to \mathbb{R}$ is said to have a *pseudo-gradient factor* if it is uniformly bounded and we have a factorization of functionals $\widetilde{\mathfrak{g}} = \mathfrak{g} \cdot \mathfrak{f}$ such that the following constraints are satisfied:

- We have $\mathfrak{f}_{T,X}(\eta) = \mathfrak{f}_{0,0}(\tau_{-X}\eta_T)$. The support of $\mathfrak{f}_{0,0}$ is bounded but may be $N$-dependent. We still require $|\mathfrak{f}_{0,0}| \lesssim 1$.
- The factor $\mathfrak{g}$ is a pseudo-gradient, and the $\eta$-wise supports of $\mathfrak{g}_{0,0}$ and $\mathfrak{f}_{0,0}$ are disjoint subsets in $\mathbb{Z}$.
- The factor $\mathfrak{f}_{0,0}$ is an average of terms that are each a product of $\eta$-spins times uniformly bounded/deterministic constants.

Since the supports of the pseudo-gradient factor and the functional $\mathfrak{f}$ in the last class of functionals are disjoint subsets, it is easy to see $\widetilde{\mathfrak{g}}$-functionals above are also pseudo-gradients. However, observe that the support of the reference functional $\mathfrak{f}_{0,0}$ is allowed to grow with $N \in \mathbb{Z}_{>0}$ so the same is true of $\widetilde{\mathfrak{g}}$. Our analysis for pseudo-gradients deteriorates in the length-scale of their support so we will not be able to efficiently study these $\widetilde{\mathfrak{g}}$-terms via the same ideas. The point of introducing the previous class of functionals is to highlight the probing of a pseudo-gradient factor with uniformly bounded support we will need to do.

*2.2. Approximate SHE.* Dynamics of the microscopic Cole-Hopf transform will be driven by the following heat operators.

**Definition 2.6.** Let $\mathfrak{P}^N$ be the heat kernel with $\mathfrak{P}_{S,S,x,y}^N = \mathbf{1}_{x=y}$ solving the semi-discrete parabolic equation below where the operator $\mathscr{L}^{!!}$, which is also defined below, acts on the backwards spatial variable of the heat kernel:

$$\partial_T \mathfrak{P}_{S,T,x,y}^N = \mathscr{L}^{!!}\mathfrak{P}_{S,T,x,y}^N \tag{2.1}$$

The $\mathscr{L}^{!!}$ operator is the discrete-type Laplacian defined below in which $\widetilde{\alpha}_k = \alpha_k + \mathscr{O}(N^{-1})$ are defined in Lemma 1.2 of [10]:

$$\mathscr{L}^{!!} \doteq 2^{-1}\sum\nolimits_{k=1}^{\mathfrak{m}} \widetilde{\alpha}_k \Delta_k^{!!}. \tag{2.2}$$

We introduced $\Delta_k\varphi_x \doteq \varphi_{x+k} + \varphi_{x-k} - 2\varphi_x$ and $\Delta_k^{!!} \doteq N^2\Delta_k$ given any function $\varphi : \mathbb{Z} \to \mathbb{R}$. Additionally for $(T, x) \in \mathbb{R}_{\geqslant 0} \times \mathbb{Z}$, we define space-time and spatial heat/convolution-operators acting on space-time test functions $\varphi : \mathbb{R}_{\geqslant 0} \times \mathbb{Z} \to \mathbb{R}$:

$$\mathfrak{P}_{T,x}^N\varphi \doteq \mathfrak{P}_{T,x}^N\varphi_{S,y} \doteq \mathfrak{P}_{T,x}^N\varphi_{\bullet,\bullet} \doteq \int_0^T \sum\nolimits_{y\in\mathbb{Z}}\mathfrak{P}_{S,T,x,y}^N \cdot \varphi_{S,y}\,\mathrm{d}S \tag{2.3}$$

$$\mathfrak{P}_{T,x}^{N,\mathbf{X}}\varphi \doteq \mathfrak{P}_{T,x}^{N,\mathbf{X}}\varphi_{0,y} \doteq \mathfrak{P}_{T,x}^{N,\mathbf{X}}\varphi_{0,\bullet} \doteq \sum\nolimits_{y\in\mathbb{Z}}\mathfrak{P}_{0,T,x,y}^N \cdot \varphi_{0,y}. \tag{2.4}$$

We occasionally use convexity of the spatial heat operator, so for any norm $\|\|$ and $p \in \mathbb{R}_{\geqslant 1}$ we have $\|\mathfrak{P}_{T,x}^{N,\mathbf{X}}\varphi\|^p \leqslant \mathfrak{P}_{T,x}^N\|\varphi\|^p$.

The main result of this section is the following SDE-type equation for the microscopic Cole-Hopf transform. By SDE-type, we mean stochastic equation with discrete noise $\xi^N$. The following borrows largely from Sect. 2 of [10].

**Proposition 2.7.** *Consider $\varepsilon_{X,1} \in \mathbb{R}_{>0}$ arbitrarily small but universal and $\beta_X \doteq \frac{1}{3} + \varepsilon_{X,1}$. We have, with notation defined after,*

$$\mathrm{d}\mathfrak{Z}_{T,x}^N = \mathscr{L}^{!!}\mathfrak{Z}_{T,x}^N\mathrm{d}T + \mathfrak{Z}_{T,x}^N\mathrm{d}\xi_{T,x}^N + \Phi_{T,x}^{N,2}\mathrm{d}T + \Phi_{T,x}^{N,3}\mathrm{d}T \tag{2.5a}$$

$$\mathfrak{Z}_{T,x}^N = \mathfrak{P}_{T,x}^{N,\mathbf{X}}\mathfrak{Z}_{0,\bullet}^N + \mathfrak{P}_{T,x}^N(\mathfrak{Z}^N\mathrm{d}\xi^N) + \mathfrak{P}_{T,x}^N(\Phi^{N,2}) + \mathfrak{P}_{T,x}^N(\Phi^{N,3}). \tag{2.5b}$$

*The first $\Phi^{N,2}$-term contains the "pseudo-gradient content" in the $\mathfrak{Z}^N$-equation. We use notation to be defined after:*

$$\Phi_{T,x}^{N,2} \doteq N^{\frac{1}{2}}\mathscr{A}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{T,x}) \cdot \mathfrak{Z}_{T,x}^N + N^{\beta_X}\left(\widetilde{\sum}_{\mathfrak{l}=1,\ldots,N^{\beta_X}}\widetilde{\mathfrak{g}}_{T,x}^{\mathfrak{l}}\right)\mathfrak{Z}_{T,x}^N$$
$$+ N^{-\frac{1}{2}}\widetilde{\sum}_{\mathfrak{l}=1,\ldots,N^{\beta_X}}\nabla_{-7\mathfrak{l}\mathfrak{m}}^!(\mathfrak{b}_{T,x}^{\mathfrak{l}}\mathfrak{Z}_{T,x}^N) \tag{2.6}$$

*Meanwhile the $\Phi^{N,3}$-term contains the "weakly-vanishing content" in the $\mathfrak{Z}^N$-equation for which we also use notation defined after:*

$$\Phi_{T,x}^{N,3} \doteq \mathfrak{w}_{T,x}\mathfrak{Z}_{T,x}^N + \sum\nolimits_{k=-2\mathfrak{m}}^{2\mathfrak{m}}c_k\nabla_k^!(\mathfrak{w}_{T,x}^k\mathfrak{Z}_{T,x}^N). \tag{2.7}$$

- *The martingale integrator $\mathrm{d}\xi^N$ is defined in (2.4) of [10]. It generalizes to any maximal-length in straightforward fashion.*
- *We define a spatial-average "operator" where $\mathfrak{g}_{T,x} = \mathfrak{g}_{0,0}(\tau_{-x}\eta_T)$ is a pseudo-gradient whose support has size at most $5\mathfrak{m}$:*

$$\mathscr{A}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{T,x}) \doteq \widetilde{\sum}_{\mathfrak{l}=1,\ldots,N^{\beta_X}}\tau_{-7\mathfrak{l}\mathfrak{m}}\mathfrak{g}_{T,x}. \tag{2.8}$$

*The summands in the $\mathscr{A}^{\mathbf{X},-}(\mathfrak{g})$-term have disjoint supports as spatial shifts are with respect to multiples of $7\mathfrak{m}$. We also note the superscript "$\mathbf{X}, -$" emphasizes a spatial average in the negative spatial direction. The support of $\mathscr{A}^{\mathbf{X},-}(\mathfrak{g})$ has size $\mathscr{O}(N^{\beta_X})$.*

- *The terms* $\widetilde{\mathfrak{g}}^{\mathfrak{l}}_{T,x} = \widetilde{\mathfrak{g}}^{\mathfrak{l}}_{0,0}(\tau_{-x}\eta_T)$ *have support of length* $\mathcal{O}(N^{\beta_X})$. *Each* $\widetilde{\mathfrak{g}}^{\mathfrak{l}}$-*term admits a pseudo-gradient factor.*
- *The terms* $\mathfrak{b}^{\mathfrak{l}}$ *are uniformly bounded. The term* $\mathfrak{w}$ *is the sum of a uniformly bounded number of weakly vanishing terms.*
- *The terms* $\mathfrak{w}^k$ *for* $|k| \leqslant 2\mathfrak{m}$ *are weakly vanishing, and the deterministic coefficients* $c_k \in \mathbb{R}$ *are uniformly bounded.*
- *Provided* $k \in \mathbb{Z}$, *define the discrete gradient* $\nabla_k \varphi_x \overset{\bullet}{=} \varphi_{x+k} - \varphi_x$ *for any* $\varphi : \mathbb{Z} \to \mathbb{R}$ *and its continuum rescaling* $\nabla_k^! \overset{\bullet}{=} N\nabla_k$.

*Remark 2.8.* The multiples $5\mathfrak{m}$ and $7\mathfrak{m}$ can be replaced with any uniformly bounded multiples; it will not change our analysis of (2.5a). All we need is $\mathscr{A}^{\mathbf{X},-}$ in Proposition 2.7 is an average of pseudo-gradients with *disjoint* supports.

*Remark 2.9.* With respect to any canonical ensemble pseudo-gradients are fluctuations, so averaging them in $\mathscr{A}^{\mathbf{X},-}$ above will provide better cancellation as the length-scale of averaging increases. The error in replacing the pseudo-gradient $\mathfrak{g}$ by a spatial-average will grow in the length-scale of spatial-averaging which is where the second term in $\Phi^{N,2}$ comes from; this will come in the proof of Proposition 2.7. Lastly, cancellation at canonical ensembles will eventually give cancellation at non-equilibrium.

To prove Proposition 2.7 we need an elementary calculation relating asymmetry to "specialized" speeds in Assumption 1.6.

**Lemma 2.10.** *We have the exact identity* $\sum_{k=1}^{\mathfrak{m}} k\alpha_k \gamma_k = \sum_{k=1}^{\mathfrak{m}} k\alpha_k \bar{\gamma}_k$.

*Proof of Proposition 2.7.* The stochastic Eq. (2.5a) of SDE-type is derived using almost entirely Proposition 2.3 from [10] and the derivation of $\mathfrak{Z}^N$-dynamics in [10] that was done prior to Proposition 2.3 in [10]. We provide only extra ingredients in additional spatial averaging and reorganizing with Taylor expansion of the exponential formula for the microscopic Cole-Hopf transform similar to the proof for Proposition 2.3 in [10]. We give the additional ingredients below starting with notation.

- Define $\mathfrak{f}^{\mathfrak{m}}_{\mathfrak{k}} \overset{\bullet}{=} \tau_{\mathfrak{k}\mathfrak{m}}\mathfrak{f}_{T,x}$ for any $\mathfrak{k} \in \mathbb{Z}$ and $\mathfrak{f} : \Omega_{\mathbb{Z}} \to \mathbb{R}$ as we only shift by $\mathfrak{m}$-multiples. We also define $\nabla^{\mathfrak{m}}_{\mathfrak{k}} \overset{\bullet}{=} \nabla_{\mathfrak{k}\mathfrak{m}}$.
- We also declare that all functionals, including the microscopic Cole-Hopf transform, are evaluated at $(T, x)$.

Proposition 2.3 from [10] gives the desired result but with cubic nonlinearities in place of the $(\bar{\mathfrak{Z}}^N)^{-1}\Phi^{N,2}$-terms, up to another difference that we address at the end of this proof that concerns quadratic polynomials in spins. The cubic nonlinearities have support contained in $[\![-\mathfrak{m}, \mathfrak{m}]\!] \subseteq \mathbb{Z}$. We will replace this cubic nonlinearity with its spatial average on the length-scale $N^{\beta_X}$. The error terms, which come from Taylor expansion, are the remaining terms in $\Phi^{N,2}$ and additional weakly vanishing terms.

The cubic nonlinearities in Proposition 2.3 of [10] are pseudo-gradients as canonical ensembles are permutation-invariant. If $\mathfrak{c}$ denotes the contribution of these nonlinearities, an elementary discrete "product/Leibniz rule" gets the following identities for which we define $\widetilde{\mathfrak{c}} \overset{\bullet}{=} \mathfrak{c}^{\mathfrak{m}}_{-2}$ and for which we give a little more explanation afterwards:

$$N^{\frac{1}{2}}\mathfrak{c}\mathfrak{Z}^N = N^{\frac{1}{2}}\mathfrak{c}^{\mathfrak{m}}_{-2}\mathfrak{Z}^N + N^{\frac{1}{2}}\mathfrak{c}^{\mathfrak{m}}_{-2}\left(\nabla^{\mathfrak{m}}_{-2}\mathfrak{Z}^N\right) - N^{\frac{1}{2}}\nabla^{\mathfrak{m}}_{-2}(\mathfrak{c}\mathfrak{Z}^N) \tag{2.9}$$

$$= N^{\frac{1}{2}}\widetilde{\mathfrak{c}}^{\mathfrak{m}}_{-7\mathfrak{l}}\mathfrak{Z}^N + N^{\frac{1}{2}}\widetilde{\mathfrak{c}}^{\mathfrak{m}}_{-7\mathfrak{l}}\left(\nabla^{\mathfrak{m}}_{-7\mathfrak{l}}\mathfrak{Z}^N\right) - N^{\frac{1}{2}}\nabla^{\mathfrak{m}}_{-7\mathfrak{l}}\left(\widetilde{\mathfrak{c}}\mathfrak{Z}^N\right)$$

$$+ N^{\frac{1}{2}}\mathfrak{c}^{\mathfrak{m}}_{-2}\left(\nabla^{\mathfrak{m}}_{-2}\mathfrak{Z}^N\right) - N^{\frac{1}{2}}\nabla^{\mathfrak{m}}_{-2}(\mathfrak{c}\mathfrak{Z}^N). \tag{2.10}$$

The second line (2.10) follows from applying the same discrete product/Leibniz rule used to obtain (2.9) but applied to the first term within the RHS of (2.9) and with $-2$ replaced by $-7\mathfrak{l}$. We will now match each term in (2.5a) to either a weakly vanishing term contributing to $\Phi^{N,3}$ or one of the terms in $\Phi^{N,2}$ at least after averaging (2.10) over $\mathfrak{l} \in [\![1, N^{\beta_X}]\!]$.

- Up to uniformly bounded error terms that are of order $N^{-1/2}$, the second term within the RHS of (2.9) is the product of linear polynomials in the spins and the cubic polynomial $\mathfrak{c}$, all supported in some neighborhood with size at most $5\mathfrak{m}$. We emphasize there is no $N^{1/2}$ in this product. This product is also without any constant term because multiplying the cubic polynomial by a linear polynomial cannot cancel spin-factors to obtain a constant. As it is a polynomial in spins supported in the same set of size $5\mathfrak{m}$, it is weakly vanishing as any polynomial of spins without constant term vanishes in expectation with respect to $\mu_{0,\mathbb{Z}}$.
- The final term from the RHS of (2.9) has the form of a gradient term in $\Phi^{N,3}$ for $\mathfrak{k} = -2\mathfrak{m}$ since it is an *unscaled* gradient acting on the product of a uniformly bounded functional and $\mathfrak{Z}^N$ which is then multiplied by the $N^{1/2}$-factor.
- We move to the first three terms in (2.10). Defining $\mathfrak{b}^{\mathfrak{l}} \overset{\bullet}{=} \widetilde{\mathfrak{c}}$ turns the third term in (2.10) into a gradient-term in $\Phi^{N,2}$.
- As $\mathfrak{c}$ is a pseudo-gradient so is $\widetilde{\mathfrak{c}} = \mathfrak{c}^{\mathfrak{m}}_{-2}$. Thus the first term in (2.10) is the first term in $\Phi^{N,2}$ after we average over $\mathfrak{l} \in [\![1, N^{\beta_X}]\!]$.
- Analysis of the cubic nonlinearity $\mathfrak{c}$ in Proposition 2.3 of [10] now amounts to computing the second term in (2.10). For this we compute the gradient of the microscopic Cole-Hopf transform by Taylor expansion of its exponential formula in terms of $\eta$-spins. Taylor expansion in this fashion is done in the proof of Proposition 2.3 in [10] for example. The result of such Taylor expansion gives a representation of the second term in (2.10) that we describe as follows. First we emphasize the support of the $\widetilde{\mathfrak{c}}^{\mathfrak{m}}_{-7\mathfrak{l}}$-factor in the second term in (2.10) is contained strictly to the left of $x - 7\mathfrak{l}\mathfrak{m}$. Taylor expansion gives

$$N^{\frac{1}{2}}\widetilde{\mathfrak{c}}^{\mathfrak{m}}_{-7\mathfrak{l}}\left(\nabla^{\mathfrak{m}}_{-7\mathfrak{l}}\mathfrak{Z}^N\right) = \sum_{\mathfrak{k}=1}^{\infty} N^{-\frac{1}{2}\mathfrak{k}+\frac{1}{2}}\lambda^{\mathfrak{k}}(\mathfrak{k}!)^{-1}\left(\widetilde{\mathfrak{c}}^{\mathfrak{m}}_{-7\mathfrak{l}}\left(\nabla^{\mathfrak{m}}_{-7\mathfrak{l}}\mathfrak{h}^N\right)^{\mathfrak{k}}\right) \qquad (2.11)$$

$$= \sum_{\mathfrak{k}=1}^{\infty} N^{-\frac{1}{2}\mathfrak{k}+\beta_X\mathfrak{k}+\frac{1}{2}}\lambda^{\mathfrak{k}}(\mathfrak{k}!)^{-1}\left(\widetilde{\mathfrak{c}}^{\mathfrak{m}}_{-7\mathfrak{l}}\left(N^{-\beta_X}\nabla^{\mathfrak{m}}_{-7\mathfrak{l}}\mathfrak{h}^N\right)^{\mathfrak{k}}\right). \qquad (2.12)$$

The Eq. (2.12) may be interpreted as the Taylor series for $\mathfrak{Z}^N$ and recalling $\mathfrak{Z}^N$ is the exponential of $\mathfrak{h}^N$. We now observe that the $\mathfrak{h}^N$-gradient is a linear polynomial in $\eta$-spins which are contained a neighborhood with length at most $\kappa_{\mathfrak{m}} N^{\beta_X}$ strictly to the right of $x - 7\mathfrak{l}\mathfrak{m}$ and thus disjoint from the support of $\widetilde{\mathfrak{c}}^{\mathfrak{m}}_{-7\mathfrak{l}}$. We thereby additionally observe that per sum-index $\mathfrak{k} \in \mathbb{Z}_{\geq 1}$ the corresponding summand in (2.12) with pseudo-gradient factor $\widetilde{\mathfrak{c}}^{\mathfrak{m}}_{-7\mathfrak{l}}$. Indeed observe the remaining functional-factor $N^{-\beta_X}\nabla^{\mathfrak{m}}_{-7\mathfrak{l}}\mathfrak{h}^N$ is uniformly bounded if $|\mathfrak{l}| \lesssim N^{\beta_X}$ since the $\eta$-spins are uniformly bounded. Additionally, the prefactor in the $\mathfrak{k}$-summand is at most $N^{\beta_X}N^{-(\mathfrak{k}-1)/2+(\mathfrak{k}-1)\beta_X}$ times uniformly bounded factors. Thus the infinite series (2.12) is summable as $\beta_X < \frac{1}{2}$ so it is a functional with pseudo-gradient factor $\widetilde{\mathfrak{c}}^{\mathfrak{m}}_{-7\mathfrak{l}}$ which is then scaled by $N^{\beta_X}$. After we average (2.12) over $\mathfrak{l} \in [\![1, N^{\beta_X}]\!]$ we may match the second term from (2.10) to the second term in $\Phi^{N,2}$. Alternatively instead of taking the entire infinite series

into the functional-with-pseudo-gradient-factor/second term in $\Phi^{N,2}$ we may cut this series off at $\mathfrak{k} = 4$ with an error that is uniformly vanishing in the large-$N$ limit.

We now return to the last difference between Proposition 2.3 from [10] and Proposition 2.7 concerning quadratic polynomials in $\eta$-spins that we remarked on at the beginning of this proof. Within Proposition 2.3 from [10] this quadratic polynomial was absorbed as a weakly vanishing term. However because we take the weaker Assumption 1.6 we must treat it differently here. First we introduce some notation to define this quadratic polynomial of interest.

- Define the length-$k$ neighborhood $\mathfrak{I}_{x,k} \overset{\bullet}{=} \{z_1, z_2 \in \mathbb{Z} : z_1 \leqslant x < z_2, \ z_2 - z_1 = k\}$ and define $\alpha_k \widetilde{\gamma}_k \overset{\bullet}{=} \alpha_k \gamma_k - \alpha_k \bar{\gamma}_k = \mathcal{O}(N^{-1/2})$.

This quadratic polynomial is the sum of terms below that we sum over $k \in [\![1, \mathfrak{m}]\!]$ and the neighborhood $\mathfrak{I}_{x,k}$:

$$
\begin{aligned}
N\alpha_k \widetilde{\gamma}_k \eta_{T,z_1} \eta_{T,z_2} &= N\alpha_k \widetilde{\gamma}_k (\eta_{T,z_1} \eta_{T,z_2} - \eta_{T,x} \eta_{T,x+1}) + N\alpha_k \widetilde{\gamma}_k \eta_{T,x} \eta_{T,x+1} \bullet \\
&= \Phi_{1;k,z_1,z_2} + \Phi_{2;k}.
\end{aligned} \tag{2.13}
$$

Observe the $\Phi_{1;}$-term is a pseudo-gradient because $\eta$-spins are exchangeable with respect to any canonical ensemble. This was the justification for the cubic nonlinearity in Proposition 2.3 in [10] being a pseudo-gradient as well. In particular after we sum over the two finite sets $k \in [\![1, \mathfrak{m}]\!]$ and $\mathfrak{I}_{x,k}$ we get another pseudo-gradient since pseudo-gradients are closed under uniformly bounded linear combinations. Thus we may employ the same decomposition and expansions for the sum of $\Phi_{1;}$-terms as those we used to address the cubic nonlinearity in Proposition 2.3 in [10]. We emphasize Assumption 1.6 implies $\Phi_{1;}$ is order $N^{1/2}$ like the aforementioned cubic nonlinearity. Meanwhile the second $\Phi_{2;k}$-terms are all independent of the $\mathfrak{I}_{x,k}$-variables. Therefore after we sum over $\mathfrak{I}_{x,k}$, we get from $\Phi_{2;k}$ the factor $k\alpha_k \widetilde{\gamma}_k$ times a $k$-independent factor. Summing over $k \in [\![1, \mathfrak{m}]\!]$ and applying Lemma 2.10 shows that $\Phi_{2;k}$-terms within the far RHS of (2.13) ultimately contribute zero to the $\mathfrak{Z}^N$-stochastic equation. □

*Proof of Lemma 2.10.* By definition of $\lambda \in \mathbb{R}$, it suffices to prove $\sum_{k=1}^{\mathfrak{m}} k\alpha_k \bar{\gamma}_k = \lambda \sum_{k=1}^{\mathfrak{m}} k^2 \alpha_k$. By definition of $\bar{\gamma}_k$,

$$
\begin{aligned}
\sum_{k=1}^{\mathfrak{m}} k\alpha_k \bar{\gamma}_k &= 2\lambda \sum_{k=1}^{\mathfrak{m}} \sum_{\ell=k}^{\mathfrak{m}} (\ell - k)\alpha_\ell + \lambda \sum_{k=1}^{\mathfrak{m}} k\alpha_k \tag{2.14} \\
&= 2\lambda \sum_{k=1}^{\mathfrak{m}} \sum_{\ell=k}^{\mathfrak{m}} \ell\alpha_\ell - 2\lambda \sum_{k=1}^{\mathfrak{m}} k \sum_{\ell=k}^{\mathfrak{m}} \alpha_\ell + \lambda \sum_{k=1}^{\mathfrak{m}} k\alpha_k.
\end{aligned}
$$

We rewrite both double summations on the far RHS by accumulating the resulting coefficients for all $\alpha_k$ with $k \in \mathbb{Z}_{>0}$. In the first double summation, we obtain $k\alpha_k$ a total of $k$-times provided any $k \in \mathbb{Z}_{>0}$. Inside the second double summation, we grab one copy of $j\alpha_k$ for each $j \in [\![1, k]\!]$. Combining these two observations with elementary calculations gives

$$
\begin{aligned}
2\lambda \sum_{k=1}^{\mathfrak{m}} \sum_{\ell=k}^{\mathfrak{m}} \ell\alpha_\ell &- 2\lambda \sum_{k=1}^{\mathfrak{m}} k \sum_{\ell=k}^{\mathfrak{m}} \alpha_\ell \\
&= 2\lambda \sum_{k=1}^{\mathfrak{m}} \left( k^2 - \sum_{j=1}^{k} j \right) \alpha_k = \lambda \sum_{k=1}^{\mathfrak{m}} (k^2 - k)\alpha_k. \tag{2.15}
\end{aligned}
$$

Combining the previous two displays completes the proof. □

*2.3. Strategy.* We now discuss "proof" of Theorem 1.7. We clarify that we will not present results here that we actually get in this paper. The reasons for this are technical and instead we will get slightly adapted quantitative versions of the following; we discuss these reasons at the end of this strategy discussion. For this reason, all results discussed *here* will be "pseudo-results".

The proof of Theorem 1.7 is built on the following key result that effectively says we can forget about the $\Phi^{N,2}$-contribution in the stochastic equation for $\mathfrak{Z}^N$ from Proposition 2.7. We discuss its importance and implications afterwards.

**Pseudo-Proposition 2.11.** *Define a norm* $\|\varphi\|_{\mathfrak{t};\mathbb{X}} \overset{\bullet}{=} \sup_{0 \leqslant \mathfrak{s} \leqslant \mathfrak{t}} \sup_{x \in \mathbb{X}} |\varphi_{\mathfrak{s},x}|$ *for any* $\mathfrak{t} \in \mathbb{R}_{\geqslant 0}$ *and* $\mathbb{X} \subseteq \mathbb{R}$. *We get the convergence in probability* $\|\mathfrak{P}^N(\Phi^{N,2})\|_{1;\mathbb{Z}} \to_{N \to \infty} 0$ *where* $\Phi^{N,2}$ *is defined in the statement of* Proposition 2.7.

Due to standard linear procedure, it is enough to analyze the process $\mathfrak{Y}^N$ that solves the same stochastic equation as $\mathfrak{Z}^N$ from Proposition 2.7 but without the $\Phi^{N,2}$-term and replacing all factors of $\mathfrak{Z}^N$ with $\mathfrak{Y}^N$ and while keeping the same initial data. More precisely it is left to show $\mathfrak{Y}^N$ converges to the solution of SHE. At this point, outside of the *exact* microscopic version of the SHE corresponding to the first two terms in the stochastic equation of $\mathfrak{Y}^N$ from Proposition 2.7, there is only weakly-vanishing data. Thus the SHE scaling limit for $\mathfrak{Y}^N$ can be computed using the same ideas from hydrodynamic limits that were used in the proof of Theorem 1.1 in [10]. There is a caveat to this that is ultimately negligible; see the end of the proof of Proposition 8.3.

We discuss elements of the would-be-proof of Pseudo-Proposition 2.11 below. For simplicity, let us assume that the gradient term in $\Phi^{N,2}$ is equal to 0. To control such a gradient term, we use a simple summation-by-parts argument and apply regularity estimates for the heat kernel in the heat operator $\mathfrak{P}^N$ like in [10]. Let us also assume the order $N^{\beta_X}$-term in $\Phi^{N,2}$ in Proposition 2.7 is equal to 0. It turns out that all steps we require to control this $N^{\beta_X}$-term will be steps we need to control the leading-order $N^{1/2}$-term in $\Phi^{N,2}$ in Proposition 2.7 anyway. The first step to control the order $N^{1/2}$-term in $\Phi^{N,2}$ is an introduction-of-cutoff.

**Pseudo-Proposition 2.12.** *Define the cutoff-spatial average* $\mathscr{C}^{\mathbf{X},-}$ *as equal to* $\mathscr{A}^{\mathbf{X},-}_{N^{\beta_X}}(\mathfrak{g})$ *in* $\Phi^{N,2}$ *in* Proposition 2.7 *but only when this spatial average is at most* $N^{\varepsilon-\beta_X/2}$ *in absolute value. On the complement event, define* $\mathscr{C}^{\mathbf{X},-}$ *to be 0. We have, in probability,*

$$\lim_{N\to\infty}\|\mathfrak{P}^N(N^{\frac{1}{2}}|\mathscr{A}^{\mathbf{X},-}_{N^{\beta_X}}(\mathfrak{g}_{S,y}) - \mathscr{C}^{\mathbf{X},-}_{S,y}|\mathfrak{Z}^N)\|_{1;\mathbb{Z}} = 0. \qquad (2.16)$$

It turns out that such a step will be unnecessary for the order $N^{\beta_X}$-term in $\Phi^{N,2}$ from Proposition 2.7 just because it is lower-order. We will explain both the would-be-proof and utility of the previous Pseudo-Proposition 2.12 in future subsections. The key second step towards the would-be-proof of Pseudo-Proposition 2.11 is the following replacement-by-time-average.

**Pseudo-Proposition 2.13.** *Given* $\mathfrak{t} \in \mathbb{R}_{>0}$ *let us define* $\mathscr{A}^{\mathfrak{t};\mathbf{T},+}_{S,y} \overset{\bullet}{=} \mathfrak{t}^{-1}\int_0^{\mathfrak{t}} \mathscr{C}^{\mathbf{X},-}_{S+\tau,y}\mathrm{d}\tau$ *and for* $\mathfrak{t} = 0$ *we instead define* $\mathscr{A}^{\mathfrak{t};\mathbf{T},+} = \mathscr{C}^{\mathbf{X},-}$. *In words* $\mathscr{A}^{\mathfrak{t};\mathbf{T},+}$ *is the time-average of* $\mathscr{C}^{\mathbf{X},-}$ *with respect to time-scale* $\mathfrak{t} \in \mathbb{R}_{\geqslant 0}$ *in the "positive time-direction".*

*Defining* $\mathfrak{t}_{\max} \overset{\bullet}{=} N^{-1}$, *we have the convergence in probability*

$$\lim_{N\to\infty}\|\mathfrak{P}^N\left(N^{\frac{1}{2}}\left(\mathscr{A}^{\mathfrak{t}_{\max};\mathbf{T},+}_{S,y} - \mathscr{C}^{\mathbf{X},-}_{S,y}\right)\mathfrak{Z}^N\right)\|_{1;\mathbb{Z}} = 0. \qquad (2.17)$$

The conclusion of Pseudo-Proposition 2.13 is replacement of the spatial-average-with-cutoff $\mathscr{C}^{\mathbf{X},-}$ defined in the statement of Pseudo-Proposition 2.12 with its time-average on time-scale $\mathfrak{t}_{\max} = N^{-1}$. To take advantage of such replacement we would need to estimate the time-average $\mathscr{A}^{\mathfrak{t}_{\max};\mathbf{T},+}$. Ultimately, we would get the following.

**Pseudo-Proposition 2.14.** *Admit the setting of* Pseudo-Proposition *2.13. We have the convergence in the probability*

$$\lim_{N\to\infty}\|\mathfrak{P}^N(N^{\frac{1}{2}}\mathscr{A}_{S,y}^{\mathfrak{t}_{\max};\mathbf{T},+}\mathfrak{Z}^N)\|_{1;\mathbb{Z}} = 0. \tag{2.18}$$

Pseudo-Proposition 2.11 would now follow from the triangle inequality for the norm $\|\|_{1;\mathbb{Z}}$. We now explain the would-be-proofs for each of Pseudo-Proposition 2.12, Pseudo-Proposition 2.13, and Pseudo-Proposition 2.14.

*2.4. Strategy-local equilibrium.* Before we start the would-be-proofs for Pseudo-Proposition 2.12, Pseudo-Proposition 2.13, and Pseudo-Proposition 2.14, we first introduce key invariant measure calculations. The first is a large-deviations-type estimate which we establish in Lemma 3.13 and Corollary 3.14, and the second is the Kipnis-Varadhan "Brownian" inequality which we establish in Lemma 3.10, Lemma 3.11, and Lemma 3.12.

**Lemma 2.15.** *Suppose the particle system starts at any "canonical ensemble" invariant measure so that the measure on particle configurations on any finite subset of $\mathbb{Z}$ is a convex combination of canonical ensembles in* Definition *3.1 on that subset. We have*

$$\lim_{N\to\infty}N^{\frac{1}{2}}\mathbf{E}|\mathscr{A}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{S,y}) - \mathscr{C}_{S,y}^{\mathbf{X},-}| = 0. \tag{2.19}$$

**Lemma 2.16.** *Suppose the particle system starts at a canonical ensemble invariant measure as in* Lemma *2.15. For any $\mathfrak{t} \in \mathbb{R}_{\geqslant 0}$,*

$$N^{1/2}\mathbf{E}|\mathscr{A}_{S,y}^{\mathfrak{t};\mathbf{T},+}| \lesssim N^{-1/2}\mathfrak{t}^{-1/2}N^{-\beta_X/2+\varepsilon}. \tag{2.20}$$

The proof of Lemma 2.15 and Lemma 2.16 depend heavily on invariant measures. However, we do not need these bounds *pointwise* in space-time but only after the law of the particle system is *averaged* against the heat kernel in space-time because of the "heat-operator-integrated" structure of the proposed estimates within Pseudo-Proposition 2.12, Pseudo-Proposition 2.13, and Pseudo-Proposition 2.14. From [14], in this *averaged* sense the law of the particle system is close to a convex combination of canonical ensemble invariant measures at least on mesoscopic space-time scales. We now note that the statistics we address in Lemma 2.15 and Lemma 2.16 are mesoscopic statistics. Thus, we expect that Lemma 2.15 and Lemma 2.16 still hold beyond the canonical ensemble initial measures when the absolute values therein are integrated against the heat kernel.

However, there are obstructions in actually implementing such a local equilibrium idea that concern how local equilibrium works at a more technical but still important level. We will refer to them in the would-be-proof of Pseudo-Proposition 2.14.

- Following [14], we reduce to local equilibrium using entropy production estimates. Because of the asymmetry in the model not being sufficiently weak, and because of the infinite-volume features of the lattice $\mathbb{Z}$ where the exclusion process evolves, establishing entropy production estimates in our setting is noticeably more difficult; Sect. 3 addresses this.

• We reduce to a local equilibrium by comparing the averaged law of the particle system to an invariant measure at the level of relative entropy using the entropy inequality in Appendix 1.8 in [21] which connects relative entropy to large deviations, for example. We also require a log-Sobolev inequality for the symmetric simple exclusion process established in [25] to control relative entropy by Dirichlet form in order to use entropy production from the first bullet point. The LSI in [25] is quadratic in the length-scale we want to reduce to local equilibrium on, so to reduce to local equilibrium on bigger length-scales, we need better entropy production or control, at a level of large deviations, on the functional that we are estimating in expectation.

### 2.5. Pseudo-Proposition 2.12.

Assume Lemma 2.15 holds without expectation and the $\mathfrak{Z}^N$-process is uniformly bounded as it is supposed to look like the SHE solution. By Lemma 2.15, there is nothing to do for invariant measure initial data. However, because the statistic we estimate in Lemma 2.15 is mesoscopic, in general we can finish by reducing to local equilibrium.

### 2.6. Pseudo-Proposition 2.13.

We use the fast-variable/slow-variable idea from stochastic homogenization. This proposes that on the time-scale on which we want to replace $\mathscr{C}^{\mathbf{X},-}$ by time-average, the heat kernel and the $\mathfrak{Z}^N$-process both exhibit little variation. The error for such a replacement-by-time-average is thus controlled by time-regularity of the heat kernel times $\mathscr{C}^{\mathbf{X},-}$ and time-regularity of the $\mathfrak{Z}^N$-process times the size of $\mathscr{C}^{\mathbf{X},-}$. The heat kernel is smooth with an integrable "enough" short-time singularity. On the other hand, we also expect that $\mathfrak{Z}^N$ has time-regularity matching that of the SHE solution, and this is Holder regularity of exponent $(\frac{1}{4})^-$. Thus, the time-regularity for $\mathfrak{Z}^N$ is worse than that for the heat kernel, so we control the error given by time-regularity of $\mathfrak{Z}^N$ times $\mathscr{C}^{\mathbf{X},-}$. For any time-scale $\mathfrak{t} \in \mathbb{R}_{\geqslant 0}$, we thus expect the resulting error to be at most, recalling $\mathscr{C}^{\mathbf{X},-}$ has an a priori bound by definition in Pseudo-Proposition 2.12, of order roughly equal to:

$$N^{1/2}\|\mathscr{C}^{\mathbf{X},-}\|_\infty \mathfrak{t}^{1/4} \lesssim N^{1/2-\beta_X/2+\varepsilon}\mathfrak{t}^{1/4} \lesssim N^{1/3}\mathfrak{t}^{1/4}. \tag{2.21}$$

Recall $\beta_X = 1/3 + \varepsilon_{X,1}$ in the statement of Proposition 2.7. Pseudo-Proposition 2.13 proposes that we take $\mathfrak{t} = \mathfrak{t}_{\max} = N^{-1}$ for which the RHS of the estimate (2.21) blows up. To remedy this we will appeal to the following multiscale procedure. Roughly it proposes an initial replacement of $\mathscr{C}^{\mathbf{X},-}$ by its time-average on *some* time-scale and applying Lemma 2.16 and local equilibrium to boost our a priori bound for $\mathscr{C}^{\mathbf{X},-}$ after time-average. First we assume for simplicity that Lemma 2.16 holds without expectation.

• The bound (2.21) tells us we can pick the time-scale $\mathfrak{t} = \mathfrak{t}_{1,1} = N^{-2+\varepsilon_1}$ for $\varepsilon_1 \in \mathbb{R}_{>0}$ arbitrarily small but still universal, so we first replace $\mathscr{C}^{\mathbf{X},-}$ with its time-average $\mathscr{A}^{\mathfrak{t}_{1,1};\mathbf{T},+}$ with respect to this preliminary time-scale $\mathfrak{t}_{1,1} = N^{-2+\varepsilon_1}$.

• Let us now replace the time-average $\mathscr{A}^{\mathfrak{t}_{1,1};\mathbf{T},+}$ with respect to the time-scale $\mathfrak{t}_{1,1} = N^{-2+\varepsilon_1}$ with the time-average $\mathscr{A}^{\mathfrak{t}_{1,2};\mathbf{T},+}$ with respect to the time-scale $\mathfrak{t}_{1,2} = N^{\varepsilon_1}\mathfrak{t}_{1,1}$ inside of the heat operator integrated against both the heat kernel and the $\mathfrak{Z}^N$-process. To this end, we first observe that $\mathscr{A}^{\mathfrak{t}_{1,2};\mathbf{T},+}$ is basically a time-average of $\mathscr{A}^{\mathfrak{t}_{1,1};\mathbf{T},+}$ on the latter/larger time-scale $\mathfrak{t}_{1,2}$. Therefore, the error we get is (2.21) as before but now with $\mathscr{A}^{\mathfrak{t}_{1,1};\mathbf{T},+}$ in place of $\mathscr{C}^{\mathbf{X},-}$ and with $\mathfrak{t} = \mathfrak{t}_{1,2} = N^{\varepsilon_1}\mathfrak{t}_{1,1}$. Lemma 2.16 gives

the following ultimately vanishing bound for the error in this second "replacement-by-time-average" step:

$$
\begin{aligned}
N^{1/2}\|\mathscr{A}^{\mathfrak{t}_{1,1};\mathbf{T},+}\|_\infty \mathfrak{t}_{1,2}^{1/4} &\lesssim N^{-1/2-\beta_X/2+\varepsilon}\mathfrak{t}_{1,1}^{-1/2}\mathfrak{t}_{1,2}^{1/4} \lesssim N^{-2/3+\varepsilon}\mathfrak{t}_{1,1}^{-1/4}\mathfrak{t}_{1,2}^{1/4}\cdot\mathfrak{t}_{1,1}^{-1/4} \\
&= N^{-2/3+\varepsilon+\varepsilon_1}\mathfrak{t}_{1,1}^{-1/4} \\
&\lesssim N^{-1/6+\varepsilon+\varepsilon_1}.
\end{aligned} \tag{2.22}
$$

- The rest of the multiscale procedure is then replacing $\mathscr{A}^{\mathfrak{t}_{1,n};\mathbf{T},+}$ with $\mathscr{A}^{\mathfrak{t}_{1,n+1};\mathbf{T},+}$, in which $\mathfrak{t}_{1,n+1} = N^{\varepsilon_1}\mathfrak{t}_{1,n}$, until we arrive at the final time-scale $\mathfrak{t}_{max} = N^{-1}$. As $\mathfrak{t}_{1,1} = N^{-2+\varepsilon_1}$, we only require a $\varepsilon_1$-dependent number of steps/errors to get to $\mathfrak{t}_{max} = N^{-1}$.

*2.7. Pseudo-Proposition 2.14.* Lemma 2.16 with $\mathfrak{t} = \mathfrak{t}_{max} = N^{-1}$ yields $N^{1/2}|\mathscr{A}^{\mathfrak{t}_{max};\mathbf{T},+}| \lesssim N^{-\beta_X/2+\varepsilon}$, so the proposed bound in Pseudo-Proposition 2.14 holds if the exclusion process starts at a canonical ensemble invariant measure. We would then like to perform a reduction to local equilibrium since $\mathscr{A}^{\mathfrak{t}_{max};\mathbf{T},+}$ is a mesoscopic statistic. However, there is a serious obstruction that cannot be easily circumvented. To explain this, let us observe that for longer time-scales $\mathfrak{t} \in \mathbb{R}_{\geqslant 0}$, the statistic $\mathscr{A}^{\mathfrak{t};\mathbf{T},+}$ depends on spins on a larger $\mathfrak{t}$-dependent block. For example, the LSI for exclusion processes from [25] grows in the length-scale of the set on which the exclusion process lives. For the time-scale $\mathfrak{t}_{max} = N^{-1}$, the associated length-scale is too large for us to apply local equilibrium. We resolve this via the following multiscale procedure that makes local equilibrium accessible by slowly boosting a priori large-deviations estimates for time-averages. Recall better large-deviations estimates help with local equilibrium.

- The issue with local equilibrium was that the time-scale $\mathfrak{t}_{max} = N^{-1}$ was too long. With this in mind, we first consider shorter time-scales. We write $\mathscr{A}^{\mathfrak{t}_{max};\mathbf{T},+}$ as an average of time-shifted time-averages on a time-scale $0 < \mathfrak{t}_{2,1} \leqslant \mathfrak{t}_{max}$ to be determined:

$$
\mathscr{A}^{\mathfrak{t}_{max};\mathbf{T},+}_{S,y} = \widetilde{\sum}_{\mathfrak{l}=0}^{\mathfrak{t}_{max}\mathfrak{t}_{2,1}^{-1}-1}\mathscr{A}^{\mathfrak{t}_{2,1};\mathbf{T},+}_{S+\mathfrak{l}\mathfrak{t}_{2,1},y}. \tag{2.23}
$$

The identity (2.23) follows by noting that averaging on a large time-scale is the same as averages on a smaller time-scale and averaging all the small-scale averages together. If $\mathfrak{t}_{2,1} \in \mathbb{R}_{>0}$ is a "short enough" time-scale, we may analyze every summand within the RHS of (2.23) using Lemma 2.16 and a reduction to local equilibrium. However, if we choose $\mathfrak{t}_{2,1} \in \mathbb{R}_{>0}$ too small our estimates for the summands on the RHS of (2.23) will not vanish in the large-$N$ limit when equipped with the additional factor of $N^{1/2}$ in the proposed estimate in the statement of Pseudo-Proposition 2.14. Because we cannot pick $\mathfrak{t}_{2,1} \in \mathbb{R}_{>0}$ large enough to get a vanishing estimate for the summands on the RHS of (2.23) while still being able to employ local equilibrium, we instead do the following with motivation given shortly. Take $\beta_{2,1} = \frac{1}{2}\beta_X - \varepsilon + \varepsilon_2$; building on (2.23), we introduce cutoff:

$$
\mathscr{A}^{\mathfrak{t}_{max};\mathbf{T},+}_{S,y} = \widetilde{\sum}_{\mathfrak{l}=0}^{\mathfrak{t}_{max}\mathfrak{t}_{2,1}^{-1}-1}\mathscr{A}^{\mathfrak{t}_{2,1};\mathbf{T},+}_{S+\mathfrak{l}\mathfrak{t}_{2,1},y} = \widetilde{\sum}_{\mathfrak{l}=0}^{\mathfrak{t}_{max}\mathfrak{t}_{2,1}^{-1}-1}\mathscr{A}^{\mathfrak{t}_{2,1};\mathbf{T},+}_{S+\mathfrak{l}\mathfrak{t}_{2,1},y}\mathbf{1}_{|\mathscr{A}^{\mathfrak{t}_{2,1};\mathbf{T},+}_{S+\mathfrak{l}\mathfrak{t}_{2,1},y}|\lesssim N^{-\beta_{2,1}}}
$$

$$
\tag{2.24}
$$

$$+ \sum_{\mathfrak{l}=0}^{\widetilde{\mathfrak{t}_{\max}\mathfrak{t}_{2,1}^{-1}-1}} \mathscr{A}_{S+\mathfrak{l}\mathfrak{t}_{2,1},y}^{\mathfrak{t}_{2,1};\mathbf{T},+} \mathbf{1}_{|\mathscr{A}_{S+\mathfrak{l}\mathfrak{t}_{2,1},y}^{\mathfrak{t}_{2,1};\mathbf{T},+}| \gtrsim N^{-\beta_{2,1}}}.$$

(2.25)

The first sum on the far RHS of the identity (2.25) gives each $\mathscr{A}^{\mathfrak{t}_{2,1};\mathbf{T},+}$-summand therein an *improved* a priori upper bound of $N^{-\beta_{2,1}}$ which is $N^{-\varepsilon_2}$ better than the a priori bound of $N^{-\beta_X/2+\varepsilon}$ for $\mathscr{A}^{\mathfrak{t}_{2,1};\mathbf{T},+}$ that is inherited from the $\mathscr{C}^{\mathbf{X},-}$ it averages. For the second sum on the far RHS of (2.25), we pick $\mathfrak{t}_{2,1} \in \mathbb{R}_{>0}$ small enough to apply local equilibrium but large enough so that the summands are negligible with high-probability with respect to canonical ensembles. To see where this negligible feature of the summands in this second sum on the far RHS of (2.25) comes from, we observe that if $\mathfrak{t}_{2,1} \in \mathbb{R}_{>0}$ is sufficiently large then Lemma 2.16 tells us the event $|\mathscr{A}^{\mathfrak{t}_{2,1};\mathbf{T},+}| \gtrsim N^{-\beta_{2,1}}$ happens with low-probability. We emphasize that the point for such step is to accomplish a lesser goal. We will not try to estimate $\mathscr{A}^{\mathfrak{t}_{2,1};\mathbf{T},+}$ by $N^{-1/2}$ but rather by $N^{-\beta_{2,1}}$. Because we want a less sharp estimate, we may choose a shorter time-scale $0 < \mathfrak{t}_{2,1} \leqslant \mathfrak{t}_{\max}$ that does the job and for which we can apply local equilibrium.

• At the end of the previous bullet point, we are now left with the first sum on the far RHS of (2.25). This is an average of time-averages $\mathscr{A}^{\mathfrak{t}_{2,1};\mathbf{T},+}$ with a priori upper bound cutoff of $N^{-\beta_{2,1}}$ that improves on the $N^{-\beta_X/2+\varepsilon}$-cutoff on $\mathscr{C}^{\mathbf{X},-}$ by a factor of $N^{-\varepsilon_2}$. Because each of these time-averages comes with the same upper bound cutoff of $N^{-\beta_{2,1}}$ we can group these time-averages on time-scale $\mathfrak{t}_{2,1} \in \mathbb{R}_{>0}$ into $\mathfrak{t}_{\max}\mathfrak{t}_{2,2}^{-1}$ many time-averages on a to-be-determined time-scale $\mathfrak{t}_{2,2}$ that is larger than $\mathfrak{t}_{2,1}$, and each of these scale $\mathfrak{t}_{2,2}$-averages will also inherit the a priori upper bound cutoff of $N^{-\beta_{2,1}}$ as it averages terms with such cutoff. So

$$\sum_{\mathfrak{l}=0}^{\widetilde{\mathfrak{t}_{\max}\mathfrak{t}_{2,1}^{-1}-1}} \mathscr{A}_{S+\mathfrak{l}\mathfrak{t}_{2,1},y}^{\mathfrak{t}_{2,1};\mathbf{T},+} \mathbf{1}_{|\mathscr{A}_{S+\mathfrak{l}\mathfrak{t}_{2,1},y}^{\mathfrak{t}_{2,1};\mathbf{T},+}| \lesssim N^{-\beta_{2,1}}} \approx \sum_{\mathfrak{l}=0}^{\widetilde{\mathfrak{t}_{\max}\mathfrak{t}_{2,2}^{-1}-1}} \mathscr{A}_{S+\mathfrak{l}\mathfrak{t}_{2,2},y}^{\mathfrak{t}_{2,2};\mathbf{T},+} \mathbf{1}_{|\mathscr{A}_{S+\mathfrak{l}\mathfrak{t}_{2,2},y}^{\mathfrak{t}_{2,2};\mathbf{T},+}| \lesssim N^{-\beta_{2,1}}}.$$

(2.26)

The error terms corresponding to the approximation (2.26) look sufficiently like the second sum on the far RHS of (2.25), it turns out after unfolding definitions. To be just a little more precise, the error terms are given by scale-$\mathfrak{t}_{2,1}$ time-averages with both upper and lower bound cutoffs. We then pick $\beta_{2,2} = \beta_{2,1} + \varepsilon_2$ and improve our a priori $N^{-\beta_{2,1}}$-bound similar to (2.25):

$$\sum_{\mathfrak{l}=0}^{\widetilde{\mathfrak{t}_{\max}\mathfrak{t}_{2,2}^{-1}-1}} \mathscr{A}_{S+\mathfrak{l}\mathfrak{t}_{2,2},y}^{\mathfrak{t}_{2,2};\mathbf{T},+} \mathbf{1}_{|\mathscr{A}_{S+\mathfrak{l}\mathfrak{t}_{2,2},y}^{\mathfrak{t}_{2,2};\mathbf{T},+}| \lesssim N^{-\beta_{2,1}}} = \sum_{\mathfrak{l}=0}^{\widetilde{\mathfrak{t}_{\max}\mathfrak{t}_{2,2}^{-1}-1}} \mathscr{A}_{S+\mathfrak{l}\mathfrak{t}_{2,2},y}^{\mathfrak{t}_{2,2};\mathbf{T},+} \mathbf{1}_{|\mathscr{A}_{S+\mathfrak{l}\mathfrak{t}_{2,2},y}^{\mathfrak{t}_{2,2};\mathbf{T},+}| \lesssim N^{-\beta_{2,2}}}$$

$$+ \sum_{\mathfrak{l}=0}^{\widetilde{\mathfrak{t}_{\max}\mathfrak{t}_{2,2}^{-1}-1}} \mathscr{A}_{S+\mathfrak{l}\mathfrak{t}_{2,2},y}^{\mathfrak{t}_{2,2};\mathbf{T},+} \mathbf{1}_{N^{-\beta_{2,2}} \lesssim |\mathscr{A}_{S+\mathfrak{l}\mathfrak{t}_{2,2},y}^{\mathfrak{t}_{2,2};\mathbf{T},+}| \lesssim N^{-\beta_{2,1}}}.$$

(2.27)

Again, if we pick $\mathfrak{t}_{2,2} \in \mathbb{R}_{>0}$ sufficiently large but not too much larger than $\mathfrak{t}_{2,1} \in \mathbb{R}_{>0}$, then we can use Lemma 2.16 and local equilibrium to argue the second sum within the RHS of the previous display is negligible with high-probability similar to our reasoning to control the second sum within the far RHS of (2.25). We now discuss why we can reduce to a local equilibrium on this larger time-scale $\mathfrak{t}_{2,2} \in \mathbb{R}_{>0}$. The indicator functions attached to summands in the second sum within the RHS of the previous display come equipped with a priori $N^{-\beta_{2,1}}$-upper bounds that are $N^{-\varepsilon_2}$-*better* than our a priori upper bounds for the summands in the second sum on the

far RHS of (2.25). The a priori upper bounds are deterministic and at the level of large-deviations trivially. Because we have better a priori upper bounds of $N^{-\beta_{2,1}}$ for the $\mathscr{A}^{\mathfrak{t}_{2,2};\mathbf{T},+}$-functional we are estimating in expectation, we can reduce bounds of scale $\mathfrak{t}_{2,2}$-data to local equilibrium as $\mathfrak{t}_{2,2} \in \mathbb{R}_{>0}$ is only slightly larger than the previous time-scale $\mathfrak{t}_{2,1} \in \mathbb{R}_{>0}$. We emphasize the error terms in the second sum on the RHS of the last display are time-averages with a priori upper bound and lower bound cutoffs that are $N^{\varepsilon_2}$-off from each other.

- We iterate these improvements on time-scale and a priori upper bounds until we arrive at the a priori upper bound $N^{-1/2-\varepsilon}$. Because we improve our a priori upper bounds by $N^{-\varepsilon_2}$ at every step, we require only a $\varepsilon_2$-dependent number of iterations. All error terms in such scheme are time-averages with a priori upper and lower bound cutoffs which differ by a factor of $N^{\varepsilon_2}$. Moreover, all time-scales on which we perform a time-average will be at most the $\mathfrak{t}_{\max}$-scale from Pseudo-Proposition 2.13.

*2.8. Technical comments.* We start with why/in what way Pseudo-Propositions 2.12, 2.13, and 2.14 are *pseudo*-results.

- For technical reasons it will be convenient to "compactify" the microscopic Cole-Hopf transform $\mathfrak{Z}^N$. The current microscopic Cole-Hopf transform solves a stochastic equation on the infinite set $\mathbb{R}_{\geqslant 0} \times \mathbb{Z}$. We will replace it by the solution of the same stochastic equation but "periodized-in-space" onto a set $\mathbb{R}_{\geqslant 0} \times \mathbb{T}_N$ as the $\|\|\|_{1;\mathbb{T}_N}$-norm is easier to work with than $\|\|\|_{1;\mathbb{Z}}$. Here $\mathbb{T}_N \subseteq \mathbb{Z}$ is a torus that is much larger than the macroscopic length-scale $N$. Thus, it should not change any scaling limits.
- In view of the previous bullet point, we will only prove estimates for $\Phi^{N,2}$ with respect to $\|\|\|_{1;\mathbb{T}_N}$-norms and not $\|\|\|_{1;\mathbb{Z}}$-norms.
- The details behind the version of Pseudo-Proposition 2.11 that we actually get are different at a technical level. In particular, we will actually take a different, but still morally similar, value for $\mathfrak{t}_{\max}$ from Pseudo-Proposition 2.13.
- The stochastic equation for $\mathfrak{Z}^N$ in Proposition 2.7 is multiplicative in $\mathfrak{Z}^N$, so it is not enough to study only local functionals of the particle system. We must also control $\|\mathfrak{Z}^N\|_{1;\mathbb{T}_N}$. This ends up being a technical point we treat with a continuity argument from PDE. It is basically a strategy that says control on initial data $\mathfrak{Z}^N_{0,\bullet}$ propagates itself via control on local functionals.

Let us now give an outline of the rest of this paper.

- In Sect. 3 we construct local equilibrium and give invariant measure estimates. In Sect. 4 we "compactify" $\mathbb{Z} \to \mathbb{T}_N$.
- In Sect. 5, we get the spatial/time-average estimates alluded to in the would-be-proofs of Pseudo-Propositions 2.12, 2.13, and 2.14. We do this via local equilibrium and a "dynamical one-block step"/dynamic version of the one-block step in [14].
- In Sect. 6 we establish time-regularity estimates for $\mathfrak{Z}^N$ using mostly standard ideas that are present in [10], for example.
- We establish a strong version of Pseudo-Proposition 2.11 in Sect. 7 following basically the previously outlined strategy. In Sect. 8 we apply this strong estimate to pretend $\Phi^{N,2} \approx 0$ with high probability. We then follow the arguments in [10].

As for a possibly clearer reading of this paper, the reader is invited to do the following.

- Sections 3, 5, 6, and 7 are all technical. The reader is invited to skim these sections and first take their results for granted and then read the paper in its written order to

get a "blackbox" proof of Theorem 1.7. For technical aspects behind showing the important ingredients within Sects. 3, 5, 6, and 7, the reader is invited to read Sect. 7 followed by Sects. 3, 5, then Sect. 6. We have organized Sects. 3, 5, 6, and 7 to provide proofs of main results first, deferring proofs of technical manipulations/ingredients until the end of the section.

## 3. Local Equilibrium Estimates

This section is a preliminary discussion consisting of tools from hydrodynamic limits. We organize it as follows.

- We will first introduce the key invariant measures for relevant exclusion processes. These are the canonical ensembles and grand-canonical ensembles that we used to define the three classes of functionals at the beginning of Sect. 2.
- We move to entropy production in Proposition 3.6. We use it to perform "reduction to local equilibrium" in Lemma 3.8.
- We then move to invariant measure calculations that will be important *after* our reduction to local equilibrium. These tools are given in Lemma 3.10, Lemma 3.11, Lemma 3.12, Lemma 3.13, and Corollary 3.14.

**Definition 3.1.** For $\mathfrak{I} \subseteq \mathbb{Z}$ and $\varrho \in \mathbb{R}$, let the canonical ensemble/measure $\mu_{\varrho,\mathfrak{I}}^{\mathrm{can}}$ be the uniform measure on the hyperplane

$$\mathfrak{H}_{\varrho,\mathfrak{I}} \overset{\bullet}{=} \left\{ \eta \in \Omega_{\mathfrak{I}} : \widetilde{\sum}_{x \in \mathfrak{I}} \eta_x = \varrho \right\} \subseteq \Omega_{\mathfrak{I}}. \tag{3.1}$$

We define the grand-canonical ensemble/measure $\mu_{\varrho,\mathfrak{I}}$ as a product measure on $\Omega_{\mathfrak{I}} = \{\pm 1\}^{\mathfrak{I}}$ such that $\mathbf{E}\eta_x = \varrho$ for all $x \in \mathfrak{I}$. We note that $\varrho \in \mathbb{R}$ denotes *spin density* and not *particle density* in our conventions.

**Definition 3.2.** Consider any subset $\mathfrak{I} \subseteq \mathbb{Z}$ and parameter $\varrho \in \mathbb{R}$, and suppose that $\mathfrak{f} : \Omega_{\mathfrak{I}} \to \mathbb{R}_{\geqslant 0}$ is a probability density with respect to either the canonical ensemble $\mu_{\varrho,\mathfrak{I}}^{\mathrm{can}}$ or the grand-canonical ensemble $\mu_{\varrho,\mathfrak{I}}$ depending on the context below.

- The *relative entropy* with respect to $\mu_{\varrho,\mathfrak{I}}^{\mathrm{can}}$ and $\mu_{\varrho,\mathfrak{I}}$, respectively, are the functionals

$$\mathfrak{H}_{\varrho,\mathfrak{I}}^{\mathrm{can}}(\mathfrak{f}) \overset{\bullet}{=} \mathbf{E}^{\mu_{\varrho,\mathfrak{I}}^{\mathrm{can}}} \mathfrak{f} \log \mathfrak{f} \quad \text{and} \quad \mathfrak{H}_{\varrho,\mathfrak{I}}(\mathfrak{f}) \overset{\bullet}{=} \mathbf{E}^{\mu_{\varrho,\mathfrak{I}}} \mathfrak{f} \log \mathfrak{f}. \tag{3.2}$$

- The *Dirichlet form* with respect to $\mu_{\varrho,\mathfrak{I}}^{\mathrm{can}}$ and $\mu_{\varrho,\mathfrak{I}}$, respectively, are defined as follows in which we recall that $\mathfrak{S}_{x,y}$ denotes the generator for the symmetric exclusion process with speed 1 on the bond $\{x, y\} \subseteq \mathbb{Z}$. Recall the set A of symmetric speeds $\alpha_{\bullet}$:

$$\mathfrak{D}_{\varrho,\mathfrak{I}}^{\mathrm{can}}(\mathfrak{f}) \overset{\bullet}{=} -\mathbf{E}^{\mu_{\varrho,\mathfrak{I}}^{\mathrm{can}}} \mathfrak{f}^{\frac{1}{2}} \cdot \sum_{x,y \in \mathfrak{I}} \alpha_{|x-y|} \mathfrak{S}_{x,y} \mathfrak{f}^{\frac{1}{2}} = -\sum_{x,y \in \mathfrak{I}} \alpha_{|x-y|} \mathbf{E}^{\mu_{\varrho,\mathfrak{I}}^{\mathrm{can}}} \mathfrak{f}^{\frac{1}{2}} \mathfrak{S}_{x,y} \mathfrak{f}^{\frac{1}{2}} \tag{3.3a}$$

$$\mathfrak{D}_{\varrho,\mathfrak{I}}(\mathfrak{f}) \overset{\bullet}{=} -\mathbf{E}^{\mu_{\varrho,\mathfrak{I}}} \mathfrak{f}^{\frac{1}{2}} \cdot \sum_{x,y \in \mathfrak{I}} \alpha_{|x-y|} \mathfrak{S}_{x,y} \mathfrak{f}^{\frac{1}{2}} = -\sum_{x,y \in \mathfrak{I}} \alpha_{|x-y|} \mathbf{E}^{\mu_{\varrho,\mathfrak{I}}} \mathfrak{f}^{\frac{1}{2}} \mathfrak{S}_{x,y} \mathfrak{f}^{\frac{1}{2}}. \tag{3.3b}$$

*Remark 3.3.* The Dirichlet form is usually defined without a square root of $\mathfrak{f}$, but we adopt it as it appears frequently here.

*3.1. Entropy production.* We present here an entropy production bound that will be the key ingredient for local equilibrium that is not already established somewhere else. First, we introduce notation to be used both in this section and beyond.

**Definition 3.4.** Provided any $T \in \mathbb{R}_{\geqslant 0}$, let us define $\mu_{T,N}$ as the law of the particle system on $\Omega_{\mathbb{Z}}$ after time-$T$ evolution with initial measure $\mu_{0,N}$. Define the Radon-Nikodym derivative $\mathfrak{f}_{T,N} \mathrm{d}\mu_{0,\mathbb{Z}} = \mathrm{d}\mu_{T,N}$.

**Definition 3.5.** For any random variable or probability measure $\mathfrak{V}$, define $\mathfrak{V}^{\mathfrak{I}} \overset{\bullet}{=} \mathbf{E}_{\mathfrak{I}}^{\mu_{0,\mathbb{Z}}} \mathfrak{V}$ as the conditional expectation conditioning on $\eta_x$ for $x \in \mathfrak{I} \subseteq \mathbb{Z}$. If $\mathfrak{V}$ is a measure, this conditional expectation is its marginal on the joint spins $\eta_x$ for $x \in \mathfrak{I} \subseteq \mathbb{Z}$.

**Proposition 3.6.** *Take any $T \in \mathbb{R}_{\geqslant 0}$ uniformly bounded, and consider any arbitrarily small though universal constant $\varepsilon \in \mathbb{R}_{>0}$. Uniformly over all initial probability measures $\mu_{0,N}$ on $\Omega_{\mathbb{Z}}$, we get the time-averaged grand-canonical Dirichlet form bound*

$$T^{-1} \int_0^T \mathfrak{D}_{0,\mathfrak{I}_N}(\mathfrak{f}_{S,N}^{\mathfrak{I}_N}) \, \mathrm{d}S \lesssim_\varepsilon T^{-1} N^{-\frac{3}{4}+5\varepsilon} + N^{-\frac{3}{4}+5\varepsilon} \lesssim T^{-1} N^{-\frac{3}{4}+5\varepsilon}. \qquad (3.4)$$

*We introduced $\mathfrak{I}_N \overset{\bullet}{=} [\![-N^{\frac{5}{4}+\varepsilon}, N^{\frac{5}{4}+\varepsilon}]\!] \subseteq \mathbb{Z}$. The second bound in the previous display follows as $T \in \mathbb{R}_{\geqslant 0}$ is uniformly bounded.*

Proposition 3.6 is referred to as an "entropy production bound" even though it estimates a time-averaged Dirichlet form. This is because the upper bound comes from estimating terms appearing in the time derivative, or growth, of relative entropy.

Without the $\varepsilon$-dependent terms in the exponents in the statement of Proposition 3.6, the conclusion is entropy production estimates that would match that of a *symmetric* simple exclusion process on a torus of length $N^{5/4}$ roughly speaking. Thus one interpretation of Proposition 3.6 is somehow controlling the contribution of the asymmetry on the length-scale $N^{5/4}$ by that of the symmetric part of the exclusion process. We note that doing so upon replacing the length-scale $N^{5/4} \to N^{3/2}$ is not difficult as we show in Lemma 3.15 for reasons which amount to $\frac{3}{2} = 2 - \frac{1}{2}$ and the relative strength of the asymmetry is $N^{-1/2}$.

*3.2. Local equilibrium.* First, we introduce notation that will be used both in and beyond the current section. The following construction is of a space-time averaged probability density in spirit of the one-block scheme in [14].

**Definition 3.7.** Recall the law of the particle system $\mu_{\bullet,N}$ with initial measure $\mu_{0,N}$. Recall $\mathfrak{f}_{\bullet,N} \mathrm{d}\mu_{0,\mathbb{Z}} \overset{\bullet}{=} \mathrm{d}\mu_{\bullet,N}$. Provided times $\mathfrak{t}_0, T \in \mathbb{R}_{\geqslant 0}$ and a subset $\mathfrak{I} \subseteq \mathbb{Z}$, define the following space-time averaged probability density where $\tau_x \mathfrak{f}_{\mathfrak{t},N}(\eta) \overset{\bullet}{=} \mathfrak{f}_{\mathfrak{t},N}(\tau_x \eta)$:

$$\bar{\mathfrak{f}}_{\mathfrak{t}_0,T,\mathfrak{I}} \overset{\bullet}{=} T^{-1} \int_0^T \widetilde{\sum_{x \in \mathfrak{I}}} \tau_x \mathfrak{f}_{\mathfrak{t}_0+S,N} \, \mathrm{d}S. \qquad (3.5)$$

We also define $\mathrm{d}\bar{\mu}_{\mathfrak{t}_0,T,\mathfrak{I}} \overset{\bullet}{=} \bar{\mathfrak{f}}_{\mathfrak{t}_0,T,\mathfrak{I}} \mathrm{d}\mu_{0,\mathbb{Z}}$ and $\mathrm{d}\bar{\mu}_{T,\mathfrak{I}} \overset{\bullet}{=} \mathrm{d}\bar{\mu}_{0,T,\mathfrak{I}}$ and $\bar{\mathfrak{f}}_{T,\mathfrak{I}} \overset{\bullet}{=} \bar{\mathfrak{f}}_{0,T,\mathfrak{I}}$. Provided any subset $\mathfrak{I}_2 \subseteq \mathbb{Z}$ and any $\varrho \in \mathbb{R}$, define "spin-density" probabilities and the "$\mathfrak{I}_2$-local" relative entropy and Dirichlet form that we use in proving local equilibrium:

$$\mathfrak{p}_{\mathfrak{t}_0,T,\mathfrak{I}}^{\mathfrak{I}_2,\varrho} \doteq \mathbf{E}^{\mu_{0,\mathfrak{I}_2}}\left(\bar{\mathfrak{f}}_{\mathfrak{t}_0,T,\mathfrak{I}}^{\mathfrak{I}_2}\mathbf{1}_{\widetilde{\sum}_{x\in\mathfrak{I}_2}\eta_x=\varrho}\right) \tag{3.6a}$$

$$\bar{\mathfrak{H}}_{\varrho,\mathfrak{I}_2}^{\mathrm{can}}\left(\bar{\mathfrak{f}}_{\mathfrak{t}_0,T,\mathfrak{I}}^{\mathfrak{I}_2}\right) \doteq \mathfrak{H}_{\varrho,\mathfrak{I}_2}^{\mathrm{can}}\left(\frac{\mathrm{d}\bar{\mu}_{\mathfrak{t}_0,T,\mathfrak{I}}^{\mathfrak{I}_2}}{\mathrm{d}\mu_{\varrho,\mathfrak{I}_2}^{\mathrm{can}}}\left(\mathfrak{p}_{\mathfrak{t}_0,T,\mathfrak{I}}^{\mathfrak{I}_2,\varrho}\right)^{-1}\mathbf{1}_{\widetilde{\sum}_{x\in\mathfrak{I}_2}\eta_x=\varrho}\right) \tag{3.6b}$$

$$\bar{\mathfrak{D}}_{\varrho,\mathfrak{I}_2}^{\mathrm{can}}\left(\bar{\mathfrak{f}}_{\mathfrak{t}_0,T,\mathfrak{I}}^{\mathfrak{I}_2}\right) \doteq \mathfrak{D}_{\varrho,\mathfrak{I}_2}^{\mathrm{can}}\left(\frac{\mathrm{d}\bar{\mu}_{\mathfrak{t}_0,T,\mathfrak{I}}^{\mathfrak{I}_2}}{\mathrm{d}\mu_{\varrho,\mathfrak{I}_2}^{\mathrm{can}}}\left(\mathfrak{p}_{\mathfrak{t}_0,T,\mathfrak{I}}^{\mathfrak{I}_2,\varrho}\right)^{-1}\mathbf{1}_{\widetilde{\sum}_{x\in\mathfrak{I}_2}\eta_x=\varrho}\right). \tag{3.6c}$$

The first term is the probability under $\mathrm{d}\bar{\mu}_{\mathfrak{t}_0,T,\mathfrak{I}}$ of the hyperplane in $\Omega_{\mathfrak{I}_2}$ with spin density $\varrho \in \mathbb{R}$. The second term is the relative entropy of $\mathrm{d}\bar{\mu}_{\mathfrak{t}_0,T,\mathfrak{I}}$ with respect to a canonical ensemble after projecting to $\mathfrak{I}_2$-marginals and conditioning on the hyperplane with appropriate spin density $\varrho \in \mathbb{R}$. The third term is the same as the second but Dirichlet form instead of relative entropy.

**Lemma 3.8.** *Suppose the functional $\varphi$ has support $\mathfrak{I}_\varphi \subseteq \mathfrak{I}_N$ where $\mathfrak{I}_N \subseteq \mathbb{Z}$ is the block defined above in* Proposition *3.6 but with possibly different but still arbitrarily small and universal exponent $\varepsilon \in \mathbb{R}_{>0}$. Given any $\kappa \in \mathbb{R}_{>0}$ and again for $\varepsilon \in \mathbb{R}_{>0}$ arbitrarily small but universal, we have the following "reduction to equilibrium" on $\mathfrak{I}_\varphi$ uniformly over uniformly bounded $\mathfrak{t}_0, T \in \mathbb{R}_{\geqslant 0}$:*

$$\mathbf{E}^{\mu_{0,\mathbb{Z}}}\left(\bar{\mathfrak{f}}_{\mathfrak{t}_0,T,\mathfrak{I}_N}^{\mathfrak{I}_\varphi} \cdot \varphi\right) \lesssim_\varepsilon T^{-1}\kappa^{-1}N^{-\frac{3}{4}+10\varepsilon}|\mathfrak{I}_N|^{-1}|\mathfrak{I}_\varphi|^3 + \kappa^{-1}\sup_{\varrho\in\mathbb{R}}\log\mathbf{E}^{\mu_{\varrho,\mathfrak{I}_\varphi}^{\mathrm{can}}}\mathbf{Exp}\left(\kappa\varphi\right). \tag{3.7}$$

We provide the proof of Lemma 3.8 towards the end of the current section as with the proof of Proposition 3.6. The proof of Lemma 3.8 is an application of the classical entropy inequality in Appendix 1.8 of [21], the log-Sobolev inequality of [25], and the entropy production bounds in Proposition 3.6. We point out that the dependence on the support of the functional $\varphi$ we take expectation of in the statement of Lemma 3.8 is cubic on the RHS of the estimate therein. This reflects additional difficulty in comparison with local equilibrium on bigger length-scales. We also point out that if we knew $|\varphi| \lesssim \kappa^{-1}$ with universal implied constant at least at a level of large deviations, the second term on the RHS of the estimate in the statement of Lemma 3.8 could be reduced to just an expectation of $|\varphi|$ times universal factors courtesy of uniform sublinearity of the exponential on compact sets and concavity of the logarithm. We use this feature of the estimate in the statement of Lemma 3.8 later in Sect. 5.

*3.3. Invariant measure estimates.* We begin this discussion with the following local construction.

**Definition 3.9.** Consider a subset $\mathfrak{I} \subseteq \mathbb{Z}$ and define the following infinitesimal generator in which $+_{\mathfrak{I}}$ denotes addition on $\mathfrak{I}$ equipped with periodic boundary conditions inf $\mathfrak{I} = $ sup $\mathfrak{I} + 1$ and thus realized as a torus:

$$\mathfrak{S}_{\mathfrak{I}}^{N,!!}\mathfrak{f}(\eta) \doteq N^2 \sum_{k=1}^{\mathfrak{m}}\sum_{x\in\mathfrak{I}}\alpha_k\left(1 + \frac{1}{2}N^{-\frac{1}{2}}\gamma_k\frac{1-\eta_x}{2}\frac{1+\eta_{x+_{\mathfrak{I}}k}}{2}\right.$$
$$\left. - \frac{1}{2}N^{-\frac{1}{2}}\gamma_k\frac{1+\eta_x}{2}\frac{1-\eta_{x+_{\mathfrak{I}}k}}{2}\right)\mathfrak{S}_{x,x+_{\mathfrak{I}}k}\mathfrak{f}(\eta). \tag{3.8}$$

We denote by $\eta_{\mathfrak{I};\bullet}$ the $\mathfrak{I}$-local/$\mathfrak{I}$-periodic stochastic process valued in $\Omega_{\mathfrak{I}}$ with infinitesimal generator $\mathfrak{S}_{\mathfrak{I}}^{N,!!}$. It is an exclusion process of particles that jump by the same local rules as the original exclusion process but with periodic boundary on the set $\mathfrak{I}$.

We first give a general Kipnis-Varadhan inequality to control time-averages for stationary exclusion processes by a Sobolev norm. We then control the Sobolev norm. Lastly, we present martingale LDP-type bounds to study spatial mixing of canonical ensembles in the context of pseudo-gradients. We note that these results are all basically classical.

**Lemma 3.10.** *Consider any* $\mathfrak{I} \subseteq \mathbb{Z}$ *and* $\varrho \in \mathbb{R}$. *For any* $\varphi : \Omega_{\mathfrak{I}} \to \mathbb{R}$ *and* $\mathfrak{t}_{av} \in \mathbb{R}_{\geqslant 0}$, *we have the following that we explain after:*

$$\mathbf{E}^{\mu_{\varrho,\mathfrak{I}}^{can}} \left| \mathscr{A}_{\mathfrak{I};\mathfrak{t}_{av}}^{\mathbf{T},+}(\varphi(\eta_{\mathfrak{I};\bullet})) \right|^2 + \mathbf{E}^{\mu_{\varrho,\mathfrak{I}}^{can}} \sup_{0 \leqslant \mathfrak{t} \leqslant \mathfrak{t}_{av}} \mathfrak{t}^2 \mathfrak{t}_{av}^{-2} |\mathscr{A}_{\mathfrak{I};\mathfrak{t}}^{\mathbf{T},+}(\varphi(\eta_{\mathfrak{I};\bullet}))|^2 \lesssim \mathfrak{t}_{av}^{-1} \|\varphi\|_{\dot{\mathbf{H}}_{\varrho,\mathfrak{I}}^{-1}}^2. \tag{3.9}$$

*Above, we have introduced the following time-averages and Sobolev norms adapted to the* $\mathfrak{I}$-*local stochastic process:*

- *Define the following time/dynamic-average operator acting on* $\varphi$ *with initial time* $\bullet \in \mathbb{R}_{\geqslant 0}$ *and time-scale* $\mathfrak{t} \in \mathbb{R}_{\geqslant 0}$:

$$\mathscr{A}_{\mathfrak{I};\mathfrak{t}}^{\mathbf{T},+}(\varphi(\eta_{\mathfrak{I};\bullet})) \doteq \mathbf{1}_{\mathfrak{t} \neq 0} \mathfrak{t}^{-1} \int_0^{\mathfrak{t}} \varphi(\eta_{\mathfrak{I};\bullet+\mathfrak{r}}) \, d\mathfrak{r} + \mathbf{1}_{\mathfrak{t}=0} \varphi(\eta_{\mathfrak{I};\bullet}). \tag{3.10}$$

  *We emphasize the superscript "$\mathbf{T},+$" denotes time-average in the positive time-direction. We will use and introduce this notation again later throughout the paper; see* Sect. *5, for example. We also clarify that the maximal process on the LHS of the proposed estimate is actually the supremum over time-scales* $0 \leqslant \mathfrak{t} \leqslant \mathfrak{t}_{av}$ *of the integral of* $\varphi(\eta_{\mathfrak{I};\bullet})$ *on the interval* $[0, \mathfrak{t}]$ *then weighted by* $\mathfrak{t}_{av}^{-1}$, *and then squaring this integral and* $\mathfrak{t}_{av}$-*weight. In particular, we only get the square of a time-average if we choose* $\mathfrak{t} = \mathfrak{t}_{av}$.
- *The Sobolev norm* $\dot{\mathbf{H}}^{-1}$ *is defined by the following, where* $\bar{\mathfrak{S}}_{\mathfrak{I}}^{N,!!}$ *denotes* $\mathfrak{S}_{\mathfrak{I}}^{N,!!}$ *without* $\mathfrak{S}_{x,x+\mathfrak{I}k}$-*terms for which* $x +_{\mathfrak{I}} k \neq x + k$ *and also sets* $\gamma_k = 0$, *so* $\bar{\mathfrak{S}}_{\mathfrak{I}}^{N,!!}$ *is the infinitesimal generator for a finite-range symmetric exclusion process on* $\mathfrak{I} \subseteq \mathbb{Z}$ *suppressing all attempted jumps outside the block* $\mathfrak{I} \subseteq \mathbb{Z}$. *We note that this finite-range symmetric exclusion process is also of order* $N^2$ *speed:*

$$\|\varphi\|_{\dot{\mathbf{H}}_{\varrho,\mathfrak{I}}^{-1}}^2 \doteq \sup_{\psi:\Omega_{\mathfrak{I}} \to \mathbb{R}} \left( 2\mathbf{E}^{\mu_{\varrho,\mathfrak{I}}^{can}}(\varphi\psi) + \mathbf{E}^{\mu_{\varrho,\mathfrak{I}}^{can}} \psi \bar{\mathfrak{S}}_{\mathfrak{I}}^{N,!!} \psi \right). \tag{3.11}$$

- *The expectation* $\mathbf{E}^{\mu_{\varrho,\mathfrak{I}}^{can}}$ *of the "path-space" functional* $|\mathscr{A}_{\mathfrak{I};\mathfrak{t}_{av}}^{\mathbf{T},+} \varphi(\eta_{\mathfrak{I};\bullet})|^2$ *and its running supremum is equal to* $\mathbf{E}^{\mu_{\varrho,\mathfrak{I}}^{can}} \mathbf{E}_{\eta_{\mathfrak{I};\bullet}}$, *where*
  – *The inner expectation is with respect to the path-space measure induced by the* $\mathfrak{S}_{\mathfrak{I}}^{N,!!}$-*dynamic with initial spins* $\eta_{\mathfrak{I};\bullet} \in \Omega_{\mathfrak{I}}$.
  – *The outer expectation is over the above initial spin configuration* $\eta_{\mathfrak{I};\bullet} \in \Omega_{\mathfrak{I}}$ *with respect to the canonical ensemble* $\mu_{\varrho,\mathfrak{I}}^{can}$.

*Proof.* The first term in the claimed bound is controlled by the second by taking the latter supremum at $\mathfrak{t} = \mathfrak{t}_{av}$. We note $\mu_{\varrho,\mathfrak{I}}^{can}$ is an invariant measure for the $\eta_{\mathfrak{I}}$-process, which can be checked by standard procedure, and we cite Lemma 2.4 in [20].                                    □

To use Lemma 3.10 we use the next set of estimates. The first is an orthogonality estimate that is certainly true for variance and comes by the spatial mixing of the canonical ensemble. The point is that orthogonality holds for the Sobolev norm too. The second

estimate is a spectral gap. It estimates the Sobolev norm by a variance with constant adapted to diffusive speed-scaling in the process and the support of the functional at hand. See Proposition 3.3 and Proposition 3.4 in [12] for a proof.

**Lemma 3.11.** *Consider any* $\mathfrak{I} \subseteq \mathbb{Z}$ *and* $\varrho \in \mathbb{R}$. *Suppose we have a collection of bounded functions* $\varphi_1, \ldots, \varphi_J : \Omega_{\mathfrak{I}} \to \mathbb{R}$ *so that*

- *The functional* $\varphi_j$ *has support* $\mathfrak{I}_j \subseteq \mathfrak{I}$, *and the supports* $\mathfrak{I}_1, \ldots, \mathfrak{I}_J \subseteq \mathfrak{I}$ *are mutually disjoint.*
- *Provided any* $j \in [\![1, J]\!]$ *and any* $\widetilde{\varrho} \in \mathbb{R}$, *we have the canonical ensemble mean zero condition* $\mathbf{E}^{\mu_{\widetilde{\varrho}, \mathfrak{I}_j}^{\mathrm{can}}} \varphi_j = 0$.

*We define the average* $\Phi_{J,N} \overset{\bullet}{=} \widetilde{\sum}_{j=1}^{J} \varphi_j$. *We recall* $\widetilde{\sum}$ *denotes an averaged sum. We have*

$$\sup_{\varrho \in \mathbb{R}} \|\Phi_{J,N}\|_{\dot{\mathbf{H}}_{\varrho, \mathfrak{I}}^{-1}}^2 \lesssim J^{-1} \sup_{j=1,\ldots,J} \sup_{\varrho_j \in \mathbb{R}} \|\varphi_j\|_{\dot{\mathbf{H}}_{\varrho_j, \mathfrak{I}_j}^{-1}}^2 \tag{3.12}$$

$$\sup_{\varrho_j \in \mathbb{R}} \|\varphi_j\|_{\dot{\mathbf{H}}_{\varrho_j, \mathfrak{I}_j}^{-1}}^2 \lesssim N^{-2} |\mathfrak{I}_j|^2 \sup_{\varrho_j \in \mathbb{R}} \mathbf{E}^{\mu_{\varrho_j, \mathfrak{I}_j}^{\mathrm{can}}} |\varphi_j|^2. \tag{3.13}$$

The orthogonality bound will be useful because we will take $\varphi_j = \tau_{-7j\mathfrak{m}} \mathfrak{g}$ in a future application of the above result; recall in Proposition 2.7 that these $\varphi_j$-terms have disjoint and uniformly bounded support. Let us additionally point out the double supremum on the RHS of the first estimate in the statement of Lemma 3.11. Orthogonality estimates would generally have an average over $j \in [\![1, J]\!]$ instead of the supremum over these indices. Proposition 3.3 and Proposition 3.4 in [12] actually allow for an average instead of a supremum on the RHS of the first estimate from the statement of Lemma 3.11. However, we will not need an average, and the average certainly implies an upper bound in terms of the supremum as well. The $N$-dependent factor in the second bound of Lemma 3.11 reflects diffusive scaling of the particle system and speed to invariant measure.

Our last Sobolev bound treats pseudo-gradients of large supports but whose pseudo-gradient factors have small supports. Lemma 3.11 is suboptimal in the size of the support of the underlying functional as its fluctuating factor, which is responsible for the estimate even being possible, may have small support.

**Lemma 3.12.** *If* $\varphi : \Omega_{\mathbb{Z}} \to \mathbb{R}$ *has support* $\mathfrak{I}_{\varphi} \subseteq \mathbb{Z}$ *and admits pseudo-gradient factor* $\mathfrak{g} : \Omega_{\mathbb{Z}} \to \mathbb{R}$ *with support* $\mathfrak{I} \subseteq \mathbb{Z}$, *then the Sobolev norm of the functional* $\varphi$ *is controlled by that of its pseudo-gradient factor. Precisely, with a universal implied constant*

$$\sup_{\varrho \in \mathbb{R}} \|\varphi\|_{\dot{\mathbf{H}}_{\varrho, \mathfrak{I}_{\varphi}}^{-1}} \lesssim \sup_{\varrho \in \mathbb{R}} \|\mathfrak{g}\|_{\dot{\mathbf{H}}_{\varrho, \mathfrak{I}}^{-1}}. \tag{3.14}$$

*Proof.* First observe that it suffices to assume $\varphi \overset{\bullet}{=} \mathfrak{g} \cdot \mathfrak{f}$ with $\mathfrak{f} = \prod_{i=1}^{\ell} \eta_{x_i}$. Indeed, we generally have $\varphi = \mathfrak{g} \cdot \mathfrak{f}$ for some local functional $\mathfrak{f} : \Omega_{\mathbb{Z}} \to \mathbb{R}$ that is an average of such products times uniformly bounded and deterministic constants. It then suffices to apply the triangle inequality for the Sobolev norm which is elementary to deduce via its variational definition. We also note that $\mathfrak{I} \subseteq \mathbb{Z}$ is disjoint from the support $\{x_i\}_{i=1}^{\ell}$ of $\mathfrak{f}$ by definition of admitting a pseudo-gradient factor. We start as follows:

- First, for any local $\psi : \Omega_{\mathbb{Z}} \to \mathbb{R}$, define $\hat{\psi} \overset{\bullet}{=} \psi \cdot \prod_{i=1}^{\ell} \eta_{x_i} = \psi \cdot \mathfrak{f}$, so in particular $\varphi = \hat{\mathfrak{g}}$ under our notation in this proof.

- We recall the local $\bar{\mathfrak{S}}^{N,!!}$-generators and Sobolev norms in Lemma 3.10. By negative semi-definiteness of $\mathfrak{S}_{x,y}$-generators, we have the following inequality of quadratic forms for any $\psi : \Omega_{\mathfrak{I}_\varphi} \to \mathbb{R}$; recall the containment $\mathfrak{I} \subseteq \mathfrak{I}_\varphi$:

$$\mathbf{E}^{\mu_{\varrho,\mathfrak{I}_\varphi}^{\mathrm{can}}} \psi \bar{\mathfrak{S}}^{N,!!}_{\mathfrak{I}_\varphi} \psi \ \leqslant \ \mathbf{E}^{\mu_{\varrho,\mathfrak{I}_\varphi}^{\mathrm{can}}} \psi \bar{\mathfrak{S}}^{N,!!}_{\mathfrak{I}} \psi \ \leqslant \ \mathbf{E}^{\mu_{\varrho,\mathfrak{I}_\varphi}^{\mathrm{can}}} \hat{\psi} \bar{\mathfrak{S}}^{N,!!}_{\mathfrak{I}} \hat{\psi}. \tag{3.15}$$

The last bound follows as $\bar{\mathfrak{S}}^{N,!!}_{\mathfrak{I}}$ does not act on spins outside $\mathfrak{I} \subseteq \mathbb{Z}$, as $\eta_x^2 = 1$ for all $x \in \mathbb{Z}$, and as $\{x_i\}_{i=1}^{\ell}$ is disjoint from $\mathfrak{I}$.

With the above notation and observation, we get the following with explanation given afterwards:

$$\begin{aligned} \|\varphi\|_{\dot{\mathbf{H}}^{-1}_{\varrho,\mathfrak{I}_\varphi}}^2 & \ \leqslant \ \sup_{\psi:\Omega_{\mathfrak{I}_\varphi}\to\mathbb{R}} \left( 2\mathbf{E}^{\mu_{\varrho,\mathfrak{I}_\varphi}^{\mathrm{can}}} (\varphi\psi) \ + \ \mathbf{E}^{\mu_{\varrho,\mathfrak{I}_\varphi}^{\mathrm{can}}} \psi \bar{\mathfrak{S}}^{N,!!}_{\mathfrak{I}_\varphi} \psi \right) \\ & \ \leqslant \ \sup_{\psi:\Omega_{\mathfrak{I}_\varphi}\to\mathbb{R}} \left( 2\mathbf{E}^{\mu_{\varrho,\mathfrak{I}_\varphi}^{\mathrm{can}}} (\mathfrak{g}\hat{\psi}) \ + \ \mathbf{E}^{\mu_{\varrho,\mathfrak{I}_\varphi}^{\mathrm{can}}} \hat{\psi} \bar{\mathfrak{S}}^{N,!!}_{\mathfrak{I}} \hat{\psi} \right). \end{aligned} \tag{3.16}$$

In the final bound, we note $\varphi\psi = \mathfrak{g}\hat{\psi}$ again because $\eta_x^2 = 1$. This last bound is controlled by $\|\mathfrak{g}\|^2$ via convexity of the Dirichlet form and replacing $\hat{\psi}$ on the far RHS of the above by its conditional expectation/projection on $\mathfrak{I} \subseteq \mathbb{Z}$. The result follows. $\qquad\square$

We now leverage spatial mixing of canonical ensembles in the context of pseudo-gradients supported on disjoint subsets to get LDP estimates for spatial-averages of pseudo-gradients. This will only be relevant for studying order $N^{1/2}$-terms in $\Phi^{N,2}$ in Proposition 2.7; recall there was no analog for order $N^{\beta_X}$-terms in $\Phi^{N,2}$ for Pseudo-Proposition 2.12.

**Lemma 3.13.** *Retain the setting of* Lemma *3.11, and additionally assume the following a priori estimate:*

- *We have the deterministic estimate $|\varphi_j| \lesssim 1$ uniformly in $\eta \in \Omega_{\mathbb{Z}}$ and $j \in [\![1, J]\!]$ with universal implied constant.*

*For any $\mathfrak{C} \in \mathbb{R}_{>0}$ and $\varrho \in \mathbb{R}$, we have the following sub-Gaussian tail bound with sufficiently large but universal implied constants:*

$$\mathbf{P}^{\mu_{\varrho,\mathfrak{I}}^{\mathrm{can}}}[|\Phi_{J,N}| \gtrsim \mathfrak{C}] \ \lesssim \ \mathbf{Exp}\left( -J\mathfrak{C}^2 \right). \tag{3.17}$$

*Proof.* Consider the filtration of $\sigma$-algebras $\sigma(\mathfrak{I}_1, \dots, \mathfrak{I}_j)$ for $j \in [\![1, J]\!]$. Under the canonical ensemble $\mu_{\varrho,\mathfrak{I}}^{\mathrm{can}}$ the $\varphi_j$-quantities are uniformly bounded martingale increments since the projection of any canonical ensemble onto a smaller block is a convex combination of canonical ensembles. We then use the Azuma martingale inequality. $\quad\square$

**Corollary 3.14.** *Retain the setting of* Lemma *3.13. For any $\varepsilon_1, \varepsilon_2 \in \mathbb{R}_{>0}$ and $C \in \mathbb{R}_{>0}$, with universal implied constants we have*

$$\mathbf{E}^{\mu_{\varrho,\mathfrak{I}}^{\mathrm{can}}} \mathbf{Exp}\left( N^{-\varepsilon_1} J |\Phi_{J,N}|^2 \right) \cdot \mathbf{1}\left( |\Phi_{J,N}| \gtrsim N^{\varepsilon_2} J^{-\frac{1}{2}} \right) \ \lesssim_{\varepsilon_1,\varepsilon_2,C} \ N^{-C}. \tag{3.18}$$

*Proof.* By Lemma 3.13 $\Phi_{J,N}$ has zero mean and is sub-Gaussian of variance of order $J^{-1}$, so it suffices to pretend $\Phi_{J,N}$ is indeed Gaussian with zero mean and variance of order $J^{-1}$. The result then follows by straightforward calculations for Gaussians. $\quad\square$

The rest of the section is dedicated to proof of Proposition 3.6 and Lemma 3.8. We invite the reader to skip to the beginning of the upcoming Sect. 4 as the proof of Proposition 3.6, especially, is quite technical.

*3.4. Proof of Proposition 3.6.* We first present the ingredients needed in the proof of Proposition 3.6 with brief descriptions of what they say and why they are true. We then combine these ingredients to get Proposition 3.6 and close with their proofs. For convenience, let us declare all expectations in this subsection are taken with respect to the grand-canonical measure $\mu_{0,\mathbb{Z}}$.

The first step we take for the proof of Proposition 3.6 is the following a priori estimate which serves as an easier but weaker version of Proposition 3.6. Roughly speaking, such a priori estimate reduces to an exclusion process on a torus of length-scale which is basically $N^{3/2}$. The explain the exponent $3/2$, the particles in the exclusion process perform symmetric random walks with speed $N^2$ and a speed $N^{3/2}$ asymmetry. Thus, we expect only information at length-scale $(N^2)^{1/2} + N^{3/2} \lesssim N^{3/2}$ to matter.

**Lemma 3.15.** *For $\varepsilon \in \mathbb{R}_{>0}$ arbitrarily small but universal, define $\mathfrak{I}_{N,+} \doteq [\![ -N^{\frac{3}{2}+\varepsilon},$ $N^{\frac{3}{2}+\varepsilon} ]\!] \subseteq \mathbb{Z}$. We also define the following:*

- *Consider the following perturbation $\widetilde{\mu}_{0,N}$ as a probability measure on the product $\Omega_{\mathbb{Z}} = \Omega_{\mathbb{Z}\setminus\mathfrak{I}_{N,+}} \times \Omega_{\mathfrak{I}_{N,+}}$. In words, the measure $\widetilde{\mu}_{0,N}$ is defined by taking the measure $\mu_{0,N}$, cutting it off outside $\mathfrak{I}_{N,+}$, and gluing the grand-canonical measure $\mu_{0,\mathbb{Z}}$ outside $\mathfrak{I}_{N,+}$:*

$$\mathrm{d}\widetilde{\mu}_{0,N} \doteq \mathrm{d}\mu_{0,\mathbb{Z}\setminus\mathfrak{I}_{N,+}} \otimes \mathrm{d}\mu_{0,N}^{\mathfrak{I}_{N,+}}. \tag{3.19}$$

- *Analogously define $\widetilde{\mu}_{T,N}$ as the measure/law after time-$T$ evolution with initial measure $\widetilde{\mu}_{0,N}$, and define $\widetilde{\mathfrak{f}}_{T,N}\mathrm{d}\mu_{0,\mathbb{Z}} \doteq \mathrm{d}\widetilde{\mu}_{T,N}$.*

*We first have the following a priori global entropy production estimate for the perturbation initial measure $\widetilde{\mu}_{0,N}$:*

$$\int_0^T \mathfrak{D}_{0,\mathbb{Z}}(\widetilde{\mathfrak{f}}_{S,N}) \, \mathrm{d}S \lesssim N^{-\frac{1}{2}+\varepsilon}. \tag{3.20}$$

*We have comparison of Dirichlet forms on $\mathfrak{I}_N \doteq [\![ -N^{\frac{5}{4}+\varepsilon}, N^{\frac{5}{4}+\varepsilon} ]\!]$ in* Proposition 3.6 *for any $C \in \mathbb{R}_{>0}$ uniformly over $0 \leqslant T \leqslant 100$:*

$$|\mathfrak{D}_{0,\mathfrak{I}_N}(\mathfrak{f}_{T,N}^{\mathfrak{I}_N}) - \mathfrak{D}_{0,\mathfrak{I}_N}(\widetilde{\mathfrak{f}}_{T,N}^{\mathfrak{I}_N})| \lesssim_{C,\varepsilon} N^{-C}. \tag{3.21}$$

The second bound (3.21) is the assertion that information beyond the aforementioned length-scale $N^{3/2}$, up to $N^\varepsilon$-factors, will not matter for Dirichlet forms. The first estimate above exploits the "stationary cutoff" beyond the length-scale $N^{3/2+\varepsilon}$.

We now implement the multiscale strategy in the proofs of Lemma 4.1 in [10] and of Theorem 2.1 in [22]. It relates entropy production on a block to that on a slightly-larger block. To explain this in more detail, we require the following constructions.

**Definition 3.16.** During the proof of Proposition 3.6 given in this subsection, we use the following constructions.

- Provided any $\mathfrak{l} \in \mathbb{Z}_{\geqslant 0}$, define $e_{N,\mathfrak{l}} \doteq \mathbf{Exp}(-N^{-5/4-\varepsilon}\mathfrak{l})$ and $\mathfrak{I}_{N,\mathfrak{l}} \doteq [\![ -N^{\frac{5}{4}+\varepsilon} - \mathfrak{l} \cdot N^\varepsilon, N^{\frac{5}{4}+\varepsilon} + \mathfrak{l} \cdot N^\varepsilon ]\!] \subseteq \mathbb{Z}$. This family of subsets is a progressive "fattening" of the original $\mathfrak{I}_N$-subset in Proposition 3.6. The weights $e_{N,\mathfrak{l}}$ will be pieces of a multiscale analysis that controls entropy production on the block $\mathfrak{I}_{N,\mathfrak{l}}$ by that on $\mathfrak{I}_{N,\mathfrak{l}+j}$; see proofs of Lemma 4.1 in [10] and Theorem 2.1 in [22].

- For any $x, y \in \mathbb{Z}$, define the projected densities $\mathfrak{f}_{\mathfrak{l}} \doteq \mathfrak{f}_{T,N}^{\mathfrak{I}_{N,\mathfrak{l}}}$ and $\mathfrak{f}_{\mathfrak{l},x,y} \doteq \mathfrak{f}_{T,N}^{\mathfrak{I}_{N,\mathfrak{l}} \cup \{x,y\}}$. The former has entropy production related to Dirichlet forms of the latter. The latter is not in the proof of Lemma 4.1 in [10], but it is in the proof of Theorem 2.1 in [22].
- Define the boundary $\partial \mathfrak{I}_{N,\mathfrak{l}} \doteq \{(x, y) \in \mathbb{Z} : x \in \mathfrak{I}_{N,\mathfrak{l}}, y \notin \mathfrak{I}_{N,\mathfrak{l}}, x < y\} \cup \{(x, y) \in \mathbb{Z} : x \notin \mathfrak{I}_{N,\mathfrak{l}}, y \in \mathfrak{I}_{N,\mathfrak{l}}, x < y\}$. Note that pairs of points in the boundary $\partial \mathfrak{I}_{N,\mathfrak{l}}$ do *not* have to be nearest neighbors. Roughly, entropy production on $\mathfrak{I}_{N,\mathfrak{l}}$ forces us to look at the boundary $\partial \mathfrak{I}_{N,\mathfrak{l}}$ as particles can jump in/out of $\mathfrak{I}_{N,\mathfrak{l}}$; see proofs of Theorem 2.1 in [22] and Lemma 4.1 of [10].

As mentioned earlier, we relate entropy production on block $\mathfrak{I}_{N,\mathfrak{l}}$ and entropy production on a small neighborhood $\mathfrak{I}_{N,\mathfrak{l}+\mathfrak{j}}$. This is the lemma below, which introduces boundary terms $\Phi_{\mathrm{Sym}}, \Phi_{\wedge,1}, \Phi_{\wedge,2}$ that we treat with another pair of lemmas.

**Lemma 3.17.** *Provided any $\mathfrak{l} \in \mathbb{Z}_{\geqslant 1}$, we have the following entropy production differential inequality:*

$$\partial_T \mathfrak{H}_{0,\mathfrak{I}_{N,\mathfrak{l}}}(\mathfrak{f}_{\mathfrak{l}}) + 2N^2 \mathfrak{D}_{0,\mathfrak{I}_{N,\mathfrak{l}}}(\mathfrak{f}_{\mathfrak{l}}) \leqslant \Phi_{\mathrm{Sym}} + \Phi_{\wedge,1} + \Phi_{\wedge,2}. \qquad (3.22)$$

*The RHS comes from jumps at $\partial \mathfrak{I}_{N,\mathfrak{l}}$. In particular $\Phi_{\mathrm{Sym}}$ comes from jumps at $\partial \mathfrak{I}_{N,\mathfrak{l}}$ in the symmetric part of the process and $\Phi_{\wedge,1}, \Phi_{\wedge,2}$ come from asymmetric such jumps. Recall $\mathfrak{S}_{x,y}$ is the generator of the speed 1 symmetric exclusion process on $\{x, y\} \subseteq \mathbb{Z}$:*

$$\Phi_{\mathrm{Sym}} \doteq 2N^2 \mathbf{E} \sum_{(x,y) \in \partial \mathfrak{I}_{N,\mathfrak{l}}} \alpha_{|y-x|} \left( (\mathfrak{f}_{\mathfrak{l},x,y} - \mathfrak{f}_{\mathfrak{l}}) \mathfrak{f}_{\mathfrak{l}}^{-1/2} \mathfrak{S}_{x,y}(\mathfrak{f}_{\mathfrak{l}}^{1/2}) \right) \qquad (3.23a)$$

$$\Phi_{\wedge,1} \doteq N^{\frac{3}{2}} \mathbf{E} \sum_{(x,y) \in \partial \mathfrak{I}_{N,\mathfrak{l}}} \mathbf{1}_{x \leqslant y} \alpha_{|y-x|} \gamma_{|y-x|} \mathbf{1}_{\eta_x=-1} \mathbf{1}_{\eta_y=1} \left( (\mathfrak{f}_{\mathfrak{l},x,y} - \mathfrak{f}_{\mathfrak{l}}) \mathfrak{f}_{\mathfrak{l}}^{-1/2} \mathfrak{S}_{x,y}(\mathfrak{f}_{\mathfrak{l}}^{1/2}) \right) \qquad (3.23b)$$

$$\Phi_{\wedge,2} \doteq -N^{\frac{3}{2}} \mathbf{E} \sum_{(x,y) \in \partial \mathfrak{I}_{N,\mathfrak{l}}} \mathbf{1}_{x \leqslant y} \alpha_{|y-x|} \gamma_{|y-x|} \mathbf{1}_{\eta_x=1} \mathbf{1}_{\eta_y=-1} \left( (\mathfrak{f}_{\mathfrak{l},x,y} - \mathfrak{f}_{\mathfrak{l}}) \mathfrak{f}_{\mathfrak{l}}^{-1/2} \mathfrak{S}_{x,y}(\mathfrak{f}_{\mathfrak{l}}^{1/2}) \right). \qquad (3.23c)$$

**Lemma 3.18.** *Retaining the context of the current subsection, with a universal implied constant we have the following estimate which controls boundary-terms in $\Phi_{\mathrm{Sym}}$ coming from symmetric jumps in the particle system in terms of data on larger $\mathfrak{I}_{N,\mathfrak{l}+\mathfrak{j}}$ blocks:*

$$|\Phi_{\mathrm{Sym}}| \lesssim N^{\frac{3}{2}} \sum_{(x,y) \in \partial \mathfrak{I}_{N,\mathfrak{l}}} \alpha_{|x-y|} \sum_{\mathfrak{j}=1}^{N^{5/4}} \mathbf{E} |\mathfrak{S}_{x,y} \left( \mathfrak{f}_{\mathfrak{l}+\mathfrak{j}}^{\frac{1}{2}} \right)|^2 + N^{\frac{5}{4}}. \qquad (3.24)$$

**Lemma 3.19.** *Retaining the context of the current subsection, we have $|\Phi_{\wedge,1}| + |\Phi_{\wedge,2}| \lesssim N^{3/2} (\mathfrak{D}_{0,\mathbb{Z}}(\mathfrak{f}_{T,N}))^{1/2}$.*

*Proof of Proposition 3.6.* By the second estimate in Lemma 3.15, we may take $\mathfrak{f} = \widetilde{\mathfrak{f}}$ with the notation therein. For convenience we take $T = 1$. The argument for general uniformly bounded $T \in \mathbb{R}_{\geqslant 0}$ follows identically. Lemmas 3.17, 3.18, and 3.19 give

$$\partial_T \mathfrak{H}_{0,\mathfrak{I}_{N,\mathfrak{l}}}(\mathfrak{f}_{\mathfrak{l}}) + 2N^2 \mathfrak{D}_{0,\mathfrak{I}_{N,\mathfrak{l}}}(\mathfrak{f}_{\mathfrak{l}})$$
$$\lesssim N^{\frac{3}{2}} \sum_{(x,y) \in \partial \mathfrak{I}_{N,\mathfrak{l}}} \alpha_{|x-y|} \sum_{\mathfrak{j}=1}^{N^{5/4}} \mathbf{E} |\mathfrak{S}_{x,y} \left( \mathfrak{f}_{\mathfrak{l}+\mathfrak{j}}^{\frac{1}{2}} \right)|^2 + N^{\frac{5}{4}} + N^{\frac{3}{2}} (\mathfrak{D}_{0,\mathbb{Z}}(\mathfrak{f}_{T,N}))^{\frac{1}{2}}. \qquad (3.25)$$

We integrate (3.25) in time while recalling the "worst-case" relative entropy bound $\mathfrak{H}_{0,\mathfrak{I}_{N,\mathfrak{l}}}(\mathfrak{f}_{\mathfrak{l}}) \lesssim |\mathfrak{I}_{N,\mathfrak{l}}| \lesssim N^{5/4+\varepsilon} + \mathfrak{l} \cdot N^{\varepsilon}$ to get the following integrated estimate, from which the second line follows from the Cauchy-Schwarz inequality applied to the last integral on the RHS of the first line. The last line follows by the global bound/the first estimate of Lemma 3.15:

$$\int_0^1 \mathfrak{D}_{0,\mathfrak{I}_{N,\mathfrak{l}}}(\mathfrak{f}_{\mathfrak{l}}) \, dT$$

$$\lesssim N^{-\frac{3}{4}+\varepsilon} + \mathfrak{l} \cdot N^{-2+\varepsilon} + \int_0^1 \left( N^{-\frac{1}{2}} \sum_{(x,y)\in\partial\mathfrak{I}_{N,\mathfrak{l}}} \alpha_{|x-y|} \sum_{j=1}^{N^{5/4}} \mathbf{E}|\mathfrak{S}_{x,y}\left(\mathfrak{f}_{\mathfrak{l}+j}^{\frac{1}{2}}\right)|^2 \right) dT$$

$$+ N^{-\frac{1}{2}} \int_0^1 (\mathfrak{D}_{0,\mathbb{Z}}(\mathfrak{f}_{T,N}))^{\frac{1}{2}} \, dT$$

$$\lesssim N^{-\frac{3}{4}+\varepsilon} + \mathfrak{l} \cdot N^{-2+\varepsilon} + \int_0^1 \left( N^{-\frac{1}{2}} \sum_{(x,y)\in\partial\mathfrak{I}_{N,\mathfrak{l}}} \alpha_{|x-y|} \sum_{j=1}^{N^{5/4}} \mathbf{E}|\mathfrak{S}_{x,y}\left(\mathfrak{f}_{\mathfrak{l}+j}^{\frac{1}{2}}\right)|^2 \right) dT$$

$$+ N^{-\frac{1}{2}} \left( \int_0^1 \mathfrak{D}_{0,\mathbb{Z}}(\mathfrak{f}_{T,N}) \, dT \right)^{\frac{1}{2}}$$

$$\lesssim N^{-\frac{3}{4}+\varepsilon} + \mathfrak{l} \cdot N^{-2+\varepsilon} + \int_0^1 \left( N^{-\frac{1}{2}} \sum_{(x,y)\in\partial\mathfrak{I}_{N,\mathfrak{l}}} \alpha_{|x-y|} \sum_{j=1}^{N^{5/4}} \mathbf{E}|\mathfrak{S}_{x,y}\left(\mathfrak{f}_{\mathfrak{l}+j}^{\frac{1}{2}}\right)|^2 \right) dT. \quad (3.26)$$

We average over all scales $\mathfrak{l} \in \mathbb{Z}_{\geqslant 1}$ against $e_{N,\mathfrak{l}}$-weights like the proofs of Theorem 2.1 in [22] and Lemma 4.1 in [10]. Because we average, the first two terms in (3.26) are okay. The proofs of Theorem 2.1 in [22] and Lemma 4.1 in [10] tell us the following for the third term in (3.26). This double sum is given by "local" Dirichlet forms/jumps at the boundary $\partial\mathfrak{I}_{N,\mathfrak{l}}$ of projections $\mathfrak{f}_{\mathfrak{l}+j}$ corresponding to $\mathfrak{I}_{N,\mathfrak{l}+j}$ subsets which are no more than slightly larger than the length-scale $|\mathfrak{I}_{N,\mathfrak{l}}|$. When we sum over all scales $\mathfrak{l} \in \mathbb{Z}_{\geqslant 1}$ as in proof of Theorem 2.1 in [22]/Lemma 4.1 in [10] against exponential weights $e_{N,\mathfrak{l}}$, the "local Dirichlet forms"/third term in (3.26) gets washed out by the $\mathfrak{D}_{0,\mathfrak{I}_{N,\mathfrak{l}+j}}(\mathfrak{f}_{\mathfrak{l}+j})$-terms from the $\mathfrak{I}_{N,\mathfrak{l}+j}$-equation ((3.25) for $\mathfrak{l} + j$ in place of $\mathfrak{l}$). Upon recalling the $e_{N,\mathfrak{l}}$-weights defined in Definition 3.16, this paragraph combined with the previous estimate (3.26) provides

$$N^{-\frac{5}{4}} \sum_{\mathfrak{l}=1}^{\infty} e_{N,\mathfrak{l}} \int_0^1 \mathfrak{D}_{0,\mathfrak{I}_{N,\mathfrak{l}}}(\mathfrak{f}_{\mathfrak{l}}) \, dT \lesssim N^{-\frac{5}{4}} \sum_{\mathfrak{l}=1}^{\infty} e_{N,\mathfrak{l}} \left( N^{-\frac{3}{4}+\varepsilon} + \mathfrak{l} \cdot N^{-2+\varepsilon} \right) \lesssim N^{-\frac{3}{4}+5\varepsilon}.$$

The last bound uses the elementary bounds $\sum_{\mathfrak{l}=1}^{\infty} e_{N,\mathfrak{l}} \lesssim N^{5/4+\varepsilon}$ and $\sum_{\mathfrak{l}=1}^{\infty} e_{N,\mathfrak{l}} \mathfrak{l} \lesssim N^{5/2+2\varepsilon}$. Like Lemma 4.1 in [10] and Theorem 2.1 in [22], this last estimate combined with convexity of the Dirichlet form completes the proof of Proposition 3.6. $\square$

It remains to get Lemma 3.15, Lemma 3.17, Lemma 3.18, and Lemma 3.19. We defer their proofs to the end of this section.

*3.5. Proof of Lemma 3.8.* We will first decompose the post-projection space-time averaged measure $\bar{\mu}_{\mathfrak{t}_0,T,\mathfrak{I}_N}^{\mathfrak{I}_\varphi}$ into conditional expectations onto the hyperplanes of Definition 3.1, write them as densities with respect to appropriate canonical ensembles, and apply the entropy inequality in Appendix 1.8 of [21] for constant $\kappa \in \mathbb{R}_{>0}$ for these

densities with respect to the canonical ensembles. This gives the following preliminary identity followed by an entropy inequality estimate:

$$
\mathbf{E}^{\mu_{0,\mathbb{Z}}}\left(\bar{\mathfrak{f}}_{\mathfrak{t}_0,T,\mathfrak{I}_N}^{\mathfrak{I}_\varphi}\cdot\varphi\right) \;=\; \mathbf{E}^{\bar{\mu}_{\mathfrak{t}_0,T,\mathfrak{I}_N}^{\mathfrak{I}_\varphi}}\varphi
$$

$$
= \sum_{\varrho\in\mathbb{R}}\mathfrak{p}_{\mathfrak{t}_0,T,\mathfrak{I}_N}^{\mathfrak{I}_\varphi,\varrho}\mathbf{E}^{\mu_{\varrho,\mathfrak{I}_\varphi}^{\mathrm{can}}}\left(\varphi\cdot\frac{\mathrm{d}\bar{\mu}_{\mathfrak{t}_0,T,\mathfrak{I}_N}^{\mathfrak{I}_\varphi}}{\mathrm{d}\mu_{\varrho,\mathfrak{I}_\varphi}^{\mathrm{can}}}[\mathfrak{p}_{\mathfrak{t}_0,T,\mathfrak{I}_N}^{\mathfrak{I}_\varphi,\varrho}]^{-1}\mathbf{1}_{\widetilde{\sum}_{x\in\mathfrak{I}_\varphi}\eta_x=\varrho}\right) \tag{3.27}
$$

$$
\lesssim \kappa^{-1}\sum_{\varrho\in\mathbb{R}}\mathfrak{p}_{\mathfrak{t}_0,T,\mathfrak{I}_N}^{\mathfrak{I}_\varphi,\varrho}\bar{\mathfrak{H}}_{\varrho,\mathfrak{I}_\varphi}^{\mathrm{can}}(\bar{\mathfrak{f}}_{\mathfrak{t}_0,T,\mathfrak{I}_N}^{\mathfrak{I}_\varphi}) + \kappa^{-1}\sum_{\varrho\in\mathbb{R}}\mathfrak{p}_{\mathfrak{t}_0,T,\mathfrak{I}_N}^{\mathfrak{I}_\varphi,\varrho}\log\mathbf{E}^{\mu_{\varrho,\mathfrak{I}_\varphi}^{\mathrm{can}}}\mathbf{Exp}\,(\kappa\varphi).
$$
$$\tag{3.28}$$

Sums over $\mathbb{R}$ are over finite sets of possible spin densities on $\mathfrak{I}_\varphi$. As $\mathfrak{p}$-terms sum to 1 on the finite set of possible $\varrho\in\mathbb{R}$ because they are probabilities under the same probability measure of disjoint hyperplanes, we get

$$
\kappa^{-1}\sum\nolimits_{\varrho\in\mathbb{R}}\mathfrak{p}_{\mathfrak{t}_0,T,\mathfrak{I}_N}^{\mathfrak{I}_\varphi,\varrho}\log\mathbf{E}^{\mu_{\varrho,\mathfrak{I}_\varphi}^{\mathrm{can}}}\mathbf{Exp}\,(\kappa\varphi) \;\leqslant\; \kappa^{-1}\sup_{\varrho\in\mathbb{R}}\log\mathbf{E}^{\mu_{\varrho,\mathfrak{I}_\varphi}^{\mathrm{can}}}\mathbf{Exp}\,(\kappa\varphi). \tag{3.29}
$$

Concerning the relative entropy, we apply the LSI within Theorem A in [25] to introduce a $\mathfrak{I}_\varphi$-local Dirichlet form estimate in terms of canonical ensemble Dirichlet forms which glue to a grand-canonical Dirichlet form:

$$
\kappa^{-1}\sum\nolimits_{\varrho\in\mathbb{R}}\mathfrak{p}_{\mathfrak{t}_0,T,\mathfrak{I}_N}^{\mathfrak{I}_\varphi,\varrho}\bar{\mathfrak{H}}_{\varrho,\mathfrak{I}_\varphi}^{\mathrm{can}}\left(\bar{\mathfrak{f}}_{\mathfrak{t}_0,T,\mathfrak{I}_N}^{\mathfrak{I}_\varphi}\right) \;\lesssim\; \kappa^{-1}|\mathfrak{I}_\varphi|^2\sum\nolimits_{\varrho\in\mathbb{R}}\mathfrak{p}_{\mathfrak{t}_0,T,\mathfrak{I}_N}^{\mathfrak{I}_\varphi,\varrho}\bar{\mathfrak{D}}_{\varrho,\mathfrak{I}_\varphi}^{\mathrm{can}}\left(\bar{\mathfrak{f}}_{\mathfrak{t}_0,T,\mathfrak{I}_N}^{\mathfrak{I}_\varphi}\right)
$$

$$
= \kappa^{-1}|\mathfrak{I}_\varphi|^2\mathfrak{D}_{0,\mathfrak{I}_\varphi}\left(\bar{\mathfrak{f}}_{\mathfrak{t}_0,T,\mathfrak{I}_N}^{\mathfrak{I}_\varphi}\right) \tag{3.30}
$$

where the last identity follows from the definition of the $\bar{\mathfrak{D}}$-functional, or equivalently by definition of conditional probability and conditional expectation. Via convexity of the Dirichlet form in Corollary 10.3 in Appendix 1.10 of [21], as in the classical one-block estimate in Theorem 2.4 in [14], the final Dirichlet form is controlled by the LHS of the bound of Proposition 3.6 but with an additional factor of $|\mathfrak{I}_N|^{-1}$ from spatial averaging and $|\mathfrak{I}_\varphi|$ from redundancies in counting contributions from jumps. Here in our application of Proposition 3.6, we choose $\varepsilon\in\mathbb{R}_{>0}$ therein to be $2\varepsilon$ where the $\varepsilon\in\mathbb{R}_{>0}$-factor in $2\varepsilon$ is the $\varepsilon$-value we picked in the current Lemma 3.8. Equivalently, we use Proposition 3.6 for twice the $\varepsilon$-exponent we are using here. This gives

$$
\kappa^{-1}|\mathfrak{I}_\varphi|^2\mathfrak{D}_{0,\mathfrak{I}_\varphi}\left(\bar{\mathfrak{f}}_{\mathfrak{t}_0,T,\mathfrak{I}_N}^{\mathfrak{I}_\varphi}\right) \;\lesssim_\varepsilon\; T^{-1}\kappa^{-1}N^{-\frac{3}{4}+10\varepsilon}|\mathfrak{I}_N|^{-1}|\mathfrak{I}_\varphi|^3. \tag{3.31}
$$

We clarify that we use Proposition 3.6 with the initial measure therein to be $\mu_{\mathfrak{t}_0,N}$ here. This completes the proof. $\qquad\square$

*3.6. Proofs of Technical Estimates.* In this last subsection, we provide proofs of the technical results Lemma 3.15, Lemma 3.17, Lemma 3.18, and Lemma 3.19 that we used in the proof of Proposition 3.6. We prove them in the order they are written.

*Proof of Lemma 3.15.* The first estimate (3.20) is standard. We cite Theorem 9.2 in Appendix 1 of [21] and use the observation that $\mathfrak{H}_{0,\mathbb{Z}}(\bar{\mathfrak{f}}_{0,N})\lesssim N^{3/2+\varepsilon}$ and use the diffusive scaling of $N^2$ in the infinitesimal generator $\mathfrak{S}^{N,!!}$. To obtain (3.21), we construct a coupling via two-species process to relate $\mathfrak{f}$ and $\bar{\mathfrak{f}}$ from which we get a Dirichlet form comparison by a total variation bound.

- Consider a coupling of initial spin configurations distributed via $\mu_{0,N}$ and $\widetilde{\mu}_{0,N}$ respectively by sampling a first configuration via $\mu_{0,N}$, and then sampling a configuration via $\widetilde{\mu}_{0,N}$ by taking the projection of the first configuration via $\mu_{0,N}$ onto $\mathfrak{I}_{N,+} \subseteq \mathbb{Z}$ and sampling the remaining spins via the grand-canonical ensemble $\mu_{0,\mathbb{Z}\setminus\mathfrak{I}_{N,+}}$. Recall $\mathfrak{I}_{N,+}$ in the statement of Lemma 3.15.
- We couple the above configurations in the following fashion. First, we observe the underlying symmetric exclusion process may be realized as attaching Poisson clocks to every *bond* connecting points inside $\mathbb{Z}$. Every step in the symmetric exclusion process, or equivalently every ringing of any Poisson clock, then corresponds to swapping the spins at those two points in the corresponding bond. We will first couple the dynamics of the symmetric exclusion model underlying the total dynamic of the two configurations by coupling these "spin-swap" dynamics. We emphasize that this is not the basic coupling. In particular, one configuration swaps possibly identical spins along a bond via this "spin-swap" dynamic whenever the other one does.
- The remaining Poisson clocks associated to the totally asymmetric finite-range exclusion process are equipped with the basic coupling, so particles in the two configurations jump together when possible when the totally asymmetric part rings.
- Initially, all discrepancies between these two configurations are supported outside $\mathfrak{I}_{N,+} \subseteq \mathbb{Z}$. We then consider the following decomposition $\mathbb{Z} \setminus \mathfrak{I}_{N,+} = \cup_{\mathfrak{l}=1}\mathfrak{I}_{N,\mathfrak{l}}$, where each sub-lattice $\mathfrak{I}_{N,\mathfrak{l}}$ satisfies $|\mathfrak{I}_{N,\mathfrak{l}}| \lesssim N^2$ with universal implied constant and the distance between $\mathfrak{I}_{N,\mathfrak{l}}$ and $\mathfrak{I}_N$ is at least $N^{3/2+\varepsilon}\mathfrak{l}$ times universal factors. Discrepancies among the two initial configurations originally in $\mathfrak{I}_{N,\mathfrak{l}}$ must then travel a distance of at least $N^{3/2+\varepsilon}\mathfrak{l}$, times universal factors, to land in $\mathfrak{I}_N \overset{\bullet}{=} [\![ -N^{5/4+\varepsilon}, N^{5/4+\varepsilon} ]\!]$.
- Under the above semi-basic coupling, the dynamics of any tagged discrepancy are a *free and unsuppressed* symmetric finite-range random walk of speed $N^2$ and randomly suppressed/killed asymmetric finite-range random walk of speed $N^{3/2}$. We emphasize the free/unsuppressed nature of the symmetric random walk with speed $N^2$ comes from coupling the symmetric parts of the two-species exclusion process via spin-swap bond dynamics rather than the basic coupling.
- So, the probability we find a discrepancy among the respective particle configurations originally in $\mathfrak{I}_{N,\mathfrak{l}} \subseteq \mathbb{Z}$ that propagates into $\mathfrak{I}_N \subseteq \mathfrak{I}_{N,+} \subseteq \mathbb{Z}$ before time 100 is bounded above by the probability that a symmetric finite-range random walk with an asymmetry of speed of order $N^{3/2}$ traveled a distance $N^{3/2+\varepsilon}\mathfrak{l}$. Because each $\mathfrak{I}_{N,\mathfrak{l}}$ subset has at most $N^2$-many discrepancies by considering the size of this block, the probability that we find any discrepancy appearing in $\mathfrak{I}_N \subseteq \mathfrak{I}_{N,+} \subseteq \mathbb{Z}$ that originally came from $\mathfrak{I}_{N,\mathfrak{l}}$ is at most the previous exit probability for random walks with an additional factor $N^2$ coming from the union bound over all possible discrepancies originally in $\mathfrak{I}_{N,\mathfrak{l}}$. We then take another union bound and sum over all subsets $\mathfrak{I}_{N,\mathfrak{l}}$ with $\mathfrak{l} \in \mathbb{Z}_{\geqslant 1}$. This probability also provides an estimate for total variation distance between $\widetilde{\mathfrak{f}}^{\mathfrak{I}_N}$ and $\mathfrak{f}^{\mathfrak{I}_N}$ as is usual for couplings, because each is a probability density with respect to $\mu_{0,\mathfrak{I}_N}$ for the law of one of the two configurations at time $T \in \mathbb{R}_{\geqslant 0}$. If $\Sigma$ is the previous random walk with $\Sigma_0 = 0$, the Doob maximal inequality then gives, for any $C \in \mathbb{R}_{>1}$, the total variation bound

$$\sup_{0 \leqslant T \leqslant 100} \mathbf{E}^{\mu_{0,\mathfrak{I}_N}} |\widetilde{\mathfrak{f}}^{\mathfrak{I}_N}_{T,N} - \mathfrak{f}^{\mathfrak{I}_N}_{T,N}|$$

$$\lesssim \sum_{\mathfrak{l}=1}^{\infty} N^2 \mathbf{P}\left( \sup_{0 \leqslant t \leqslant 200} |\Sigma_t| \gtrsim N^{\frac{3}{2}+\varepsilon}\mathfrak{l} \right) \lesssim \sum_{\mathfrak{l}=1}^{\infty} \mathfrak{l}^{-C} N^2 N^{-C\varepsilon} N^{-\frac{3}{2}C} \mathbf{E}|\Sigma_{200}|^C$$

$$\lesssim_C \ N^{2-C\varepsilon}.$$

Above, we observe $\mathbf{E}|\Sigma_{200}|^C \lesssim_C N^{\frac{3}{2}C}$ by standard random walk estimates because it is symmetric of speed $N^2$ with speed-$N^{3/2}$ asymmetry. We now estimate the Dirichlet form in terms of the LHS. In what follows all expectations are taken with respect to the invariant measure $\mu_{0,\mathbb{Z}}$. Note $\mu_{0,\mathbb{Z}}$ is invariant under spin-swaps; thus it is an invariant measure under the $\mathfrak{S}_{x,y}$-generators. We now make a list of elementary but useful observations for the next step.

- First observe the deterministic identities $a^2 - b^2 = (a-b)(a+b)$ and $(a+b)^2 \lesssim a^2 + b^2$ with $a, b \in \mathbb{R}_{\geqslant 0}$. Given $a, b \in \mathbb{R}_{\geqslant 0}$, we also have $(a-b)^2 \lesssim |a^2-b^2|$. If $a, b = 0$ this is trivial, and otherwise $(a-b)^2 \lesssim |a-b|(a+b) \cdot |a-b||a+b|^{-1} \lesssim |a^2-b^2|$.
- Let $\sigma_{x,y} : \Omega_{\mathbb{Z}} \to \Omega_{\mathbb{Z}}$ denote the involution that swaps spins at the deterministic points $x, y \in \mathbb{Z}$. In particular, we have the identity $\mathfrak{S}_{x,y}\varphi = \varphi \circ \sigma_{x,y} - \varphi$ for any functional $\varphi : \Omega_{\mathbb{Z}} \to \mathbb{R}$. We additionally note that the measure $\mu_{0,\mathbb{Z}}$ is $\sigma_{x,y}$-invariant for the same reason it is $\mathfrak{S}_{x,y}$-invariant. Thus, we have $\mathbf{E}|\mathfrak{S}_{x,y}\varphi|^2 \lesssim \mathbf{E}|\varphi \circ \sigma_{x,y}|^2 + \mathbf{E}|\varphi|^2 \lesssim \mathbf{E}|\varphi|^2$.
- As $\mathfrak{f}$ and $\widetilde{\mathfrak{f}}$ are probability densities, the previous bullet point gives $\mathbf{E}|\mathfrak{S}_{x,y}\mathfrak{f}^{1/2}|^2 \lesssim \mathbf{E}\mathfrak{f} = 1$ and $\mathbf{E}|\mathfrak{S}_{x,y}\widetilde{\mathfrak{f}}^{1/2}|^2 \lesssim \mathbf{E}\widetilde{\mathfrak{f}} = 1$.
- The first bullet point and an elementary calculation give $|\mathfrak{S}_{x,y}\varphi - \mathfrak{S}_{x,y}\psi|^2 = |\varphi \circ \sigma_{x,y} - \psi \circ \sigma_{x,y}|^2 + |\varphi - \psi|^2$.

Using these bullet points and the Cauchy-Schwarz inequality gives the following with more explanation given after:

$$\mathbf{E}|\mathfrak{S}_{x,y}(\mathfrak{f}_{T,N}^{\mathfrak{J}_N})^{\frac{1}{2}}|^2 - \mathbf{E}|\mathfrak{S}_{x,y}(\widetilde{\mathfrak{f}}_{T,N}^{\mathfrak{J}_N})^{\frac{1}{2}}|^2$$

$$\lesssim \ \mathbf{E}(|\mathfrak{S}_{x,y}(\mathfrak{f}_{T,N}^{\mathfrak{J}_N})^{\frac{1}{2}} - \mathfrak{S}_{x,y}(\widetilde{\mathfrak{f}}_{T,N}^{\mathfrak{J}_N})^{\frac{1}{2}}| \cdot |\mathfrak{S}_{x,y}(\mathfrak{f}_{T,N}^{\mathfrak{J}_N})^{\frac{1}{2}} + \mathfrak{S}_{x,y}(\widetilde{\mathfrak{f}}_{T,N}^{\mathfrak{J}_N})^{\frac{1}{2}}|) \tag{3.32}$$

$$\lesssim \ \left(\mathbf{E}|\mathfrak{S}_{x,y}(\mathfrak{f}_{T,N}^{\mathfrak{J}_N})^{\frac{1}{2}} - \mathfrak{S}_{x,y}(\widetilde{\mathfrak{f}}_{T,N}^{\mathfrak{J}_N})^{\frac{1}{2}}|^2\right)^{\frac{1}{2}} \left(\mathbf{E}|\mathfrak{S}_{x,y}(\mathfrak{f}_{T,N}^{\mathfrak{J}_N})^{\frac{1}{2}} + \mathfrak{S}_{x,y}(\widetilde{\mathfrak{f}}_{T,N}^{\mathfrak{J}_N})^{\frac{1}{2}}|^2\right)^{\frac{1}{2}} \tag{3.33}$$

$$\lesssim \ \left(\mathbf{E}|\mathfrak{S}_{x,y}(\mathfrak{f}_{T,N}^{\mathfrak{J}_N})^{\frac{1}{2}} - \mathfrak{S}_{x,y}(\widetilde{\mathfrak{f}}_{T,N}^{\mathfrak{J}_N})^{\frac{1}{2}}|^2\right)^{\frac{1}{2}} \left(\mathbf{E}|\mathfrak{S}_{x,y}(\mathfrak{f}_{T,N}^{\mathfrak{J}_N})^{\frac{1}{2}}|^2 + \mathbf{E}|\mathfrak{S}_{x,y}(\widetilde{\mathfrak{f}}_{T,N}^{\mathfrak{J}_N})^{\frac{1}{2}}|^2\right)^{\frac{1}{2}} \tag{3.34}$$

$$\lesssim \ \left(\mathbf{E}|(\mathfrak{f}_{T,N}^{\mathfrak{J}_N})^{\frac{1}{2}} - (\widetilde{\mathfrak{f}}_{T,N}^{\mathfrak{J}_N})^{\frac{1}{2}}|^2 + \mathbf{E}|(\mathfrak{f}_{T,N}^{\mathfrak{J}_N} \circ \sigma_{x,y})^{\frac{1}{2}} - (\widetilde{\mathfrak{f}}_{T,N}^{\mathfrak{J}_N} \circ \sigma_{x,y})^{\frac{1}{2}}|^2\right)^{\frac{1}{2}} \tag{3.35}$$

$$\lesssim \ \left(\mathbf{E}|\mathfrak{f}_{T,N}^{\mathfrak{J}_N} - \widetilde{\mathfrak{f}}_{T,N}^{\mathfrak{J}_N}| + \mathbf{E}|\mathfrak{f}_{T,N}^{\mathfrak{J}_N} \circ \sigma_{x,y} - \widetilde{\mathfrak{f}}_{T,N}^{\mathfrak{J}_N} \circ \sigma_{x,y}|\right)^{\frac{1}{2}} \tag{3.36}$$

$$\lesssim \ \left(\mathbf{E}|\mathfrak{f}_{T,N}^{\mathfrak{J}_N} - \widetilde{\mathfrak{f}}_{T,N}^{\mathfrak{J}_N}|\right)^{\frac{1}{2}}. \tag{3.37}$$

The Dirichlet form sums at most $N^{100}$ such terms, so the last two displays and elementary asymptotics yield (3.21). We justify the previous list of estimates in case of interest. The first estimate above follows from applying the difference of squares in the first bullet point from our list with $a = \mathfrak{S}_{x,y}\mathfrak{f}^{1/2}$ and $b = \mathfrak{S}_{x,y}\widetilde{\mathfrak{f}}^{1/2}$. The second line follows from the Cauchy-Schwarz inequality. The third line follows from applying the second bound in the first bullet point preceding the above list of estimates for $a = \mathfrak{S}_{x,y}\mathfrak{f}$ and $b = \mathfrak{S}_{x,y}\widetilde{\mathfrak{f}}$ within the second expectation factor. The fourth line follows by the third bullet point

from the list preceding the above estimates applied for the second expectation factor and applying the final bullet point in said list for the first expectation factor. The fifth bound follows by $(a-b)^2 \lesssim |a^2 - b^2|$ we gave in that list. The last bound follows by $\sigma_{x,y}$-invariance of $\mu_{0,\mathbb{Z}}$. □

*Proof of Lemma 3.17.* The proposed inequality follows immediately from (4.14) and (4.15) within the proof of Lemma 4.1 in [10]. We emphasize that the proof of Lemma 4.1 in [10] does not depend on the maximal jump-length $\mathfrak{m} \in \mathbb{Z}_{\geqslant 1}$. □

*Proof of Lemma 3.18.* Note the Cauchy-Schwarz inequality $ab \lesssim N^{3/4}a^2 + N^{-3/4}b^2$ for $a = \mathfrak{S}_{x,y}\mathfrak{f}_{\mathfrak{l}}^{1/2}$ and $b = (\mathfrak{f}_{\mathfrak{l},x,y} - \mathfrak{f}_{\mathfrak{l}})\mathfrak{f}_{\mathfrak{l}}^{-1/2}$:

$$N^2 \mathbf{E} \sum_{(x,y)\in\partial\mathfrak{I}_{N,\mathfrak{l}}} \alpha_{|y-x|}(\mathfrak{f}_{\mathfrak{l},x,y} - \mathfrak{f}_{\mathfrak{l}})\mathfrak{f}_{\mathfrak{l}}^{-\frac{1}{2}}\mathfrak{S}_{x,y}\mathfrak{f}_{\mathfrak{l}}^{\frac{1}{2}} \lesssim N^{\frac{11}{4}}\mathbf{E} \sum_{(x,y)\in\partial\mathfrak{I}_{N,\mathfrak{l}}} \alpha_{|y-x|}|\mathfrak{S}_{x,y}\mathfrak{f}_{\mathfrak{l}}^{\frac{1}{2}}|^2$$
$$+ N^{\frac{5}{4}}\mathbf{E} \sum_{(x,y)\in\partial\mathfrak{I}_{N,\mathfrak{l}}} \alpha_{|y-x|}|\mathfrak{f}_{\mathfrak{l},x,y} - \mathfrak{f}_{\mathfrak{l}}|^2\mathfrak{f}_{\mathfrak{l}}^{-1}. \tag{3.38}$$

We first bound the first term on the RHS of (3.38) by convexity of the Dirichlet form; this lets us replace the probability density $\mathfrak{f}_{\mathfrak{l}}$ in this term by projections onto larger subsets. Combining this with the observation $\frac{11}{4} = \frac{3}{2} + \frac{5}{4}$ gives

$$N^{\frac{11}{4}}\mathbf{E} \sum_{(x,y)\in\partial\mathfrak{I}_{N,\mathfrak{l}}} \alpha_{|y-x|}|\mathfrak{S}_{x,y}\mathfrak{f}_{\mathfrak{l}}^{\frac{1}{2}}|^2 = N^{\frac{11}{4}} \sum_{(x,y)\in\partial\mathfrak{I}_{N,\mathfrak{l}}} \alpha_{|y-x|}\mathbf{E}|\mathfrak{S}_{x,y}\mathfrak{f}_{\mathfrak{l}}^{\frac{1}{2}}|^2$$
$$\lesssim N^{\frac{3}{2}} \sum_{(x,y)\in\partial\mathfrak{I}_{N,\mathfrak{l}}} \alpha_{|y-x|} \sum_{j=1}^{N^{\frac{5}{4}}} \mathbf{E}|\mathfrak{S}_{x,y}\mathfrak{f}_{\mathfrak{l}+j}^{\frac{1}{2}}|^2. \tag{3.39}$$

The RHS of (3.39) appears in the upper bound of the result Lemma 3.18 we are currently proving. We move to the second term on the RHS of (3.38). For this, we first observe the inequality $\mathfrak{f}_{\mathfrak{l},x,y}\mathfrak{f}_{\mathfrak{l}}^{-1} \lesssim 1$. This follows from the fact that $\mathfrak{f}_{\mathfrak{l}}$ is the conditional expectation of $\mathfrak{f}_{\mathfrak{l},x,y}$ over a uniformly bounded set with uniformly positive probabilities. This was used in the proof of Lemma 4.1 in [10] for example. As $\mathfrak{f}_{\mathfrak{l},x,y}$ and $\mathfrak{f}_{\mathfrak{l}}$ are probability densities, we use $|a-b|^2 \lesssim a^2 + b^2$ with $a = \mathfrak{f}_{\mathfrak{l},x,y}$ and $b = \mathfrak{f}_{\mathfrak{l}}$ to get

$$N^{\frac{5}{4}}\mathbf{E}\sum_{(x,y)\in\partial\mathfrak{I}_{N,\mathfrak{l}}} \alpha_{|y-x|}|\mathfrak{f}_{\mathfrak{l},x,y} - \mathfrak{f}_{\mathfrak{l}}|^2\mathfrak{f}_{\mathfrak{l}}^{-1}$$
$$\lesssim N^{\frac{5}{4}}\sum_{(x,y)\in\partial\mathfrak{I}_{N,\mathfrak{l}}} \alpha_{|y-x|}\mathbf{E}|\mathfrak{f}_{\mathfrak{l},x,y}|^2\mathfrak{f}_{\mathfrak{l}}^{-1} + N^{\frac{5}{4}}\sum_{(x,y)\in\partial\mathfrak{I}_{N,\mathfrak{l}}} \alpha_{|y-x|}\mathbf{E}\mathfrak{f}_{\mathfrak{l}} \tag{3.40}$$
$$\lesssim N^{\frac{5}{4}}. \tag{3.41}$$

The last bound follows by finite support of $\alpha$-coefficients and inequalities $\mathbf{E}|\mathfrak{f}_{\mathfrak{l},x,y}|^2\mathfrak{f}_{\mathfrak{l}}^{-1} \lesssim \mathbf{E}\mathfrak{f}_{\mathfrak{l},x,y} = 1$ and $\mathbf{E}\mathfrak{f}_{\mathfrak{l}} = 1$. □

*Proof of Lemma 3.19.* We study $\Phi_\wedge \overset{\bullet}{=} \Phi_{\wedge,1}$. The proof to estimate $\Phi_{\wedge,2}$ follows from identical considerations because the sign and the direction of the asymmetric jump will not matter in what follows. Via the finite support of $\alpha, \gamma$ coefficients, we get

$$|\Phi_\wedge| \lesssim N^{\frac{3}{2}} \sum_{(x,y)\in\partial\mathfrak{I}_{N,\mathfrak{l}}} \alpha_{|x-y|}\gamma_{|x-y|}\mathbf{E}|\mathfrak{f}_{\mathfrak{l},x,y} + \mathfrak{f}_{\mathfrak{l}}| \cdot (\mathfrak{f}_{\mathfrak{l}})^{-\frac{1}{2}} \cdot |\mathfrak{S}_{x,y}(\mathfrak{f}_{\mathfrak{l}}^{\frac{1}{2}})| \lesssim$$
$$N^{\frac{3}{2}} \sup_{(x,y)\in\partial\mathfrak{I}_{N,\mathfrak{l}}} \mathbf{E}|\mathfrak{f}_{\mathfrak{l},x,y} + \mathfrak{f}_{\mathfrak{l}}| \cdot (\mathfrak{f}_{\mathfrak{l}})^{-\frac{1}{2}} \cdot |\mathfrak{S}_{x,y}(\mathfrak{f}_{\mathfrak{l}}^{\frac{1}{2}})|. \tag{3.42}$$

We now apply the Cauchy-Schwarz inequality with respect to the expectation to get

$$N^{\frac{3}{2}}\mathbf{E}|\mathfrak{f}_{\mathfrak{l},x,y} + \mathfrak{f}_{\mathfrak{l}}| \cdot (\mathfrak{f}_{\mathfrak{l}})^{-1/2} \cdot |\mathfrak{S}_{x,y}(\mathfrak{f}_{\mathfrak{l}}^{1/2})|$$
$$\lesssim N^{\frac{3}{2}}\left(\mathbf{E}|\mathfrak{f}_{\mathfrak{l},x,y} + \mathfrak{f}_{\mathfrak{l}}|^2\mathfrak{f}_{\mathfrak{l}}^{-1}\right)^{1/2}\left(\mathbf{E}|\mathfrak{S}_{x,y}(\mathfrak{f}_{\mathfrak{l}}^{\frac{1}{2}})|^2\right)^{1/2} \lesssim N^{\frac{3}{2}}\left(\mathbf{E}|\mathfrak{S}_{x,y}(\mathfrak{f}_{\mathfrak{l}}^{\frac{1}{2}})|^2\right)^{1/2}. \tag{3.43}$$

The last estimate (3.43) follows by the three bounds $|\mathfrak{f}_{\mathfrak{l},x,y} + \mathfrak{f}_{\mathfrak{l}}|^2\mathfrak{f}_{\mathfrak{l}}^{-1} \lesssim \mathfrak{f}_{\mathfrak{l},x,y}^2\mathfrak{f}_{\mathfrak{l}}^{-1} + \mathfrak{f}_{\mathfrak{l}}$ and $\mathbf{E}|\mathfrak{f}_{\mathfrak{l},x,y}|^2\mathfrak{f}_{\mathfrak{l}}^{-1} \lesssim \mathbf{E}\mathfrak{f}_{\mathfrak{l},x,y} = 1$ and $\mathbf{E}\mathfrak{f}_{\mathfrak{l}} = 1$ in the proof of Lemma 3.18. The last expectation on the far RHS of (3.43) is bounded by the global $\mathfrak{D}_{0,\mathbb{Z}}$-Dirichlet form of $\mathfrak{f}_{\mathfrak{l}}$ as each bond in the Dirichlet form contributes non-negative amount and $\mathfrak{S}_{x,y}$ corresponds to just one bond in the $\mathfrak{D}_{0,\mathbb{Z}}$-form. Last, the global $\mathfrak{D}_{0,\mathbb{Z}}$-Dirichlet form of $\mathfrak{f}_{\mathfrak{l}}$ is bounded by that of the unprojected global Radon-Nikodym derivative $\mathfrak{f}_{T,N}$ of the time-$T$ probability measure $\mu_{T,N}$ by convexity of the Dirichlet form. We combine this with (3.42) and (3.43) to finish the proof.                                                                     □

## 4. Analytic Compactification

We replace the microscopic Cole-Hopf transform with a "compactified" version at the level of stochastic heat-operator-type equations such as (2.5b). We start by "compactifying" the heat operators. Below $\varepsilon_{\mathrm{cpt}} \in \mathbb{R}_{>0}$ is arbitrarily small but universal.

**Definition 4.1.** Define $\mathbb{T}_N \overset{\bullet}{=} [\![-N^{\frac{5}{4}+\varepsilon_{\mathrm{cpt}}}, N^{\frac{5}{4}+\varepsilon_{\mathrm{cpt}}}]\!]$ and let $\bar{\mathfrak{P}}^N$ be the heat kernel on $\mathbb{R}_{\geqslant 0}^2 \times \mathbb{T}_N^2$ satisfying $\bar{\mathfrak{P}}_{S,S,x,y}^N = \mathbf{1}_{x=y}$ and

$$\partial_T \bar{\mathfrak{P}}_{S,T,x,y}^N = \mathscr{L}^{!!}\bar{\mathfrak{P}}_{S,T,x,y}^N. \tag{4.1}$$

Above, we have introduced $\mathscr{L}^{!!} \overset{\bullet}{=} 2^{-1}\sum_{k=1}^{\mathfrak{m}}\widetilde{\alpha}_k\bar{\Delta}_k^{!!}$, in which $\bar{\Delta}_k^{!!} \overset{\bullet}{=} N^2\bar{\Delta}_k$ and $\bar{\Delta}_k\varphi_x \overset{\bullet}{=} \varphi_{x+k} + \varphi_{x-k} - 2\varphi_x$ provided any test function $\varphi : \mathbb{T}_N \to \mathbb{R}$. Here, the barred discrete differential operators mean that addition/subtraction are both defined on $\mathbb{T}_N$ with periodic boundary conditions $\inf \mathbb{T}_N = \sup \mathbb{T}_N + 1$, thus realizing $\mathbb{T}_N$ as a torus. For $\varphi : \mathbb{R}_{\geqslant 0} \times \mathbb{T}_N \to \mathbb{R}$, we also define

$$\bar{\mathfrak{P}}_{T,x}^N\varphi \overset{\bullet}{=} \bar{\mathfrak{P}}_{T,x}^N\varphi_{S,y} \overset{\bullet}{=} \bar{\mathfrak{P}}_{T,x}^N\varphi_{\bullet,\bullet} \overset{\bullet}{=} \int_0^T\sum_{y\in\mathbb{T}_N}\bar{\mathfrak{P}}_{S,T,x,y}^N \cdot \varphi_{S,y}\, \mathrm{d}S \tag{4.2a}$$

$$\bar{\mathfrak{P}}_{T,x}^{N,\mathbf{X}}\varphi \overset{\bullet}{=} \bar{\mathfrak{P}}_{T,x}^{N,\mathbf{X}}\varphi_{0,y} \overset{\bullet}{=} \bar{\mathfrak{P}}_{T,x}^{N,\mathbf{X}}\varphi_{0,\bullet} \overset{\bullet}{=} \sum_{y\in\mathbb{T}_N}\bar{\mathfrak{P}}_{0,T,x,y}^N \cdot \varphi_{0,y}. \tag{4.2b}$$

*Remark 4.2.* Although $\bar{\mathfrak{P}}^N$ is not the full-line heat kernel $\mathfrak{P}^N$ for which heat kernel estimates are established in Proposition A.1 in [10], the heat kernel $\bar{\mathfrak{P}}^N$ admits a classical explicit representation given by spatial translations of $\mathfrak{P}^N$. Thus, Proposition A.1 in [10] holds for $\bar{\mathfrak{P}}^N$ for uniformly bounded times upon replacing the distance on $\mathbb{Z}$ with the geodesic distance on $\mathbb{T}_N$. The same is true if we replace "Proposition A.1" with "Corollary A.2" in [10] for the same reason. We organize this in Lemma A.1.

The aforementioned "compactification" of the microscopic Cole-Hopf transform defined here is basically given by replacing heat operators in (2.5b) by compactifications defined above. We will also perturb the initial data for reasons which are briefly discussed in

Remark 4.5. Roughly speaking, like [10] our analysis of stochastic heat-type equations will use regularity of their initial data, and our perturbation of the initial data in Definition 4.3 below ensures such regularity with respect to the geodesic distance on the corresponding geometry $\mathbb{T}_N$. However, the perturbations are not detectable in any $\mathbf{C}(\mathbb{K})$ in the large-$N$ limit so the details of these perturbations are not too delicate and can be modified in many ways without changing much.

**Definition 4.3.** Define $\widetilde{\mathfrak{Z}}^N \triangleq \mathfrak{Z}^N - \bar{\mathfrak{Z}}^N$ where $\bar{\mathfrak{Z}}^N$ is the solution to the following equation with terms to be defined afterwards:

$$\bar{\mathfrak{Z}}_{T,x}^N \triangleq \bar{\mathfrak{P}}_{T,x}^{N,\mathbf{X}}(\chi_\bullet \mathfrak{Z}_{0,\bullet}^N) + \bar{\mathfrak{P}}_{T,x}^N(\bar{\mathfrak{Z}}^N d\xi^N) + \bar{\mathfrak{P}}_{T,x}^N(\bar{\Phi}^{N,2}) + \bar{\mathfrak{P}}_{T,x}^N(\bar{\Phi}^{N,3}). \qquad (4.3)$$

Above $\chi : \mathbb{Z} \to \mathbb{R}_{\geq 0}$ is a cutoff function that satisfies the following support and derivative-type constraints:

- We have $\chi_y = 1$ for $|y| \leq \frac{1}{3} \sup \mathbb{T}_N$ and $\chi_y = 0$ for $|y| \geq \frac{2}{3} \sup \mathbb{T}_N$, so the support of $\chi$ is contained *in the interior of* $\mathbb{T}_N$.
- We have a macroscopic-length-scale Lipschitz bound $|\chi_y - \chi_{y'}| \leq N^{-1}|y - y'|$ and a uniform bound $|\chi_y| \leq 1$ for all $y, y' \in \mathbb{Z}$.

Defining $N^{-1}\bar{\nabla}_k^! \varphi = \varphi_{x+k} - \varphi_x$, where barred-gradients are gradients on $\mathbb{T}_N$ with periodic boundary conditions, we have

$$\bar{\Phi}_{T,x}^{N,2} \triangleq N^{\frac{1}{2}} \mathscr{A}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{T,x}) \cdot \bar{\mathfrak{Z}}_{T,x}^N + N^{\beta_X}\left(\widetilde{\sum}_{\mathfrak{l}=1,\ldots,N^{\beta_X}} \widetilde{\mathfrak{g}}_{T,x}^{\mathfrak{l}}\right)\bar{\mathfrak{Z}}_{T,x}^N$$
$$+ N^{-\frac{1}{2}}\widetilde{\sum}_{\mathfrak{l}=1,\ldots,N^{\beta_X}} \bar{\nabla}_{-7\mathfrak{l}\mathfrak{m}}^! (\mathfrak{b}_{T,x}^{\mathfrak{l}} \bar{\mathfrak{Z}}_{T,x}^N). \qquad (4.4)$$

The term $\bar{\Phi}^{N,3}$ is the compactification of $\Phi^{N,3}$ with weakly vanishing content defined in the statement of Proposition 2.7:

$$\bar{\Phi}_{T,x}^{N,3} \triangleq \mathfrak{w}_{T,x}\bar{\mathfrak{Z}}_{T,x}^N + \sum_{k=-2\mathfrak{m}}^{2\mathfrak{m}} c_k \bar{\nabla}_k^!(\mathfrak{w}_{T,x}^k \bar{\mathfrak{Z}}_{T,x}^N). \qquad (4.5)$$

We recall the functionals inside $\bar{\Phi}^{N,2}$ and $\bar{\Phi}^{N,3}$, namely upon dropping all $\bar{\mathfrak{Z}}^N$-terms therein, are defined in Proposition 2.7. We emphasize that although $\bar{\mathfrak{Z}}^N$ evolves on $\mathbb{T}_N$, the particle system and its functionals evolve via the global $\mathfrak{S}^{N,!!}$-dynamic on $\Omega_{\mathbb{Z}}$.

*Remark 4.4.* The term $\bar{\mathfrak{Z}}^N d\xi^N$ is the martingale associated to the Poisson process whose jumps are those of $dM$ in (2.4) of [10] then scaled by $\bar{\mathfrak{Z}}^N$ at the same space-time point. We will make similar constructions later in the paper.

*Remark 4.5.* The $\chi$-cutoff guarantees the initial data of the $\bar{\mathfrak{Z}}^N$ process has spatial regularity with respect to geodesic distance on $\mathbb{T}_N$, inherited via a priori near-stationary regularity bounds and $\chi$-regularity bounds, *but without changing the initial data in a fashion that is detectable at macroscopic length-scales*. Indeed, this $\chi$-cutoff allows us to forget about any boundary conditions, given its support, while doing nothing at lengths of order $|\mathbb{T}_N|$, which is order $N^{5/4+\varepsilon_{\mathrm{cpt}}}$, about the origin; see Definition 4.3.

The primary goal of this section is the following estimate which compares $\mathfrak{Z}^N$ and $\bar{\mathfrak{Z}}^N$. In particular, one consequence of this next result is a comparison of these two processes on compact space-time sets which are the sets of interest in Theorem 1.7.

**Proposition 4.6.** *Take any $C_1 \in \mathbb{R}_{>0}$ arbitrarily large but universal. We consider the following data for any $\mathfrak{t} \in \mathbb{R}_{\geq 0}$ and $\mathbb{X} \subseteq \mathbb{Z}$.*

- *Define an "interior set"* $\widetilde{\mathbb{T}}_N \doteq [\![-N^{\frac{5}{4}+\frac{1}{2}\varepsilon_{\mathrm{cpt}}}, N^{\frac{5}{4}+\frac{1}{2}\varepsilon_{\mathrm{cpt}}}]\!] \subseteq \mathbb{T}_N = [\![-N^{\frac{5}{4}+\varepsilon_{\mathrm{cpt}}}, N^{\frac{5}{4}+\varepsilon_{\mathrm{cpt}}}]\!]$ *but far from the boundary, so* $|\widetilde{\mathbb{T}}_N| \ll |\mathbb{T}_N|$.
- *Define the time-discretized norm* $[\varphi]_{\mathfrak{t};\mathbb{X}} \doteq \sup_{(T,x)\in\mathfrak{I}_{\mathfrak{t}}\times\mathbb{X}} |\varphi_{T,x}|$ *with time-discretization* $\mathfrak{I}_{\mathfrak{t}} \doteq \{\mathfrak{t}\mathfrak{j}N^{-100}\}_{\mathfrak{j}=0}^{N^{100}}$.

*There exists* $C_2 \in \mathbb{R}_{>0}$ *depending only on* $C_1 \in \mathbb{R}_{>0}$ *satisfying* $C_2 \gtrsim C_1$ *with universal implied constant if* $C_1 \in \mathbb{R}_{>0}$ *is sufficiently big such that for any deterministic time-horizon* $0 \leqslant \mathfrak{t}_{\mathfrak{f}} \leqslant 1$, *we have the following outside an event of probability at most* $\kappa_{C_1,C_2}N^{-C_2}$:

$$[\widetilde{\mathfrak{Z}}^N]_{\mathfrak{t}_{\mathfrak{f}};\widetilde{\mathbb{T}}_N} \lesssim N^{-C_1}. \tag{4.6}$$

*Remark 4.7.* We will eventually upgrade the estimate in Proposition 4.6 to an estimate of the supremum over the semi-discrete set $[0, 1] \times \widetilde{\mathbb{T}}_N$ as opposed to the supremum over a fully discrete set. This may be plausible given that for time-scales well below the microscopic time-scale we expect to see very little occur. However, because the stochastic equation for $\bar{\mathfrak{Z}}^N$ is multiplicative in $\bar{\mathfrak{Z}}^N$, this will require a priori estimates for $\bar{\mathfrak{Z}}^N$ which we are not quite ready to establish. We mention this, however, to explain the eventual utility behind Proposition 4.6. Such upgrade from the supremum over a totally discrete set to the supremum over a semi-discrete set will all be done in this final section of this paper concerning the proof of Theorem 1.7; see Lemma 8.6 and its proof for details. In particular, the reader is invited to treat this section as a reason to study the compactification of $\bar{\mathfrak{Z}}^N$ until we return to upgrading the comparison in Proposition 4.6 to an estimate for the supremum over the semi-discrete set $[0, 1] \times \widetilde{\mathbb{T}}_N$.

*Remark 4.8.* To illustrate the ideas behind Proposition 4.6 and where the proof comes from, let us pretend

$$\mathfrak{Z}^N = \mathfrak{P}^{N,\mathbf{X}}\mathfrak{Z}^N_{0,\bullet} \quad \text{and} \quad \bar{\mathfrak{Z}}^N = \bar{\mathfrak{P}}^{N,\mathbf{X}}\mathfrak{Z}^N_{0,\bullet}. \tag{4.7}$$

In this case, the difference $\mathfrak{Z}^N - \bar{\mathfrak{Z}}^N$ on $\widetilde{\mathbb{T}}_N$ is controlled by comparing the pair of heat kernels $\mathfrak{P}^N$ and $\bar{\mathfrak{P}}^N$ on uniformly bounded time sets with backwards spatial variable in $\widetilde{\mathbb{T}}_N$. These heat kernels correspond to random walks that differ with exponentially small probability when the backwards spatial variable/initial starting point is in the set $\widetilde{\mathbb{T}}_N$. This basically gives Proposition 4.6 in this "idealized" case. We will deal with the other terms in the stochastic equations for $\mathfrak{Z}^N$ and $\bar{\mathfrak{Z}}^N$ perturbatively.

*4.1. Stochastic fundamental solutions.* In spirit of Remark 4.8, we first consider the following fundamental solutions.

**Definition 4.9.** We define $\mathfrak{J}^N$ as a space-time process on $\mathbb{R}^2_{\geqslant 0} \times \mathbb{Z}^2$ via $\mathfrak{J}^N_{S,S,x,y} = \mathbf{1}_{x=y}$ and, for any $T \in \mathbb{R}_{\geqslant 0}$ satisfying $T \geqslant S$,

$$\mathfrak{J}^N_{S,T+\mathfrak{t},x,y} = \mathfrak{P}^{N,\mathbf{X}}_{\mathfrak{t},x}\mathfrak{J}^N_{S,T,\bullet,y} + \mathfrak{P}^N_{\mathfrak{t},x}(\mathfrak{J}^N_{S,T+\bullet,\bullet,y}\mathrm{d}\xi^N) + \mathfrak{P}^N_{\mathfrak{t},x}(\Phi^{N,2}_{S,T+\bullet,\bullet,y}) + \mathfrak{P}^N_{\mathfrak{t},x}(\Phi^{N,3}_{S,T+\bullet,\bullet,y}). \tag{4.8}$$

Above, we introduced the following $\Phi^{N,2}$/pseudo-gradient content from Proposition 2.7, but adapted to $\mathfrak{J}^N$, for any $T \geqslant S$:

$$\Phi^{N,2}_{S,T,x,y} \doteq N^{\frac{1}{2}}\mathscr{A}^{\mathbf{X},-}_{N^{\beta_X}}(\mathfrak{g}_{T,x}) \cdot \mathfrak{J}^N_{S,T,x,y} + N^{\beta_X}\left(\widetilde{\sum}_{\mathfrak{l}=1,\ldots,N^{\beta_X}}\widetilde{\mathfrak{g}}^{\mathfrak{l}}_{T,x}\right)\mathfrak{J}^N_{S,T,x,y}$$
$$+ N^{-\frac{1}{2}}\widetilde{\sum}_{\mathfrak{l}=1,\ldots,N^{\beta_X}}\nabla^!_{-7\mathfrak{l}\mathfrak{m}}(\mathfrak{b}^{\mathfrak{l}}_{T,x}\mathfrak{J}^N_{S,T,x,y}). \tag{4.9}$$

Again the $\Phi^{N,3}$-term contains weakly vanishing content from Proposition 2.7 but with a multiplicative $\mathfrak{J}^N$-factor:

$$\Phi^{N,3}_{S,T,x,y} \overset{\bullet}{=} \mathfrak{w}_{T,x}\mathfrak{J}^N_{S,T,x,y} + \sum_{k=-2\mathfrak{m}}^{2\mathfrak{m}} c_k \nabla^!_k(\mathfrak{w}^k_{T,x}\mathfrak{J}^N_{S,T,x,y}). \tag{4.10}$$

We analogously define $\bar{\mathfrak{J}}^N$ as a space-time field on $\mathbb{R}^2_{\geqslant 0} \times \mathbb{T}^2_N$ via $\bar{\mathfrak{J}}^N_{S,S,x,y} = \mathbf{1}_{x=y}$ and, for any $T \in \mathbb{R}_{\geqslant 0}$ satisfying $T \geqslant S$,

$$\bar{\mathfrak{J}}^N_{S,T+\mathfrak{t},x,y} = \bar{\mathfrak{P}}^{N,\mathbf{X}}_{\mathfrak{t},x}\bar{\mathfrak{J}}^N_{S,T,\bullet,y} + \bar{\mathfrak{P}}^N_{\mathfrak{t},x}(\bar{\mathfrak{J}}^N_{S,T+\bullet,\bullet,y}\mathrm{d}\xi^N) + \bar{\mathfrak{P}}^N_{\mathfrak{t},x}(\bar{\Phi}^{N,2}_{S,T+\bullet,\bullet,y}) + \bar{\mathfrak{P}}^N_{\mathfrak{t},x}(\bar{\Phi}^{N,3}_{S,T+\bullet,\bullet,y}). \tag{4.11}$$

We introduced the following data for any $T \geqslant S$ where gradients $\bar{\nabla}$ are defined with respect to addition on the torus $\mathbb{T}_N$:

$$\bar{\Phi}^{N,2}_{S,T,x,y} \overset{\bullet}{=} N^{\frac{1}{2}}\mathscr{A}^{\mathbf{X},-}_{N^{\beta_X}}(\mathfrak{g}_{T,x}) \cdot \bar{\mathfrak{J}}^N_{S,T,x,y} + N^{\beta_X}\left(\widetilde{\sum_{\mathfrak{l}=1,\ldots,N^{\beta_X}}\widetilde{\mathfrak{g}}^{\mathfrak{l}}_{T,x}}\right)\bar{\mathfrak{J}}^N_{S,T,x,y}$$
$$+ N^{-\frac{1}{2}}\widetilde{\sum_{\mathfrak{l}=1,\ldots,N^{\beta_X}}}\bar{\nabla}^!_{-7\mathfrak{l}\mathfrak{m}}(\mathfrak{b}^{\mathfrak{l}}_{T,x}\bar{\mathfrak{J}}^N_{S,T,x,y}). \tag{4.12}$$

The following $\bar{\Phi}^{N,3}$ is the compactification of the $\Phi^{N,3}$ in the $\mathfrak{J}^N$-equation introduced above:

$$\bar{\Phi}^{N,3}_{S,T,x,y} \overset{\bullet}{=} \mathfrak{w}_{T,x}\bar{\mathfrak{J}}^N_{S,T,x,y} + \sum_{k=-2\mathfrak{m}}^{2\mathfrak{m}} c_k \bar{\nabla}^!_k(\mathfrak{w}^k_{T,x}\bar{\mathfrak{J}}^N_{S,T,x,y}). \tag{4.13}$$

We conclude by defining the difference in fundamental solutions $\widetilde{\mathfrak{J}}^N \overset{\bullet}{=} \mathfrak{J}^N - \bar{\mathfrak{J}}^N$.

Elementary calculation of the forwards-time-differential of $\mathfrak{J}^N$ and $\bar{\mathfrak{J}}^N$ that resembles the verification of the Duhamel principle for parabolic equations gives the following SDE-type equations for $\mathfrak{J}^N$ and $\bar{\mathfrak{J}}^N$; we give a clarifying point after:

$$\mathrm{d}_\mathfrak{t}\mathfrak{J}^N_{S,\mathfrak{t},x,y} = \mathscr{L}^{!!}\mathfrak{J}^N_{S,\mathfrak{t},x,y}\mathrm{d}\mathfrak{t} + \mathfrak{J}^N_{S,\mathfrak{t},x,y}\mathrm{d}\xi^N_{\mathfrak{t},x} + \Phi^{N,2}_{S,\mathfrak{t},x,y}\mathrm{d}\mathfrak{t} + \Phi^{N,3}_{S,\mathfrak{t},x,y}\mathrm{d}\mathfrak{t} \tag{4.14a}$$

$$\mathrm{d}_\mathfrak{t}\bar{\mathfrak{J}}^N_{S,\mathfrak{t},x,y} = \bar{\mathscr{L}}^{!!}\bar{\mathfrak{J}}^N_{S,\mathfrak{t},x,y}\mathrm{d}\mathfrak{t} + \bar{\mathfrak{J}}^N_{S,\mathfrak{t},x,y}\mathrm{d}\xi^N_{\mathfrak{t},x} + \bar{\Phi}^{N,2}_{S,\mathfrak{t},x,y}\mathrm{d}\mathfrak{t} + \bar{\Phi}^{N,3}_{S,\mathfrak{t},x,y}\mathrm{d}\mathfrak{t}. \tag{4.14b}$$

- The operators $\mathscr{L}^{!!}$ and $\bar{\mathscr{L}}^{!!}$ are operators from the construction of the pair of heat operators $\mathfrak{P}^N$ and $\bar{\mathfrak{P}}^N$, respectively. These operators act on $\mathfrak{J}^N$ and $\bar{\mathfrak{J}}^N$ through the backwards spatial variable $x \in \mathbb{Z}$ and $x \in \mathbb{T}_N$, respectively.

Provided Remark 4.8, we aim to get diffusive tail estimates for fundamental solutions $\mathfrak{J}^N$ and $\bar{\mathfrak{J}}^N$. To this end, throughout this section we use the following exponential weights, which are exponentials of non-negative quantities, to quantify diffusive tails.

**Definition 4.10.** Define $e_{N,x,y} \overset{\bullet}{=} \mathbf{Exp}(N^{-3/4}\mathbf{d}_{x,y})$ and $e^{\mathrm{cpt}}_{N,x,y} \overset{\bullet}{=} \mathbf{Exp}(N^{-3/4}\mathbf{d}^{\mathrm{cpt}}_{x,y})$ where $\mathbf{d}^{\mathrm{cpt}}$ is geodesic distance on the torus $\mathbb{T}_N$ and $\mathbf{d}$ is the usual absolute-value-distance on $\mathbb{Z}$. We also define $\|\,\|_{\omega;p}$ as the $p$-norm with respect to all randomness.

**Proposition 4.11.** *Provided any $p \in \mathbb{R}_{\geqslant 1}$, uniformly in $S, T \in \mathbb{R}_{\geqslant 0}$ such that $S \leqslant T \leqslant 1$, we have*

$$\sup_{x,y\in\mathbb{Z}}e_{N,x,y}\|\mathfrak{J}^N_{S,T,x,y}\|_{\omega;2p} + \sup_{x,y\in\mathbb{T}_N}(e^{\mathrm{cpt}}_{N,x,y})\|\bar{\mathfrak{J}}^N_{S,T,x,y}\|_{\omega;2p} \lesssim \mathbf{Exp}\left(\kappa_p N^{1/2}\right). \tag{4.15}$$

*Recall* $\widetilde{\mathbb{T}}_N \overset{\bullet}{=} [\![ -N^{\frac{5}{4}+\frac{1}{2}\varepsilon_{\text{cpt}}}, N^{\frac{5}{4}+\frac{1}{2}\varepsilon_{\text{cpt}}} ]\!]$ *from* Proposition *4.6. Uniformly in* $x \in \widetilde{\mathbb{T}}_N$ *and* $y \in \mathbb{T}_N$, *we have the comparison estimate*

$$\| \widetilde{\mathfrak{J}}^N_{S,T,x,y} \|_{\omega;2p} \overset{\bullet}{=} \| \mathfrak{J}^N_{S,T,x,y} - \bar{\mathfrak{J}}^N_{S,T,x,y} \|_{\omega;2p} \lesssim_p \mathbf{Exp}\left(-N^{1/2}\right). \tag{4.16}$$

*Remark 4.12.* We note (4.15) and (4.16) agree with Laplace transforms for $\mathfrak{P}^N$ and $\bar{\mathfrak{P}}^N$ and Gaussian kernels of time-scale $N^2$.

*Proof of Proposition 4.6.* For convenience let us first ignore $\chi$ from the initial data of $\bar{\mathfrak{Z}}^N$. By uniqueness of solutions to defining linear equations of $\mathfrak{Z}^N$ and $\bar{\mathfrak{Z}}^N$, we obtain the following fundamental solution representation for the difference $\widetilde{\mathfrak{Z}}^N = \mathfrak{Z}^N - \bar{\mathfrak{Z}}^N$:

$$\widetilde{\mathfrak{Z}}^N_{T,x} = \sum_{y \in \mathbb{Z}} \mathfrak{J}^N_{0,T,x,y} \mathfrak{Z}^N_{0,y} - \sum_{y \in \mathbb{T}_N} \bar{\mathfrak{J}}^N_{0,T,x,y} \mathfrak{Z}^N_{0,y} = \sum_{y \in \mathbb{T}_N} \widetilde{\mathfrak{J}}^N_{0,T,x,y} \mathfrak{Z}^N_{0,y}$$

$$+ \sum_{y \in \mathbb{Z} \setminus \mathbb{T}_N} \mathfrak{J}^N_{0,T,x,y} \mathfrak{Z}^N_{0,y}. \tag{4.17}$$

By Proposition 4.11, the Holder inequality, a priori bounds for near-stationary data, and $|\mathbb{T}_N| \lesssim N^2$, for $p \in \mathbb{R}_{\geqslant 1}$ and $C_1 \in \mathbb{R}_{>0}$ and $x \in \widetilde{\mathbb{T}}_N$ we get the following pair of upper bounds with further explanation that we detail afterwards:

$$\left\| \left( \sum_{y \in \mathbb{T}_N} \widetilde{\mathfrak{J}}^N_{0,T,x,y} \mathfrak{Z}^N_{0,y} \right) \right\|_{\omega;p} \leqslant \sum_{y \in \mathbb{T}_N} \| \widetilde{\mathfrak{J}}^N_{0,T,x,y} \|_{\omega;2p} \| \mathfrak{Z}^N_{0,y} \|_{\omega;2p}$$

$$\lesssim_p \sum_{y \in \mathbb{T}_N} \| \widetilde{\mathfrak{J}}^N_{0,T,x,y} \|_{\omega;2p} \lesssim_{p,C_1} N^{-2C_1}. \tag{4.18}$$

The first bound in (4.18) follows by the triangle inequality and the Holder inequality $\| \varphi \psi \|_{\omega;p} \leqslant \| \varphi \|_{\omega;2p} \| \psi \|_{\omega;2p}$. The second inequality follows from the uniform bound on $\| \mathfrak{Z}^N_{0,\bullet} \|_{\omega;2p}$ from the definition of near-stationary initial data. In the third step, we apply the second estimate in Proposition 4.11 and then we multiply by $|\mathbb{T}_N| \lesssim N^2$ to account for the sum over $y \in \mathbb{T}_N$. Because the exponentially small upper bound from the second estimate in Proposition 4.11 beats out any power of $N \in \mathbb{Z}_{>0}$, we get the far RHS of (4.18) as an upper bound for the far LHS. In particular the aforementioned $|\mathbb{T}_N|$-factor can be ignored to get (4.18).

Because $|\mathbb{T}_N| \gg |\widetilde{\mathbb{T}}_N|$, for $x \in \widetilde{\mathbb{T}}_N$ and $y \notin \mathbb{T}_N$, note $\mathbf{d}_{x,y} \geqslant \kappa N^{5/4+\varepsilon_{\text{cpt}}}$ with $\kappa \in \mathbb{R}_{>0}$ universal, so $e^{-1}_{N,x,y} \leqslant \mathbf{Exp}(-\kappa N^{1/2+\varepsilon_{\text{cpt}}})$. Thus, we get the same upper bound for the second term on the far RHS of (4.17) by using (4.15) which we explain afterwards:

$$\left\| \left( \sum_{y \in \mathbb{Z} \setminus \mathbb{T}_N} \mathfrak{J}^N_{0,T,x,y} \mathfrak{Z}^N_{0,y} \right) \right\|_{\omega;p} \lesssim_p \sum_{y \in \mathbb{Z} \setminus \mathbb{T}_N} \| \mathfrak{J}^N_{0,T,x,y} \|_{\omega;2p}$$

$$\lesssim \sum_{y \in \mathbb{Z} \setminus \mathbb{T}_N} \mathbf{Exp}\left( \kappa_p N^{1/2} \right) e^{-1}_{N,x,y} \lesssim_{p,C_1} N^{-2C_1}. \tag{4.19}$$

The first bound follows again by the triangle inequality for the $\| \|_{\omega;p}$-norm, the Holder inequality $\| \varphi \psi \|_{\omega;p} \leqslant \| \varphi \|_{\omega;2p} \| \psi \|_{\omega;2p}$, and moment bounds for initial data. The second/third bounds follow by the diffusive tail in the first bound in Proposition 4.11 combined with an elementary summation estimate resembling tail probabilities of Brownian motion. This bound also requires the observation that the negative exponent in such tail estimate dwarfs the positive exponent $N^{1/2}$ on the RHS of the first bound in Proposition 4.11 when $x \in \widetilde{\mathbb{T}}_N$ and we sum outside $\mathbb{T}_N$, again since for $x \in \widetilde{\mathbb{T}}_N$ and $y \notin \mathbb{T}_N$, we

have $\mathbf{d}_{x,y} \gtrsim N^{5/4+\varepsilon_{\mathrm{cpt}}}$. Using the last display with (4.17) and (4.18) gives the following where we control a sup by an *unaveraged* sum of absolute values:

$$\||\widetilde{\mathfrak{Z}}_{T,x}^{N}\|_{\mathfrak{t}_{\mathfrak{f}};\widetilde{\mathbb{T}}_{N}}\|_{\omega;p} \leqslant \|\Big(\sum\nolimits_{\mathfrak{I}_{\mathfrak{t}_{\mathfrak{f}}}\times\widetilde{\mathbb{T}}_{N}}|\widetilde{\mathfrak{Z}}_{T,x}^{N}|\Big)\|_{\omega;p} \leqslant \sum\nolimits_{\mathfrak{I}_{\mathfrak{t}_{\mathfrak{f}}}\times\widetilde{\mathbb{T}}_{N}}\|\widetilde{\mathfrak{Z}}_{T,x}^{N}\|_{\omega;p}$$
$$\lesssim_{p,C_1} |\mathfrak{I}_{\mathfrak{t}_{\mathfrak{f}}}||\widetilde{\mathbb{T}}_{N}|N^{-2C_1}. \tag{4.20}$$

Combining the estimate (4.20) with the Chebyshev inequality gives the desired estimate (4.6) if we choose $C_2 = C_1 - 103$ and then choose $C_1 \in \mathbb{R}_{>0}$ sufficiently/arbitrarily large but universal for example. To make this precise, the Markov inequality and (4.20) with $p = 1$ gives the following estimate in which all implied constants depend only on $C_1$ and $C_2 = C_1 - 103$:

$$\mathbf{P}\Big(\|\widetilde{\mathfrak{Z}}_{T,x}^{N}\|_{\mathfrak{t}_{\mathfrak{f}};\widetilde{\mathbb{T}}_{N}} \gtrsim N^{-C_1}\Big) \lesssim N^{C_1}\||\widetilde{\mathfrak{Z}}_{T,x}^{N}\|_{\mathfrak{t}_{\mathfrak{f}};\widetilde{\mathbb{T}}_{N}}\|_{\omega;1}$$
$$\lesssim N^{C_1}|\mathfrak{I}_{\mathfrak{t}_{\mathfrak{f}}}||\widetilde{\mathbb{T}}_{N}|N^{-2C_1} \lesssim N^{-C_1}N^{100+2} \lesssim N^{-C_2}. \tag{4.21}$$

We choose $C_1 \in \mathbb{R}_{>0}$ sufficiently/arbitrarily large but universal to complete the proof in the case where we ignore $\chi$/set $\chi \equiv 1$. The error we get when we ignore $\chi$ comes from the initial data $\bar{\mathfrak{Z}}_{0,\bullet}^{N}$. It is ultimately the following error in replacing $\chi$ by 1:

$$|\mathrm{Error}| \leqslant \sum\nolimits_{y\in\mathbb{T}_{N}}|\bar{\mathfrak{Z}}_{0,T,x,y}^{N}||\chi_y - 1|\mathfrak{Z}_{0,y}^{N}. \tag{4.22}$$

Via $|\mathbb{T}_{N}| \gg |\widetilde{\mathbb{T}}_{N}|$ and definition of $\chi$ in Definition 4.3, the support of $\chi_y - 1$ is $\kappa N^{5/4+\varepsilon_{\mathrm{cpt}}}$ away from $x \in \widetilde{\mathbb{T}}_{N}$ for $\kappa \in \mathbb{R}_{>0}$ universal; we made a similar observation prior to (4.19). We also have $|\chi_y - 1| \leqslant 2$, so we may use tail bounds for $\bar{\mathfrak{Z}}^{N}$ in the first bound in Proposition 4.11 and a priori bounds for near-stationary initial data, like in our proof of (4.19 ), to account for ignoring $\chi$. □

### 4.2. Proof of Proposition 4.11.

The first step is a short-time moment estimate with a proof almost identical to that of Proposition 3.2 in [10]. The short-time nature allows us to control both $\mathfrak{Z}^{N}$ and $\bar{\mathfrak{Z}}^{N}$ on exponential scales $e_{N,x,y}$ and $e_{N,x,y}^{\mathrm{cpt}}$, respectively.

**Lemma 4.13.** *Take* $S, T \in \mathbb{R}_{\geqslant 0}$ *with* $S \leqslant T \leqslant 1$ *and define* $\mathfrak{t}_{\mathrm{short}} \overset{\bullet}{=} N^{-1/2}$. *For* $p \in \mathbb{R}_{\geqslant 1}$, *we have, recalling* Definition *4.10,*

$$\sup\nolimits_{0\leqslant\mathfrak{t}\leqslant\mathfrak{t}_{\mathrm{short}}}\sup\nolimits_{x,y\in\mathbb{Z}}e_{N,x,y}\|\mathfrak{Z}_{S,T+\mathfrak{t},x,y}^{N}\|_{\omega;2p} \lesssim_p \sup\nolimits_{x,y\in\mathbb{Z}}e_{N,x,y}\|\mathfrak{Z}_{S,T,x,y}^{N}\|_{\omega;2p} \tag{4.23a}$$
$$\sup\nolimits_{0\leqslant\mathfrak{t}\leqslant\mathfrak{t}_{\mathrm{short}}}\sup\nolimits_{x,y\in\mathbb{T}_{N}}e_{N,x,y}^{\mathrm{cpt}}\|\bar{\mathfrak{Z}}_{S,T+\mathfrak{t},x,y}^{N}\|_{\omega;2p} \lesssim_p \sup\nolimits_{x,y\in\mathbb{T}_{N}}e_{N,x,y}^{\mathrm{cpt}}\|\bar{\mathfrak{Z}}_{S,T,x,y}^{N}\|_{\omega;2p}. \tag{4.23b}$$

We defer proof of Lemma 4.13 to the end of the section. The first bound in Proposition 4.11 will come from iterating Lemma 4.13. We now turn to the second ingredient for the proof of Proposition 4.11. In particular, for (4.16) we set up a Duhamel-type formula that gives an explicit relationship for $\bar{\mathfrak{Z}}^{N}$ and $\mathfrak{Z}^{N}$ that is convenient to analyze with tail estimates for $\bar{\mathfrak{Z}}^{N}$ and $\mathfrak{Z}^{N}$.

**Lemma 4.14.** *Provided any $S$, $T \in \mathbb{R}_{\geqslant 0}$ satisfying $S \leqslant T$ and provided any $x$, $y \in \mathbb{T}_N$, we have, with notation explained after,*

$$
\bar{\mathfrak{J}}^N_{S,T,x,y} = \mathfrak{J}^N_{S,T,x,y}
$$
$$
+ \int_S^T \sum_{w \in \mathbb{T}_N} \bar{\mathfrak{J}}^N_{R,T,x,w} \cdot N^2 \sum_{|\mathfrak{l}| \leqslant 10 N^{\beta_X} \mathfrak{m}} \left( \mathfrak{f}^{\mathfrak{l},1}_{R,w} \mathfrak{J}^N_{S,R,w+\mathfrak{l},y} + \mathfrak{f}^{\mathfrak{l},2}_{R,w} \mathfrak{J}^N_{S,R,w\widetilde{+}\mathfrak{l},y} \right) \mathrm{d}R.
$$
(4.24)

*The functionals $\mathfrak{f}^{\mathfrak{l},i}$ have supports outside the ball $\mathfrak{B}$ of radius $N^{\frac{5}{4}+\frac{2}{3}\varepsilon_{\mathrm{cpt}}}$ around $0 \in \mathbb{Z}$ and they are uniformly bounded. Moreover $+$ denotes addition on $\mathbb{Z}$ and $\widetilde{+}$ denotes addition on the torus $\mathbb{T}_N$ in the statement and proof of this result.*

*Proof of Proposition 4.11.* As we noted after the statement of Lemma 4.13, the estimate (4.15) follows by its validity at $T = S$, which follows almost entirely by definition, then iterating Lemma 4.13 and gathering factors a total of order $N^{1/2}$-many times. To prove (4.16), recall $\beta_X \leqslant 1$ and $S$, $T \leqslant 1$. By Lemma 4.14 and the Holder inequality $\|\varphi \psi\|_{\omega;p} \leqslant \|\varphi\|_{\omega;2p} \|\psi\|_{\omega;2p}$, we get

$$
\|\widetilde{\mathfrak{J}}^N_{S,T,x,y}\|_{\omega;p} \lesssim_{\mathfrak{m}} N^3 \sup_{S \leqslant R \leqslant T} \left( \sum_{w \in \mathbb{T}_N} \mathbf{1}_{w \notin \mathfrak{B}} \|\widetilde{\mathfrak{J}}^N_{R,T,x,w}\|_{\omega;2p} \cdot \sup_{w \in \mathbb{Z}} \|\mathfrak{J}^N_{S,R,w,y}\|_{\omega;2p} \right).
$$
(4.25)

We emphasize that the cutoff indicator function inside the supremum on the RHS of the previous display follows by the support constraints on the $\mathfrak{f}^{\mathfrak{l},1}$ and $\mathfrak{f}^{\mathfrak{l},2}$ functionals from the statement of Lemma 4.14. We assume $S$, $T \in \mathbb{R}_{\geqslant 0}$ satisfy $S \leqslant T \leqslant 1$ and take any $x \in \widetilde{\mathbb{T}}_N$ and $y \in \mathbb{T}_N$. For this case, we observe if $w \notin \mathfrak{B}$ then it satisfies $|w| \gtrsim N^{5/4+2\varepsilon_{\mathrm{cpt}}/3}$, thus $\mathbf{d}^{\mathrm{cpt}}_{x,w} \gtrsim N^{5/4+\varepsilon_{\mathrm{cpt}}/2}$, so for these points $w \notin \mathfrak{B}$ the diffusive tail estimate (4.15) yields the following estimate which says the diffusive tails beat the exponentially growing factor on the RHS of (4.15) thus rendering such exponentially growing factor asymptotically irrelevant. Here $\kappa \gtrsim 1$:

$$
\mathbf{1}_{w \notin \mathfrak{B}} \|\bar{\mathfrak{J}}^N_{R,T,x,w}\|_{\omega;2p} \lesssim \mathbf{Exp}\left(\kappa_p N^{1/2}\right) \mathbf{1}_{w \notin \mathfrak{B}} e^{-1}_{N,x,y}
$$
$$
\lesssim \mathbf{Exp}\left(\kappa_p N^{1/2}\right) \mathbf{Exp}\left(-\kappa N^{1/2+\varepsilon_{\mathrm{cpt}}/2}\right).
$$
(4.26)

We now use (4.15) and uniform boundedness of $e^{-1}_{N,\bullet,\bullet}$-factors to obtain the following estimate with universal implied constant. Roughly speaking it forgets the diffusive off-diagonal/tail behavior of $\mathfrak{J}^N$ and uses only the exponential control for its growth. As we note again shortly such exponential control is dwarfed by the negative-exponent factor on the far RHS of (4.26):

$$
\|\mathfrak{J}^N_{S,R,w,y}\|_{\omega;2p} \lesssim \mathbf{Exp}\left(\kappa_p N^{1/2}\right).
$$
(4.27)

We plug (4.26) and (4.27) into (4.25) to establish the following for which we recall $|\mathbb{T}_N| \lesssim N^2$ and then bound the supremum from the RHS (4.25) by $|\mathbb{T}_N|$ times the space-time supremum of the product of $\|\;\|_{\omega;2p}$-norms. We ultimately obtain the estimate below in which the final conclusion is that the negative-exponent factor on the far RHS of (4.26) dwarfs all other factors:

$$
\|\widetilde{\mathfrak{J}}^N_{S,T,x,y}\|_{\omega;p} \lesssim_{\mathfrak{m}} N^3 |\mathbb{T}_N| \mathbf{Exp}\left(2\kappa_p N^{1/2}\right) \mathbf{Exp}\left(-\kappa N^{1/2+\varepsilon_{\mathrm{cpt}}/2}\right)
$$
$$
\lesssim_p \mathbf{Exp}\left(-N^{1/2+\varepsilon_{\mathrm{cpt}}/3}\right) \leqslant \mathbf{Exp}\left(-N^{1/2}\right).
$$
(4.28)

The second estimate on the far RHS of the previous display follows by elementary asymptotics in $N \in \mathbb{Z}_{>0}$ as $\kappa_p \lesssim_p 1$. $\qquad\qquad\qquad\qquad\qquad\qquad\square$

We start the proof of the auxiliary results in Lemma 4.13 and Lemma 4.14 with notation that captures diffusive tail behavior of the heat kernels $\mathfrak{P}^N$ and $\bar{\mathfrak{P}}^N$ and that will be used in the proof of Lemma 4.13. Let us first recall Definition 4.10.

**Definition 4.15.** Given any $\kappa \in \mathbb{R}_{>0}$, any $S, T \in \mathbb{R}_{\geqslant 0}$ so that $S \leqslant T$, and any $x, y \in \mathbb{Z}$ or $\mathbb{T}_N$, define the following exponential weights in which $\mathbf{d}_{x,y} = |x - y|$ is usual distance on $\mathbb{Z}$, in which $\mathfrak{s}_{S,T} = |T - S|$, and in which $\mathbf{d}^{\mathrm{cpt}}$ is geodesic distance on $\mathbb{T}_N$:

$$e_{S,T,x,y}^{N,\kappa} \; \overset{\bullet}{=} \; \mathbf{Exp}\left(\kappa \frac{\mathbf{d}_{x,y}}{N \mathfrak{s}_{S,T}^{1/2} \vee 1}\right), \quad e_{S,T,x,y}^{N,\kappa,\mathrm{cpt}} \; \overset{\bullet}{=} \; \mathbf{Exp}\left(\kappa \frac{\mathbf{d}_{x,y}^{\mathrm{cpt}}}{N \mathfrak{s}_{S,T}^{1/2} \vee 1}\right). \qquad (4.29)$$

We will require later for the proof of Lemma 4.13 the following elementary summation estimate, which follows by a geometric series bound, and the following elementary control on $e_{N,x,y}$ and $e_{N,x,y}^{\mathrm{cpt}}$ for short times. We take $\kappa \in \mathbb{R}_{>0}$ strictly positive:

$$(N \mathfrak{s}_{S,T}^{1/2} \vee 1)^{-1} \left(\sum_{w \in \mathbb{Z}} e_{S,T,x,w}^{N,-\kappa} + \sum_{w \in \mathbb{T}_N} e_{S,T,x,w}^{N,-\kappa,\mathrm{cpt}}\right) \lesssim_\kappa 1 \qquad (4.30\mathrm{a})$$

$$\mathbf{1}_{\mathfrak{t} \leqslant N^{-1/2}} \left(e_{0,\mathfrak{t},x,y}^{N,-2} e_{N,x,y}^2 + e_{0,\mathfrak{t},x,y}^{N,-2,\mathrm{cpt}} (e_{N,x,y}^{\mathrm{cpt}})^2\right) \lesssim 1. \qquad (4.30\mathrm{b})$$

*Proof of Lemma 4.13.* We use notation of Definitions 4.10, 4.15; recall $\mathfrak{t}_{\mathrm{short}} = N^{-1/2}$. We note the following estimates for both $\mathfrak{P}^N$ and $\bar{\mathfrak{P}}^N$ which follow by Lemma A.1 as we explain soon. Just for this proof, we define $\mathfrak{t}_N \overset{\bullet}{=} N^2 \mathfrak{t} \vee 1$ as a function of $\mathfrak{t} \in \mathbb{R}_{\geqslant 0}$. The following estimates effectively give a short-time scaling of heat kernels for random walks with speed $N^2$ with exponential off-diagonal weight that is of diffusive type. In the following pair of estimates 8 and 2 are just arbitrary positive numbers:

$$\sup_{0 \leqslant \mathfrak{t} \leqslant \mathfrak{t}_{\mathrm{short}}} \sup_{x,y \in \mathbb{Z}} \mathfrak{t}_N^{\frac{1}{2}} e_{N,x,y}^2 e_{0,\mathfrak{t},x,y}^{N,8} \mathfrak{P}_{0,\mathfrak{t},x,y}^N d + \sup_{0 \leqslant \mathfrak{t} \leqslant \mathfrak{t}_{\mathrm{short}}} \sup_{x,y \in \mathbb{T}_N} \mathfrak{t}_N^{\frac{1}{2}} (e_{N,x,y}^{\mathrm{cpt}})^2 e_{0,\mathfrak{t},x,y}^{N,8,\mathrm{cpt}} \bar{\mathfrak{P}}_{0,\mathfrak{t},x,y}^N$$

$$\lesssim 1 \qquad (4.31\mathrm{a})$$

$$\sup_{0 \leqslant \mathfrak{t} \leqslant \mathfrak{t}_{\mathrm{short}}} \sup_{x,y \in \mathbb{Z}} \mathfrak{t}_N e_{N,x,y}^2 e_{0,\mathfrak{t},x,y}^{N,8} \sup_{|k| \lesssim \mathfrak{m}} |\nabla_k^! \mathfrak{P}_{0,\mathfrak{t},x,y}^N|$$

$$+ \sup_{0 \leqslant \mathfrak{t} \leqslant \mathfrak{t}_{\mathrm{short}}} \sup_{x,y \in \mathbb{T}_N} \mathfrak{t}_N (e_{N,x,y}^{\mathrm{cpt}})^2 e_{0,\mathfrak{t},x,y}^{N,8,\mathrm{cpt}} \sup_{|k| \lesssim \mathfrak{m}} |\bar{\nabla}_k^! \bar{\mathfrak{P}}_{0,\mathfrak{t},x,y}^N| \lesssim_{\mathfrak{m}} N. \qquad (4.31\mathrm{b})$$

Indeed, for $0 \leqslant \mathfrak{t} \leqslant \mathfrak{t}_{\mathrm{short}}$, by the off-diagonal bound (A.1) for $\kappa = 10$ and the bound $e_{0,\mathfrak{t},x,y}^{N,-2} \lesssim e_{N,x,y}^{-2}$ from Definition 4.15,

$$\mathfrak{t}_N^{1/2} \mathfrak{P}_{0,\mathfrak{t},x,y}^N \; \lesssim \; e_{0,\mathfrak{t},x,y}^{N,-10} \; \lesssim \; e_{N,x,y}^{-2} e_{0,\mathfrak{t},x,y}^{N,-8}. \qquad (4.32)$$

This provides (4.31a); for (4.31b) we instead use (A.2) in Lemma A.1 and $\mathfrak{m} \lesssim 1$. We recall that $\bar{\mathfrak{P}}^N$ satisfies the same estimates for $\mathfrak{t} \lesssim 1$ although with respect to the geodesic distance $\mathbf{d}^{\mathrm{cpt}}$, so we get both estimates in (4.31a) and (4.31b).

We will now prove (4.23a). The proof of (4.23b) follows from identical methods upon elementary replacements like replacing $\mathbb{Z}$ by $\mathbb{T}_N$. We will basically copy the proof of Proposition 3.2 in [10]. By definition of $\mathfrak{J}^N$ in Definition 4.9, we get

$$\|\mathfrak{J}_{S,T+\mathfrak{t},x,y}^N\|_{\omega;2p}^2 \lesssim \|\mathfrak{P}_{\mathfrak{t},x}^{N,\mathbf{X}} \mathfrak{J}_{S,T,\bullet,y}^N\|_{\omega;2p}^2 + \|\mathfrak{P}_{\mathfrak{t},x}^N (\mathfrak{J}_{S,T+\bullet,\bullet,y}^N \mathrm{d}\xi^N)\|_{\omega;2p}^2$$

$$+ \|\mathfrak{P}_{\mathfrak{t},x}^N (\Phi_{S,T+\bullet,\bullet,y}^{N,2})\|_{\omega;2p}^2 + \|\mathfrak{P}_{\mathfrak{t},x}^N (\Phi_{S,T+\bullet,\bullet,y}^{N,3})\|_{\omega;2p}^2. \qquad (4.33)$$

Following the proof of Proposition 3.2 in [10], the estimate (4.31a) and the elementary inequality $e_{N,x,y} \leqslant e_{N,x,w} e_{N,w,y}$, which follows by the triangle inequality, and the $\mathbb{Z}$-sum bound (4.30a) for $S = 0$ and $\kappa = 8$ from Definition 4.15 and convexity of $\mathfrak{P}^{N,\mathbf{X}}$ give the following spatial-heat-operator bound; recall $\mathfrak{t}_N = N^2 \mathfrak{t} \vee 1$ from earlier in this proof and $e$-weights in Definition 4.10:

$$e_{N,x,y}^2 \|\mathfrak{P}_{\mathfrak{t},x}^{N,\mathbf{X}} \mathfrak{I}_{S,T,\bullet,y}^N\|_{\omega;2p}^2 \leqslant e_{N,x,y}^2 \sum_{w \in \mathbb{Z}} \mathfrak{P}_{0,\mathfrak{t},x,w}^N \cdot \|\mathfrak{I}_{S,T,w,y}^N\|_{\omega;2p}^2$$

$$\lesssim \sum_{w \in \mathbb{Z}} e_{N,x,w}^2 \mathfrak{P}_{0,\mathfrak{t},x,w}^N \cdot e_{N,w,y}^2 \|\mathfrak{I}_{S,T,w,y}^N\|_{\omega;2p}^2 \tag{4.34}$$

$$\lesssim \mathfrak{t}_N^{-1/2} \sum_{w \in \mathbb{Z}} e_{0,\mathfrak{t},x,w}^{N,-8} \cdot \sup_{w \in \mathbb{Z}} e_{N,w,y}^2 \|\mathfrak{I}_{S,T,w,y}^N\|_{\omega;2p}^2 \tag{4.35}$$

$$\lesssim \sup_{w \in \mathbb{Z}} e_{N,w,y}^2 \|\mathfrak{I}_{S,T,w,y}^N\|_{\omega;2p}^2. \tag{4.36}$$

We now want to use the martingale bound Lemma 3.1 in [10] as with the proof of Proposition 3.2 in [10]. However, since the martingale inequality of Lemma 3.1 from [10] is specific to the microscopic Cole-Hopf transform $\mathfrak{Z}^N$, and in particular does not generalize to arbitrary products between any adapted process and the martingale increment $\mathrm{d}\xi^N$ in (2.4) of [10], we instead use Lemma B.1 to get the first line below; the lines after follow by $e_{N,x,y}^2 e_{N,w,y}^{-2} \leqslant e_{N,x,w}^2$ and (4.31a) and $\mathfrak{t}_N^{-1/2} \sum_{w \in \mathbb{Z}} e_{0,\mathfrak{t},x,w}^{N,-8} \lesssim 1$:

$$e_{N,x,y}^2 \|\mathfrak{P}_{\mathfrak{t},x}^N(\mathfrak{I}_{S,T+\bullet,\bullet,y}^N \mathrm{d}\xi^N)\|_{\omega;2p}^2$$

$$\lesssim_p e_{N,x,y}^2 \int_0^{\mathfrak{t}} \left( \sup_{0 \leqslant \mathfrak{r} \leqslant \mathfrak{t}} \sup_{w \in \mathbb{Z}} e_{N,w,y}^2 \|\mathfrak{I}_{S,T+\mathfrak{r},w,y}^N\|_{\omega;2p}^2 \right) \mathfrak{s}_{R,\mathfrak{t}}^{-\frac{1}{2}} \sum_{w \in \mathbb{Z}} \mathfrak{P}_{R,\mathfrak{t},x,w}^N e_{N,w,y}^{-2} \, \mathrm{d}R \tag{4.37}$$

$$\lesssim_p \int_0^{\mathfrak{t}} \mathfrak{s}_{R,\mathfrak{t}}^{-\frac{1}{2}} \left( \sup_{0 \leqslant \mathfrak{r} \leqslant \mathfrak{t}} \sup_{w \in \mathbb{Z}} e_{N,w,y}^2 \|\mathfrak{I}_{S,T+\mathfrak{r},w,y}^N\|_{\omega;2p}^2 \right) \sum_{w \in \mathbb{Z}} e_{N,x,w}^2 \mathfrak{P}_{R,\mathfrak{t},x,w}^N \, \mathrm{d}R \tag{4.38}$$

$$\lesssim_p \int_0^{\mathfrak{t}} \mathfrak{s}_{R,\mathfrak{t}}^{-\frac{1}{2}} \left( \sup_{0 \leqslant \mathfrak{r} \leqslant \mathfrak{t}} \sup_{w \in \mathbb{Z}} e_{N,w,y}^2 \|\mathfrak{I}_{S,T+\mathfrak{r},w,y}^N\|_{\omega;2p}^2 \right) (\mathfrak{s}_{R,\mathfrak{t}})_N^{-\frac{1}{2}} \sum_{w \in \mathbb{Z}} e_{R,\mathfrak{t},x,w}^{N,-8} \, \mathrm{d}R \tag{4.39}$$

$$\lesssim \int_0^{\mathfrak{t}} \mathfrak{s}_{R,\mathfrak{t}}^{-\frac{1}{2}} \left( \sup_{0 \leqslant \mathfrak{r} \leqslant \mathfrak{t}} \sup_{w \in \mathbb{Z}} e_{N,w,y}^2 \|\mathfrak{I}_{S,T+\mathfrak{r},w,y}^N\|_{\omega;2p}^2 \right) \mathrm{d}R. \tag{4.40}$$

We have used the notation $\mathfrak{s}_{S,R}^{-1/2} \overset{\bullet}{=} |S - R|^{-1/2}$ within Definition 4.15. We also clarify our earlier notation $(\mathfrak{s}_{R,\mathfrak{t}})_N = N^2 \mathfrak{s}_{R,\mathfrak{t}} \vee 1$. For the third term on the RHS of (4.33), we first observe that the heat kernel $\mathfrak{P}^N$ is a probability measure on $\mathbb{Z}$ with respect to the forwards spatial-variable. We then apply the Cauchy-Schwarz inequality with respect to the space-time convolution "integral" via heat kernel. As $\mathfrak{t} \lesssim N^{-1/2}$, calculations similar to the derivations of (4.36) and (4.40) give, for $\Psi$ defined shortly, the bounds

$$e_{N,x,y}^2 \|\mathfrak{P}_{t,x}^N(\Phi_{S,T+\bullet,\bullet,y}^{N,2})\|_{\omega;2p}^2$$

$$\lesssim e_{N,x,y}^2 N\mathfrak{t} \int_0^{\mathfrak{t}} \sum_{w\in\mathbb{Z}} \mathfrak{P}_{R,\mathfrak{t},x,w}^N \cdot \|\mathfrak{I}_{S,T+R,w,y}^N\|_{\omega;2p}^2 \, \mathrm{d}R + \Psi \tag{4.41}$$

$$\lesssim N^{\frac{1}{2}} \int_0^{\mathfrak{t}} \sum_{w\in\mathbb{Z}} e_{N,x,w}^2 \mathfrak{P}_{R,\mathfrak{t},x,w}^N \cdot e_{N,w,y}^2 \|\mathfrak{I}_{S,T+R,w,y}^N\|_{\omega;2p}^2 \, \mathrm{d}R + \Psi \tag{4.42}$$

$$\lesssim N^{\frac{1}{2}} \int_0^{\mathfrak{t}} (\mathfrak{s}_{R,\mathfrak{t}})_N^{-\frac{1}{2}} \sum_{w\in\mathbb{Z}} e_{R,\mathfrak{t},x,w}^{N,-8} \cdot e_{N,w,y}^2 \|\mathfrak{I}_{S,T+R,w,y}^N\|_{\omega;2p}^2 \, \mathrm{d}R + \Psi \tag{4.43}$$

$$\lesssim N^{\frac{1}{2}} \int_0^{\mathfrak{t}} \sup_{w\in\mathbb{Z}} e_{N,w,y}^2 \|\mathfrak{I}_{S,T+R,w,y}^N\|_{\omega;2p}^2 \, \mathrm{d}R + \Psi. \tag{4.44}$$

The last quantity $\Psi$ is slightly more subtle and in particular unaddressed in Proposition 3.2 in [10]. It is defined as

$$\Psi \overset{\bullet}{=} e_{N,x,y}^2 \mathfrak{t} \int_0^{\mathfrak{t}} \sum_{w\in\mathbb{Z}} \mathfrak{P}_{R,\mathfrak{t},x,w}^N \cdot \widetilde{\sum}_{\mathfrak{l}=1}^{N^{\beta_X}} N\|\nabla_{-7\mathfrak{lm}}(\mathfrak{b}_{T+R,w}^{\mathfrak{l}} \mathfrak{I}_{S,T+R,w,y}^N)\|_{\omega;2p}^2 \, \mathrm{d}S. \tag{4.45}$$

We will not perform summation-by-parts for the $\Psi$ quantity. We instead expand the gradient while recalling the $\mathfrak{b}^{\mathfrak{l}}$-functionals are uniformly bounded. This bounds the $\|\|_{\omega;2p}$-term in the $\Psi$-definition by the same norm of $\mathfrak{I}^N$ evaluated at points $7\mathfrak{lm}$-apart. Noting $e_{N,w,y}^2 \lesssim e_{N,w,w+\mathfrak{k}}^2 e_{N,w+\mathfrak{k},y}^2 \lesssim e_{N,w+\mathfrak{k},y}^2$ for all $\mathfrak{k} \in \mathbb{Z}$ with $|\mathfrak{k}| \lesssim N^{\beta_X}$, where $\beta_X = \frac{1}{3} + \varepsilon_{X,1} \leqslant \frac{1}{2}$, and recalling $\mathfrak{t} \lesssim N^{-1/2}$, this paragraph gives the following when combined with the same considerations as in derivations of (4.36), (4.40), and (4.44):

$$\Psi \lesssim N\mathfrak{t} \int_0^{\mathfrak{t}} \sum_{w\in\mathbb{Z}} e_{N,x,w}^2 \mathfrak{P}_{R,\mathfrak{t},x,w}^N \cdot \widetilde{\sum}_{\mathfrak{l}=1}^{N^{\beta_X}} e_{N,w,y}^2 \||\mathfrak{I}_{S,T+R,w,y}^N|$$
$$+ |\mathfrak{I}_{S,T+R,w-7\mathfrak{lm},y}^N|\|_{\omega;2p}^2 \, \mathrm{d}R \tag{4.46}$$

$$\lesssim N^{\frac{1}{2}} \int_0^{\mathfrak{t}} (\mathfrak{s}_{R,\mathfrak{t}})_N^{-\frac{1}{2}} \sum_{w\in\mathbb{Z}} e_{R,\mathfrak{t},x,w}^{N,-8} \cdot \widetilde{\sum}_{\mathfrak{l}=1}^{N^{\beta_X}} e_{N,w,y}^2 \||\mathfrak{I}_{S,T+R,w,y}^N|$$
$$+ |\mathfrak{I}_{S,T+R,w-7\mathfrak{lm},y}^N|\|_{\omega;2p}^2 \, \mathrm{d}R \tag{4.47}$$

$$\lesssim N^{\frac{1}{2}} \int_0^{\mathfrak{t}} (\mathfrak{s}_{R,\mathfrak{t}})_N^{-\frac{1}{2}} \sum_{w\in\mathbb{Z}} e_{R,\mathfrak{t},x,w}^{N,-8} \cdot \widetilde{\sum}_{\mathfrak{l}=1}^{N^{\beta_X}}$$
$$\left( e_{N,w,y}^2 \|\mathfrak{I}_{S,T+R,w,y}^N\|_{\omega;2p} + e_{N,w-7\mathfrak{lm},y}^2 \|\mathfrak{I}_{S,T+R,w-7\mathfrak{lm},y}^N\|_{\omega;2p}^2 \right) \, \mathrm{d}R \tag{4.48}$$

$$\lesssim N^{\frac{1}{2}} \int_0^{\mathfrak{t}} \sup_{w\in\mathbb{Z}} e_{N,w,y}^2 \|\mathfrak{I}_{S,T+R,w,y}^N\|_{\omega;2p}^2 \, \mathrm{d}R. \tag{4.49}$$

Lastly, we treat the final term from the RHS of (4.33). Following the weakly vanishing-term and gradient-term estimates from the proof of Proposition 3.2 in [10], which resemble the proof of (4.44) but with an additional summation-by-parts, we get

$$e_{N,x,y}^2 \| \mathfrak{P}_{\mathfrak{t},x}^N(\Phi_{S,T+\bullet,\bullet,y}^{N,3}) \|_{\omega;2p}^2$$

$$\lesssim \int_0^{\mathfrak{t}} \sum_{w \in \mathbb{Z}} e_{N,x,w}^2 \left( \mathfrak{P}_{R,\mathfrak{t},x,w}^N + \sup_{|k| \lesssim \mathfrak{m}} |\nabla_k^! \mathfrak{P}_{R,\mathfrak{t},x,w}^N| \right) \cdot e_{N,w,y}^2 \| \mathfrak{I}_{S,T+R,w,y}^N \|_{\omega;2p}^2 \, dR \qquad (4.50)$$

$$\lesssim \int_0^{\mathfrak{t}} \sup_{w \in \mathbb{Z}} e_{N,w,y}^2 \| \mathfrak{I}_{S,T+R,w,y}^N \|_{\omega;2p}^2 \sum_{w \in \mathbb{Z}} e_{N,x,w}^2 \left( \mathfrak{P}_{R,\mathfrak{t},x,w}^N + \sup_{|k| \lesssim \mathfrak{m}} |\nabla_k^! \mathfrak{P}_{R,\mathfrak{t},x,w}^N| \right) dR \qquad (4.51)$$

$$\lesssim \int_0^{\mathfrak{t}} \sup_{w \in \mathbb{Z}} e_{N,w,y}^2 \| \mathfrak{I}_{S,T+R,w,y}^N \|_{\omega;2p}^2 \cdot \left( (\mathfrak{s}_{R,\mathfrak{t}})_N^{-\frac{1}{2}} \sum_{w \in \mathbb{Z}} e_{R,\mathfrak{t},x,w}^{N,-8} + N(\mathfrak{s}_{R,\mathfrak{t}})_N^{-1} \sum_{w \in \mathbb{Z}} e_{R,\mathfrak{t},x,w}^{N,-8} \right) dR \qquad (4.52)$$

$$\lesssim \int_0^{\mathfrak{t}} \left( 1 + N(\mathfrak{s}_{R,\mathfrak{t}})_N^{-\frac{1}{2}} \right) \sup_{w \in \mathbb{Z}} e_{N,w,y}^2 \| \mathfrak{I}_{S,T+R,w,y}^N \|_{\omega;2p}^2 \, dR \qquad (4.53)$$

$$\lesssim \int_0^{\mathfrak{t}} \mathfrak{s}_{R,\mathfrak{t}}^{-\frac{1}{2}} \sup_{w \in \mathbb{Z}} e_{N,w,y}^2 \| \mathfrak{I}_{S,T+R,w,y}^N \|_{\omega;2p}^2 \, dR. \qquad (4.54)$$

The previous display additionally requires (4.31b) to treat the gradient of the heat kernel that we have not touched yet for this proof along with the estimate $\mathfrak{t}_N^{-1/2} \sum_{w \in \mathbb{Z}} e_{0,\mathfrak{t},x,w}^{N,-8} \lesssim 1$ that we have been using frequently for this proof. We also use the estimate $(\mathfrak{s}_{R,\mathfrak{t}})_N \gtrsim N^2 \mathfrak{s}_{R,\mathfrak{t}}$ which follows by definition and the estimate $1 \geqslant \mathfrak{s}_{R,\mathfrak{t}}$ for $R \leqslant \mathfrak{t} \leqslant 1$ to get the last bound (4.54). Otherwise, we proceed as in the derivations of (4.36), (4.40), (4.44), and (4.49) to arrive at (4.54) above. We now define

$$\Phi_R \overset{\bullet}{=} \sup_{w \in \mathbb{Z}} e_{N,w,y}^2 \| \mathfrak{I}_{S,T+R,w,y}^N \|_{\omega;2p}^2. \qquad (4.55)$$

We gather estimates in (4.33), (4.36), (4.40), (4.44), (4.49), and (4.54) to deduce

$$\Phi_{\mathfrak{t}} \lesssim_p \Phi_0 + \int_0^{\mathfrak{t}} \mathfrak{s}_{R,\mathfrak{t}}^{-\frac{1}{2}} \sup_{0 \leqslant \mathfrak{r} \leqslant \mathfrak{t}} \Phi_{\mathfrak{r}} \, dR + \int_0^{\mathfrak{t}} \mathfrak{s}_{R,\mathfrak{t}}^{-\frac{1}{2}} \Phi_R \, dR + N^{\frac{1}{2}} \int_0^{\mathfrak{t}} \Phi_R \, dR$$

$$\lesssim \Phi_0 + \mathfrak{t}^{\frac{1}{2}} \sup_{0 \leqslant \mathfrak{r} \leqslant \mathfrak{t}} \Phi_{\mathfrak{r}} + \int_0^{\mathfrak{t}} \mathfrak{s}_{R,\mathfrak{t}}^{-\frac{1}{2}} \Phi_R \, dR + N^{\frac{1}{2}} \int_0^{\mathfrak{t}} \Phi_R \, dR. \qquad (4.56)$$

The first integral on the far RHS of (4.56) admits a singular factor in the integrand, but it is integrable so that we can still apply the Gronwall inequality. The Gronwall inequality ultimately gives the following uniformly in $0 \leqslant \mathfrak{t} \leqslant N^{-1/2}$:

$$\sup_{0 \leqslant \mathfrak{t} \leqslant N^{-1/2}} \Phi_{\mathfrak{t}} \lesssim_p \Phi_0 + \sup_{0 \leqslant \mathfrak{t} \leqslant N^{-1/2}} \mathfrak{t}^{\frac{1}{2}} \sup_{0 \leqslant \mathfrak{r} \leqslant \mathfrak{t}} \Phi_{\mathfrak{r}} \lesssim \Phi_0 + N^{-\frac{1}{4}} \sup_{0 \leqslant \mathfrak{r} \leqslant N^{-1/2}} \Phi_{\mathfrak{r}}. \qquad (4.57)$$

We complete the proof upon moving the second term on the far RHS to the far LHS. □

*Proof of Lemma 4.14.* We proceed as with classical parabolic semigroups and check the RHS of the claimed identity solves the defining SDE-type stochastic equation for $\bar{\mathfrak{J}}^N$ given an appropriate choice of functionals. This would provide the desired result as the initial data at $T = S$ match on both sides of the proposed identity, from which

we employ the elementary uniqueness for solutions to the linear SDE-type equation for $\tilde{\mathfrak{J}}^N$. In particular, all of our reasoning besides tracking the support of both $\mathfrak{f}^{\mathfrak{l},1}$ and $\mathfrak{f}^{\mathfrak{l},2}$ functionals will be algebraic. Defining $\Phi_T$ as the entire RHS of the proposed identity and $\mathrm{d}_T$ as the $T$-time differential,

$$
\mathrm{d}_T \Phi_T = \mathrm{d}_T \tilde{\mathfrak{J}}^N_{S,T,x,y} + \mathrm{d}_T \int_S^T \sum_{w \in \mathbb{T}_N} \bar{\mathfrak{J}}^N_{R,T,x,w} \cdot N^2 \sum_{|\mathfrak{l}| \leqslant 10 N^{\beta_X} \mathfrak{m}} \left( \mathfrak{f}^{\mathfrak{l},1}_{R,w} \tilde{\mathfrak{J}}^N_{S,R,w+\mathfrak{l},y} \right.
$$
$$
\left. + \mathfrak{f}^{\mathfrak{l},2}_{R,w} \tilde{\mathfrak{J}}^N_{S,R,w\mp\mathfrak{l},y} \right) \mathrm{d}R. \tag{4.58}
$$

Denote the first term on the RHS of (4.58) by $\Phi_1$ and the second term by $\Phi_2$. Let us define $\Psi^{N,\mathfrak{l}}_{R,w} \overset{\bullet}{=} \mathfrak{f}^{\mathfrak{l},1}_{R,w} \tilde{\mathfrak{J}}^N_{S,R,w+\mathfrak{l},y} + \mathfrak{f}^{\mathfrak{l},2}_{R,w} \tilde{\mathfrak{J}}^N_{S,R,w\mp\mathfrak{l},y}$ as what is integrated in $\Phi_2$ on the RHS of (4.58). We start with computing $\Phi_2$ using the chain rule, where the first term on the RHS of (4.59) comes by applying $\mathrm{d}_T$ to $\bar{\mathfrak{J}}^N$ and the second term comes by applying $\mathrm{d}_T$ to the integral while recalling $\tilde{\mathfrak{J}}^N_{T,T,x,y} = \mathbf{1}_{x=y}$:

$$
\Phi_2 = \int_S^T \sum_{w \in \mathbb{T}_N} \mathrm{d}_T \bar{\mathfrak{J}}^N_{R,T,x,w} \cdot N^2 \sum_{|\mathfrak{l}| \leqslant 10 N^{\beta_X} \mathfrak{m}} \Psi^{N,\mathfrak{l}}_{R,w} \mathrm{d}R + N^2 \sum_{|\mathfrak{l}| \leqslant 10 N^{\beta_X} \mathfrak{m}} \Psi^{N,\mathfrak{l}}_{T,x} \mathrm{d}T. \tag{4.59}
$$

Let $\bar{\mathscr{T}}$ be the linear operator acting on functions $\varphi : \mathbb{R}_{\geqslant 0} \times \mathbb{T}_N \to \mathbb{R}$ so that $\mathrm{d}_T \bar{\mathfrak{J}}^N_{S,T,x,y} = \bar{\mathscr{T}} \bar{\mathfrak{J}}^N_{S,T,x,y} \mathrm{d}T$. By (4.58) and (4.59) and recalling the definition of $\Phi_T$ as the RHS of the proposed estimate, we get the following which we explain afterwards:

$$
\mathrm{d}_T \Phi_T = \mathrm{d}_T \tilde{\mathfrak{J}}^N_{S,T,x,y} + \left( \int_S^T \sum_{w \in \mathbb{T}_N} \bar{\mathscr{T}} \bar{\mathfrak{J}}^N_{R,T,x,w} \cdot N^2 \sum_{|\mathfrak{l}| \leqslant 10 N^{\beta_X} \mathfrak{m}} \Psi^{N,\mathfrak{l}}_{R,w} \mathrm{d}R \right) \mathrm{d}T
$$
$$
+ N^2 \sum_{|\mathfrak{l}| \leqslant 10 N^{\beta_X} \mathfrak{m}} \Psi^{N,\mathfrak{l}}_{T,x} \mathrm{d}T
$$
$$
= \mathrm{d}_T \tilde{\mathfrak{J}}^N_{S,T,x,y} + \bar{\mathscr{T}} \left( \int_S^T \sum_{w \in \mathbb{T}_N} \bar{\mathfrak{J}}^N_{R,T,x,w} \cdot N^2 \sum_{|\mathfrak{l}| \leqslant 10 N^{\beta_X} \mathfrak{m}} \Psi^{N,\mathfrak{l}}_{R,w} \mathrm{d}R \right) \mathrm{d}T
$$
$$
+ N^2 \sum_{|\mathfrak{l}| \leqslant 10 N^{\beta_X} \mathfrak{m}} \Psi^{N,\mathfrak{l}}_{T,x} \mathrm{d}T
$$
$$
= \bar{\mathscr{T}} \left( \mathfrak{J}^N_{S,T,x,y} + \int_S^T \sum_{w \in \mathbb{T}_N} \bar{\mathfrak{J}}^N_{R,T,x,w} \cdot N^2 \sum_{|\mathfrak{l}| \leqslant 10 N^{\beta_X} \mathfrak{m}} \Psi^{N,\mathfrak{l}}_{R,w} \mathrm{d}R \right) \mathrm{d}T
$$
$$
+ \left( \mathrm{d}_T \tilde{\mathfrak{J}}^N_{S,T,x,y} - \bar{\mathscr{T}} \tilde{\mathfrak{J}}^N_{S,T,x,y} \mathrm{d}T \right) + N^2 \sum_{|\mathfrak{l}| \leqslant 10 N^{\beta_X} \mathfrak{m}} \Psi^{N,\mathfrak{l}}_{T,x} \mathrm{d}T
$$
$$
= \bar{\mathscr{T}} \Phi_T \mathrm{d}T + \left( \mathrm{d}_T \tilde{\mathfrak{J}}^N_{S,T,x,y} - \bar{\mathscr{T}} \tilde{\mathfrak{J}}^N_{S,T,x,y} \mathrm{d}T \right) + N^2 \sum_{|\mathfrak{l}| \leqslant 10 N^{\beta_X} \mathfrak{m}} \Psi^{N,\mathfrak{l}}_{T,x} \mathrm{d}T. \tag{4.60}
$$

To be precise, the first line above follows from combining (4.58) and (4.59). The second identity above follows by pulling out the time-independent $\bar{\mathscr{T}}$-operator, acting on $\tilde{\mathfrak{J}}^N$ through the $x$-variable, outside the integral. The third line replaces $\mathrm{d}_T\tilde{\mathfrak{J}}^N$ with $\bar{\mathscr{T}}\tilde{\mathfrak{J}}^N$ and isolates the resulting error in such replacement. The last line follows by definition of $\Phi_T$.

Recalling the definition of $\Psi^{N,\mathfrak{l}}$, it is left to show that the second and third terms within (4.60) cancel for appropriate choices of functionals $\mathfrak{f}^\mathfrak{l}$. To this end, first observe that the second term in (4.60) vanishes as long as the usual distance between $x \in \mathbb{T}_N$ and the boundary of $\mathbb{T}_N$ is at least $10N^{\beta_X}\mathfrak{m}$, thus for these points we may choose $\mathfrak{f}^{\mathfrak{l},i} = 0$. On the other hand, for points within $10N^{\beta_X}\mathfrak{m}$ of the boundary, the second term involves contributions from $\tilde{\mathfrak{J}}^N$ evaluated at sites in a neighborhood in $\mathbb{T}_N$ of radius $10N^{\beta_X}\mathfrak{m}$ under the geodesic distance on $\mathbb{T}_N$ and sites in a neighborhood in $\mathbb{Z}$ of the same radius of $10N^{\beta_X}\mathfrak{m}$ but under the usual distance on $\mathbb{Z}$, all afterwards scaled by $N^2$ and additional uniformly bounded functionals. This paragraph may be checked just by unfolding definition of the local terms/operators $\mathrm{d}_T\tilde{\mathfrak{J}}^N$ and $\bar{\mathscr{T}}$. The second set of sites where we pick $\mathfrak{f}^\mathfrak{l} \neq 0$ are necessarily bounded below by $N^{5/4+2\varepsilon_{\mathrm{cpt}}/3}$ in absolute value by required proximity to the boundary of $\mathbb{T}_N$. This completes the proof. $\qquad\square$

## 5. Dynamical One-Block Analysis

We study the first two/non-gradient terms in $\bar{\Phi}^{N,2}$ of Definition 4.3 by "local equilibrium" in Lemma 3.8, invariant measure bounds in Lemmas 3.10, 3.11, 3.12, and Corollary 3.14, then combined with analytic inputs of heat kernel estimates for $\tilde{\mathfrak{P}}^N$.

- We first introduce estimates of main interest in this section. The proofs for all results will be deferred until after statements of all main results; we invite the reader, at least in a first reading, to skip these technical proofs and go to Sect. 6. We clarify that there are five propositions in this current section, the last four of which carry corollaries that will be crucial to the proof of Theorem 1.7. We give the proof of every corollary after its statement as each such proof will be almost trivial.
- The proof of each main result in this section will each require a few preliminary ingredients. We present these preliminaries for each result and combine them to give the proof of the corresponding result. We defer the proof of *all* preliminary lemmas to the end of this section to avoid obscuring the idea behind proofs of main results with a host of technical manipulations.
- The results are given in the following order. First, we will introduce a mechanism to replace the order $N^{1/2}$ spatial average of pseudo-gradients by a CLT-type cutoff that resembles Pseudo-Proposition 2.12. We then give an estimate that we eventually apply to estimate errors in the multiscale time-replacement strategy discussed in the would-be-proof for Pseudo-Proposition 2.13. Third, we present estimates to analyze the error terms, which we recall are time-averages of pseudo-gradient content with two-sided a priori bounds, from the multiscale "gluing" that we discussed in the would-be-proof of Pseudo-Proposition 2.14. We will write versions of the estimates for the first order $N^{1/2}$ term in $\bar{\Phi}^{N,2}$ and then for the order $N^{\beta_X}$-terms in $\bar{\Phi}^{N,2}$. We then organize resulting corollaries that turn the aforementioned estimates into multiscale results. Lastly, we give proofs.

We inherit notation of Proposition 2.7. We invite the reader to go back to Proposition 2.7 for notation throughout this section.

*5.1. Static analysis.* Let us first introduce the following cutoff basically from the statement of Pseudo-Proposition 2.12.

**Definition 5.1.** Recall $\beta_X = \frac{1}{3} + \varepsilon_{X,1}$ and $\mathscr{A}^{\mathbf{X},-}$ in Proposition 2.7. Define $\mathbf{1}[\mathscr{E}^{\mathbf{X},-}_{S,y}] \doteq \mathbf{1}[|\mathscr{A}^{\mathbf{X},-}_{N^{\beta_X}}(\mathfrak{g}_{S,y})| \lesssim N^{-\frac{1}{2}\beta_X + \frac{1}{2}\varepsilon_{X,1}}]$ and

$$\mathscr{C}^{\mathbf{X},-}_{N^{\beta_X}}(\mathfrak{g}_{S,y}) \doteq \mathscr{A}^{\mathbf{X},-}_{N^{\beta_X}}(\mathfrak{g}_{S,y}) \cdot \mathbf{1}[\mathscr{E}^{\mathbf{X},-}_{S,y}]. \tag{5.1}$$

The following result is a quantitative and $\mathbb{T}_N$-adapted version of Pseudo-Proposition 2.12. Its proof, as we will see towards the second half of this section, depends on the observation that the introduction of the CLT-type cutoff at canonical ensembles does something only with *exponentially low* probability. We then apply the local equilibrium reduction in Lemma 3.8.

**Proposition 5.2.** *With* $\mathscr{A}^{\mathbf{X},-}_{N^{\beta_X}}(\mathfrak{g})$ *in Proposition 2.7 and* $\mathscr{C}^{\mathbf{X},-}_{N^{\beta_X}}(\mathfrak{g})$ *in Definition 5.1, there exists* $\beta_{\mathrm{univ}} \in \mathbb{R}_{>0}$ *universal so that*

$$\mathbf{E}\|\bar{\mathfrak{P}}^N_{T,x}(N^{\frac{1}{2}}|\mathscr{A}^{\mathbf{X},-}_{N^{\beta_X}}(\mathfrak{g}_{S,y}) - \mathscr{C}^{\mathbf{X},-}_{N^{\beta_X}}(\mathfrak{g}_{S,y})|)\|_{1;\mathbb{T}_N} \lesssim N^{-\beta_{\mathrm{univ}}}. \tag{5.2}$$

The utility of Proposition 5.2 is in a priori bounds on time-averages of $\mathscr{C}^{\mathbf{X},-}$. This is our only advantage of $\mathscr{C}^{\mathbf{X},-}$ versus $\mathscr{A}^{\mathbf{X},-}$.

*5.2. Dynamic analysis IA.* We move onto estimates behind implementation of a quantitative $\mathbb{T}_N$-adapted version of Pseudo-Proposition 2.13. We recall that this pseudo-proposition justifies the replacement of the spatial-average-with-cutoff $\mathscr{C}^{\mathbf{X},-}$ from Definition 5.1 with its time-average on some mesoscopic time-scale. We additionally recall that the strategy behind the would-be-proof of Pseudo-Proposition 2.13 amounts to a multiscale scheme given by a replacement-by-time-average on progressively larger time-scales until we reach the "maximal" mesoscopic time-scale. We start with notation for time-averages.

**Definition 5.3.** Take $\mathfrak{t}_{\mathrm{av}} \in \mathbb{R}_{\geqslant 0}$ and define the following time-average $\mathscr{A}^{\mathbf{T},+}$ acting on the spatial-average and its cutoff:

$$\mathscr{A}^{\mathbf{T},+}_{\mathfrak{t}_{\mathrm{av}}} \mathscr{A}^{\mathbf{X},-}_{N^{\beta_X}}(\mathfrak{g}_{S,y}) \doteq \mathbf{1}_{\mathfrak{t}_{\mathrm{av}} \neq 0}\mathfrak{t}_{\mathrm{av}}^{-1}\int_0^{\mathfrak{t}_{\mathrm{av}}} \mathscr{A}^{\mathbf{X},-}_{N^{\beta_X}}(\mathfrak{g}_{S+\mathfrak{r},y})\mathrm{d}\mathfrak{r} + \mathbf{1}_{\mathfrak{t}_{\mathrm{av}}=0}\mathscr{A}^{\mathbf{X},-}_{N^{\beta_X}}(\mathfrak{g}_{S,y}) \tag{5.3a}$$

$$\mathscr{A}^{\mathbf{T},+}_{\mathfrak{t}_{\mathrm{av}}} \mathscr{C}^{\mathbf{X},-}_{N^{\beta_X}}(\mathfrak{g}_{S,y}) \doteq \mathbf{1}_{\mathfrak{t}_{\mathrm{av}} \neq 0}\mathfrak{t}_{\mathrm{av}}^{-1}\int_0^{\mathfrak{t}_{\mathrm{av}}} \mathscr{C}^{\mathbf{X},-}_{N^{\beta_X}}(\mathfrak{g}_{S+\mathfrak{r},y})\mathrm{d}\mathfrak{r} + \mathbf{1}_{\mathfrak{t}_{\mathrm{av}}=0}\mathscr{C}^{\mathbf{X},-}_{N^{\beta_X}}(\mathfrak{g}_{S,y}). \tag{5.3b}$$

Let us clarify that the first time-average $\mathscr{A}^{\mathbf{T},+}\mathscr{A}^{\mathbf{X},-}$ will not be used until later in a future subsection concerning the proof for a result dedicated towards the would-be-proof of Pseudo-Proposition 2.14. We think of $\mathscr{A}^{\mathbf{T},+}$ as a time-averaging operator.

**Proposition 5.4.** *Take any deterministic* $\mathfrak{t}_{\mathrm{av}} \in \mathbb{R}_{\geqslant 0}$ *satisfying* $N^{-2} \lesssim \mathfrak{t}_{\mathrm{av}} \lesssim N^{-1}$. *There exists* $\beta_{\mathrm{univ}} \in \mathbb{R}_{>0}$ *universal so that*

$$\mathfrak{t}_{\mathrm{av}}^{1/4}\mathbf{E}\|\bar{\mathfrak{P}}^N_{T,x}(N^{\frac{1}{2}}|\mathscr{A}^{\mathbf{T},+}_{\mathfrak{t}_{\mathrm{av}}}\mathscr{C}^{\mathbf{X},-}_{N^{\beta_X}}(\mathfrak{g}_{S,y})|)\|_{1;\mathbb{T}_N} \lesssim N^{-\beta_{\mathrm{univ}}}. \tag{5.4}$$

The factor $\mathfrak{t}_{\mathrm{av}}^{1/4}$ will come from time-regularity of the $\bar{\mathfrak{z}}^N$-process, so the nature of Proposition 5.4 matches that of (2.22).

*5.3. Dynamic analysis IIA.* We now focus on estimates for the multiscale idea in the would-be-proof for Pseudo-Proposition 2.14. As discussed earlier in the would-be-proof of Pseudo-Proposition 2.14, this means we will study time-averages of cutoffs $\mathscr{C}^{\mathbf{X},-}$ of spatial averages $\mathscr{A}^{\mathbf{X},-}$ of pseudo-gradients by analyzing two-sided cutoffs for these time-averages. First, some notation to make presenting these two-sided cutoff estimates more convenient.

**Definition 5.5.** Recall the constructions of time-averages from Definition 5.3. Define the following pair of events controlling maximal processes and time-averages. We clarify that in these following events, the process of interest is the supremum of the *integral* of $\mathscr{C}^{\mathbf{X},-}$-terms over all time-scales $0 \leqslant \mathfrak{t} \leqslant \mathfrak{t}_{av}$ that are then weighted by $\mathfrak{t}_{av}^{-1}$. These are only time-averages if $\mathfrak{t} = \mathfrak{t}_{av}$:

$$\mathbf{1}[\mathscr{G}_{\lesssim;S,y}^{\beta_-}] \stackrel{\bullet}{=} \mathbf{1}\left(\sup_{0 \leqslant \mathfrak{t} \leqslant \mathfrak{t}_{av}} \mathfrak{t} \cdot \mathfrak{t}_{av}^{-1} |\mathscr{A}_{\mathfrak{t}}^{\mathbf{T},+} \mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{S,y})| \lesssim N^{-\beta_-}\right) \tag{5.5a}$$

$$\mathbf{1}[\mathscr{G}_{\geqslant;S,y}^{\beta_+}] \stackrel{\bullet}{=} \mathbf{1}\left(\sup_{0 \leqslant \mathfrak{t} \leqslant \mathfrak{t}_{av}} \mathfrak{t} \cdot \mathfrak{t}_{av}^{-1} |\mathscr{A}_{\mathfrak{t}}^{\mathbf{T},+} \mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{S,y})| \geqslant N^{-\beta_+}\right). \tag{5.5b}$$

We will eventually take $\beta_-, \beta_+ \in \mathbb{R}_{\geqslant 0}$ deterministic such that $\beta_- \leqslant \beta_+$. The implied constant in the first event is universal.

Take any deterministic $\mathfrak{t}_{N,\varepsilon} \in \mathbb{R}_{\geqslant 0}$ and define the following "two-sided cutoffs" for the time-averages in Definition 5.3:

$$\mathscr{C}_{\mathfrak{t}_{av};\beta_-,\beta_+}^{\mathbf{T},+,\mathfrak{t}_{N,\varepsilon},1} \mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{S,y}) \stackrel{\bullet}{=} \mathscr{A}_{\mathfrak{t}_{av}}^{\mathbf{T},+} \mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{S+\mathfrak{t}_{N,\varepsilon},y}) \cdot \mathbf{1}[\mathscr{G}_{\lesssim;S+\mathfrak{t}_{N,\varepsilon},y}^{\beta_-}]\mathbf{1}[\mathscr{G}_{\geqslant;S,y}^{\beta_+}] \tag{5.6a}$$

$$\mathscr{C}_{\mathfrak{t}_{av};\beta_-,\beta_+}^{\mathbf{T},+,\mathfrak{t}_{N,\varepsilon},2} \mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{S,y}) \stackrel{\bullet}{=} \mathscr{A}_{\mathfrak{t}_{av}}^{\mathbf{T},+} \mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{S,y}) \cdot \mathbf{1}[\mathscr{G}_{\lesssim;S,y}^{\beta_-}]\mathbf{1}[\mathscr{G}_{\geqslant;S+\mathfrak{t}_{N,\varepsilon},y}^{\beta_+}]. \tag{5.6b}$$

In words, the first $\mathscr{C}^{\mathbf{T},+,1}$-operator adjusts the $\mathscr{A}^{\mathbf{T},+}$-operator with forward time-shift by $\mathfrak{t}_{N,\varepsilon} \in \mathbb{R}_{\geqslant 0}$. It also introduces an upper bound cutoff of order $N^{-\beta_-}$ for this shifted time-average, and then it introduces the lower bound cutoff of $N^{-\beta_+}$ for the *unshifted* time-average. Meanwhile the second $\mathscr{C}^{\mathbf{T},+,2}$-operator adjusts the $\mathscr{A}^{\mathbf{T},+}$-operator by introducing an upper bound cutoff of order $N^{-\beta_-}$ and then a lower bound cutoff of $N^{-\beta_+}$ for a $\mathfrak{t}_{N,\varepsilon}$-*shifted* time-average. Note that by taking $\mathfrak{t} = \mathfrak{t}_{av}$ in the suprema defining $\mathscr{G}^{\beta_-}$-events in $\mathscr{C}^{\mathbf{T},+,1}$ and $\mathscr{C}^{\mathbf{T},+,2}$ above, we get the following elementary deterministic bound with a universal implied constant:

$$|\mathscr{C}_{\mathfrak{t}_{av};\beta_-,\beta_+}^{\mathbf{T},+,\mathfrak{t}_{N,\varepsilon},1} \mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{S,y})| + |\mathscr{C}_{\mathfrak{t}_{av};\beta_-,\beta_+}^{\mathbf{T},+,\mathfrak{t}_{N,\varepsilon},2} \mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{S,y})| \lesssim N^{-\beta_-}. \tag{5.7}$$

We emphasize that these $\mathscr{C}^{\mathbf{T},+,\mathfrak{i}}$-cutoffs will be error terms that we get as the result of starting with an a priori cutoff for time-averages of $\mathscr{C}^{\mathbf{X},-}$-terms, improving cutoffs slightly, gluing time-averages with slightly improved estimates to time-averages on slightly bigger time-scales with the same slightly improved estimates, and iterating these steps starting with, again, improving slightly the last a priori estimates. This is the multiscale procedure in the would-be-proof of Pseudo-Proposition 2.14. The only difference with these $\mathscr{C}^{\mathbf{T},+,\mathfrak{i}}$-cutoffs and the error terms presented in the would-be-proof for Pseudo-Proposition 2.14 is that the previous $\mathscr{C}^{\mathbf{T},+,\mathfrak{i}}$-cutoffs are defined with cutoffs for *maximal processes* associated to the time-averages in Definition 5.3. Also the additional time-shift $\mathfrak{t}_{N,\varepsilon} \in \mathbb{R}_{\geqslant 0}$ is absent from the would-be-proof of Pseudo-Proposition 2.14. This point/difference is mostly technical, especially because the time-shift $\mathfrak{t}_{N,\varepsilon} \in \mathbb{R}_{\geqslant 0}$ will basically be of the same order as the time-average scale $\mathfrak{t}_{av} \in \mathbb{R}_{\geqslant 0}$.

The main result for this current subsection is the following control for these $\mathscr{C}^{\mathbf{T},+,\mathsf{i}}$-cutoffs against the heat operator $\bar{\mathfrak{P}}^N$ with an appropriate $\mathbb{T}_N$-adapted norm. We will emphasize here only the relationship between the time-scales for time-averages and the exponents for the $\mathscr{C}^{\mathbf{T},+,\mathsf{i}}$-cutoffs in the following result: given better a priori upper bounds, and thus for $\beta_- \in \mathbb{R}_{\geqslant 0}$ bigger, we can reduce to local equilibrium with Lemma 3.8 for a longer time-scale $\mathsf{t}_{\mathrm{av}} \in \mathbb{R}_{\geqslant 0}$ for the time-average. With a longer time-scale $\mathsf{t}_{\mathrm{av}} \in \mathbb{R}_{\geqslant 0}$ for the time-average, we can upgrade the exponent $\beta_- \in \mathbb{R}_{\geqslant 0}$ to a slightly bigger $\beta_+ \in \mathbb{R}_{\geqslant 0}$ and then control the error since *at invariant measures/statistical equilibrium* the lower-bound cutoff of $N^{-\beta_+}$ is negligible with sufficiently high probability if we pick a sufficiently large time-scale $\mathsf{t}_{\mathrm{av}} \in \mathbb{R}_{\geqslant 0}$ for the time-average.

**Proposition 5.6.** *Consider deterministic data* $\beta_{\pm} \in \mathbb{R}_{\geqslant 0}$ *and* $\mathsf{t}_{\mathrm{av}} \in \mathbb{R}_{\geqslant 0}$ *and* $\mathsf{t}_{N,\varepsilon} \in \mathbb{R}_{\geqslant 0}$ *satisfying the following constraints.*

- *We have* $\beta_- \geqslant \frac{1}{2}\beta_X - \varepsilon_{X,2}$ *and* $\beta_+ = \beta_- + \varepsilon_{X,3}$ *for* $\varepsilon_{X,2}, \varepsilon_{X,3} \in \mathbb{R}_{>0}$ *arbitrarily/sufficiently small but universal. We have* $\beta_+ \leqslant \frac{2}{3}$.
- *Define* $\mathsf{t}_{\mathrm{av}} \overset{\bullet}{=} N^{-1-\beta_X - \frac{1}{2}\beta_- + \beta_+ + \varepsilon_{X,4}}$ *for* $\varepsilon_{X,4} \in \mathbb{R}_{>0}$ *arbitrarily small but uniformly bounded below.*
- *Consider* $\mathsf{t}_{N,\varepsilon} \in \mathbb{R}_{\geqslant 0}$ *such that* $0 \leqslant \mathsf{t}_{N,\varepsilon} \lesssim N^\varepsilon \mathsf{t}_{\mathrm{av}}$ *with* $\varepsilon \in \mathbb{R}_{>0}$ *arbitrarily small but universal.*

*Define* $\|\varphi\|_{\bullet;\mathbb{X}} = \|\varphi_{\mathsf{t},x}\|_{\bullet;\mathbb{X}} = \sup_{0 \leqslant \mathsf{t} \leqslant \bullet} \sup_{x \in \mathbb{X}} |\varphi_{\mathsf{t},x}|$. *There exists* $\beta_{\mathrm{univ}} \in \mathbb{R}_{>0}$ *universal so for any deterministic* $0 \leqslant \mathsf{t}_{\mathsf{s}}, \mathsf{t}_{\mathsf{f}} \leqslant 2$, *we have the following estimate, in which* $\varepsilon_X = \max_i \varepsilon_{X,i} \vee \varepsilon$, *whose statement we clarify afterwards:*

$$\sup_{\mathsf{i}=1,2} \mathbf{E}\|\bar{\mathfrak{P}}_{T,x}^N (N^{\frac{1}{2}} |\mathscr{C}_{\mathsf{t}_{\mathrm{av}};\beta_-,\beta_+}^{\mathbf{T},+,\mathsf{t}_{N,\varepsilon},\mathsf{i}} \mathscr{C}_{N^{\beta_X}}^{\mathbb{X},-} (\mathfrak{g}_{S+\mathsf{t}_{\mathsf{s}}},y)|)\|_{\mathsf{t}_{\mathsf{f}};\mathbb{T}_N}$$
$$\lesssim N^{-\beta_{\mathrm{univ}}} + N^{-\frac{3}{4}\beta_- + 10\varepsilon_X} + N^{-\frac{1}{2}+10\varepsilon_X} + N^{-\beta_{\mathrm{univ}}} \mathsf{t}_{\mathsf{f}}. \tag{5.8}$$

*Let us emphasize the norm on the LHS is a supremum over the* $(T, x)$-*variables in the heat operator. The space-time variables* $(S, y)$ *are the integration variables in the heat operator and the constant* $\mathsf{t}_{\mathsf{s}} \in \mathbb{R}_{\geqslant 0}$ *is a deterministic time-shift for the particle system data. Lastly, the supremum over* $\mathsf{i} \in \{1, 2\}$ *is over the two choices of time-average cutoffs* $\mathscr{C}^{\mathbf{T},+,1}$ *and* $\mathscr{C}^{\mathbf{T},+,2}$ *from Definition 5.5.*

*Remark 5.7.* We could have given the previous result for $\mathsf{t}_{\mathsf{f}} = 1$ and its utility for the current paper would not have changed. We stated it for general $\mathsf{t}_{\mathsf{f}} \in \mathbb{R}_{\geqslant 0}$ as this will be important for the narrow-wedge initial measure to be addressed in a future article. Similarly, the time-shift $\mathsf{t}_{\mathsf{s}} \in \mathbb{R}_{\geqslant 0}$ will play almost no role in the proof of Proposition 5.6 and is there because of what the terms that we will eventually need to control with Proposition 5.6 actually are. See Sect. 7/the proof of Lemma 7.11 for details.

We eventually take $\varepsilon_{X,4}$ in the statement of Proposition 5.6 $N$-dependent, but bounded below uniformly, for presentational convenience in Sect. 7. We note this again in Corollary 5.12 which contains a version of Proposition 5.6 we will use later.

*5.4. Dynamic analysis IB.* This current subsection and the next are all basically versions of Proposition 5.4/Proposition 5.6 but for order $N^{\beta_X}$-terms in $\bar{\Phi}^{N,2}$. The clarifying remarks concerning Proposition 5.4 and Proposition 5.6 apply here as well. We start by basically replicating Proposition 5.4 but for order $N^{\beta_X}$-terms in $\bar{\Phi}^{N,2}$. In particular, we justify replacement of $\widetilde{\mathfrak{g}}^{\mathfrak{l}}$-terms of order $N^{\beta_X}$ with their time-averages on a mesoscopic time-scale, and the key estimate to this end is given in Proposition 5.9.

**Definition 5.8.** Consider any $t_{av} \in \mathbb{R}_{\geqslant 0}$ and define the following time-average-operator $\mathscr{A}^{\mathbf{T},+}$ acting on the functional $\widetilde{\mathfrak{g}}^{\mathfrak{l}}$ from Proposition 2.7 that admits a uniformly bounded pseudo-gradient factor:

$$\mathscr{A}^{\mathbf{T},+}_{t_{av}}(\widetilde{\mathfrak{g}}^{\mathfrak{l}}_{S,y}) \;\overset{\bullet}{=}\; \mathbf{1}_{t_{av}\neq 0} t_{av}^{-1} \int_0^{t_{av}} \widetilde{\mathfrak{g}}^{\mathfrak{l}}_{S+\mathfrak{r},y} d\mathfrak{r} \;+\; \mathbf{1}_{t_{av}=0}\widetilde{\mathfrak{g}}^{\mathfrak{l}}_{S,y}. \tag{5.9}$$

**Proposition 5.9.** *Consider any deterministic time-scale $t_{av} \in \mathbb{R}_{>0}$ satisfying $N^{-2} \lesssim t_{av} \lesssim N^{-1}$. There exists a universal constant $\beta_{univ} \in \mathbb{R}_{>0}$ so we have the following expectation estimate for a $\|\|\|_{1;\mathbb{T}_N}$-norm uniformly over terms-with-pseudo-gradient-factors:*

$$\sup_{\mathfrak{l}=1,\ldots,N^{\beta_X}} t_{av}^{1/4} \mathbf{E}\|\bar{\mathfrak{P}}^N_{T,x}(N^{\beta_X}|\mathscr{A}^{\mathbf{T},+}_{t_{av}}(\widetilde{\mathfrak{g}}^{\mathfrak{l}}_{S,y})|)\|_{1;\mathbb{T}_N} \;\lesssim\; N^{-\beta_{univ}}. \tag{5.10}$$

### 5.5. Dynamic analysis IIB.

We basically replicate Proposition 5.6 for order $N^{\beta_X}$-terms in $\bar{\Phi}^{N,2}$ in this subsection.

**Definition 5.10.** Recall the constructions of time-averages in Definition 5.8. We define the following pair of events controlling maximal processes and time-averages. Again, like Definition 5.5 the following events are concerned with suprema of *integrals* of $\widetilde{\mathfrak{g}}^{\mathfrak{l}}$-terms over time-scales $0 \leqslant t \leqslant t_{av}$ that are then weighted by $t_{av}^{-1}$. These are only time-averages if $t = t_{av}$:

$$\mathbf{1}[\mathscr{F}^{\beta_-}_{\lesssim;S,y}] \;\overset{\bullet}{=}\; \mathbf{1}\left(\sup_{0\leqslant t\leqslant t_{av}} t\cdot t_{av}^{-1}|\mathscr{A}^{\mathbf{T},+}_t(\widetilde{\mathfrak{g}}^{\mathfrak{l}}_{S,y})| \lesssim N^{-\beta_-}\right) \tag{5.11a}$$

$$\mathbf{1}[\mathscr{F}^{\beta_+}_{\gtrsim;S,y}] \;\overset{\bullet}{=}\; \mathbf{1}\left(\sup_{0\leqslant t\leqslant t_{av}} t\cdot t_{av}^{-1}|\mathscr{A}^{\mathbf{T},+}_t(\widetilde{\mathfrak{g}}^{\mathfrak{l}}_{S,y})| \geqslant N^{-\beta_+}\right). \tag{5.11b}$$

Take any deterministic $t_{N,\varepsilon} \in \mathbb{R}_{\geqslant 0}$ and define the following "two-sided cutoffs" for the time-averages in Definition 5.8:

$$\mathscr{C}^{\mathbf{T},+,t_{N,\varepsilon},1}_{t_{av};\beta_-,\beta_+}(\widetilde{\mathfrak{g}}^{\mathfrak{l}}_{S,y}) \;\overset{\bullet}{=}\; \mathscr{A}^{\mathbf{T},+}_{t_{av}}(\widetilde{\mathfrak{g}}^{\mathfrak{l}}_{S+t_{N,\varepsilon},y}) \cdot \mathbf{1}[\mathscr{F}^{\beta_-}_{\lesssim;S+t_{N,\varepsilon},y}]\mathbf{1}[\mathscr{F}^{\beta_+}_{\gtrsim;S,y}] \tag{5.12a}$$

$$\mathscr{C}^{\mathbf{T},+,t_{N,\varepsilon},2}_{t_{av};\beta_-,\beta_+}(\widetilde{\mathfrak{g}}^{\mathfrak{l}}_{S,y}) \;\overset{\bullet}{=}\; \mathscr{A}^{\mathbf{T},+}_{t_{av}}(\widetilde{\mathfrak{g}}^{\mathfrak{l}}_{S,y}) \cdot \mathbf{1}[\mathscr{F}^{\beta_-}_{\lesssim;S,y}]\mathbf{1}[\mathscr{F}^{\beta_+}_{\gtrsim;S+t_{N,\varepsilon},y}]. \tag{5.12b}$$

**Proposition 5.11.** *Consider deterministic data $\beta_\pm \in \mathbb{R}_{\geqslant 0}$ and $t_{av} \in \mathbb{R}_{\geqslant 0}$ and $t_{N,\varepsilon} \in \mathbb{R}_{\geqslant 0}$ satisfying the following constraints.*

- *We have $\beta_+ = \beta_- + \varepsilon_{X,2}$ for $\varepsilon_{X,2} \in \mathbb{R}_{>0}$ arbitrarily small but universal. We also have $\beta_+ \leqslant 5/12$.*
- *Define $t_{av} \overset{\bullet}{=} N^{-\frac{5}{4}-\frac{1}{2}\beta_-+\beta_++\varepsilon_{X,3}}$ for $\varepsilon_{X,3} \in \mathbb{R}_{>0}$ arbitrarily small but uniformly bounded below.*
- *Consider $t_{N,\varepsilon} \in \mathbb{R}_{\geqslant 0}$ such that $0 \leqslant t_{N,\varepsilon} \lesssim N^\varepsilon t_{av}$ with $\varepsilon \in \mathbb{R}_{>0}$ arbitrarily small but universal.*

*There exists $\beta_{univ} \in \mathbb{R}_{>0}$ universal so that for any deterministic $0 \leqslant t_{\mathfrak{s}}, t_{\mathfrak{f}} \leqslant 2$, we have the following in which $\varepsilon_X = \max_i \varepsilon_{X,i} \vee \varepsilon$:*

$$\sup_{i=1,2}\sup_{\mathfrak{l}=1,\ldots,N^{\beta_X}} \mathbf{E}\|\bar{\mathfrak{P}}^N_{T,x}(N^{\beta_X}|\mathscr{C}^{\mathbf{T},+,t_{N,\varepsilon},i}_{t_{av};\beta_-,\beta_+}(\widetilde{\mathfrak{g}}^{\mathfrak{l}}_{S+t_{\mathfrak{s}},y})|)\|_{t_{\mathfrak{f}};\mathbb{T}_N}$$

$$\lesssim\; N^{-\beta_{univ}} + N^{-\frac{1}{8}-\frac{3}{4}\beta_-+10\varepsilon_X} + N^{-\frac{1}{2}+10\varepsilon_X} + t_{\mathfrak{f}}N^{-\beta_{univ}}. \tag{5.13}$$

*5.6. Consequences.* We first present the corollary corresponding to Proposition 5.6 as it tells us the "maximal" time-scale for the time-average we want to replace the spatial-average-with-cutoff $\mathscr{C}^{X,-}$ by while being able to estimate each error term from the multiscale idea for Pseudo-Proposition 2.14. This will be the maximal time-scale we can study with local equilibrium.

**Corollary 5.12.** *Consider deterministic data $\{\beta_m\}_{m=0}^{\infty}$ and $\{t_{av,m}\}_{m=1}^{\infty}$ and $\{t_{N,\varepsilon,m}\}_{m=1}^{\infty}$ and $\mathfrak{m}_+ \in \mathbb{Z}_{\geqslant 0}$ defined as follows.*

- *Define $\beta_0 \stackrel{\bullet}{=} \frac{1}{2}\beta_X - \varepsilon_{X,2}$ and $\beta_m \stackrel{\bullet}{=} \beta_{m-1} + \varepsilon_{X,3}$ for $\varepsilon_{X,2}, \varepsilon_{X,3} \in \mathbb{R}_{>0}$ arbitrarily/sufficiently small but universal and $m \in \mathbb{Z}_{\geqslant 1}$.*
- *Define $t_{av,m} \stackrel{\bullet}{=} N^{-1-\beta_X-\frac{1}{2}\beta_{m-1}+\beta_m+\varepsilon_{X,4}}$ for $m \in \mathbb{Z}_{\geqslant 1}$ and $\varepsilon_{X,4} \in \mathbb{R}_{>0}$ arbitrarily small but uniformly bounded below and chosen so that $t_{av,m+1}$ is an integer multiple of $t_{av,m}$. The integer-multiple-constraint will only be useful for convenience in Sect. 7. Note this may force $\varepsilon_{X,4}$ to be $N$-dependent but this will not matter as long as $\varepsilon_{X,4}$ is arbitrarily small but uniformly bounded below.*
- *Take $t_{N,\varepsilon,m} \in \mathbb{R}_{\geqslant 0}$ to be any deterministic time satisfying $0 \leqslant t_{N,\varepsilon,m} \lesssim N^\varepsilon t_{av,m}$ with $\varepsilon \in \mathbb{R}_{>0}$ arbitrarily small but universal.*
- *Define $\mathfrak{m}_+ \in \mathbb{Z}_{\geqslant 0}$ as the smallest positive integer for which $\beta_{\mathfrak{m}_+} > \frac{1}{2} + 3\varepsilon_{X,3}$.*

*For some universal constant $\beta_{univ} \in \mathbb{R}_{>0}$ and for any deterministic $0 \leqslant t_{\mathfrak{s}} \leqslant 2$, we have the estimate*

$$\sup_{m=1,\ldots,\mathfrak{m}_+}\sup_{i=1,2}\mathbf{E}\|\bar{\mathfrak{P}}_{T,x}^N(N^{\frac{1}{2}}|\mathscr{C}_{t_{av,m};\beta_{m-1},\beta_m}^{\mathbf{T},+,t_{N,\varepsilon,m},i}\mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{S+t_{\mathfrak{s}}},y)|)\|_{2;\mathbb{T}_N}$$
$$quad \lesssim N^{-\beta_{univ}}. \tag{5.14}$$

*We have $\mathfrak{m}_+ \lesssim 1$ and $1 \leqslant t_{av,m}t_{av,m-1}^{-1} \leqslant N^{\varepsilon_{X,3}}$ for $m \in \mathbb{Z}_{\geqslant 1}$ and $t_{av,\mathfrak{m}_+} \lesssim N^{-13/12+\varepsilon_{X,5}}$ for $\varepsilon_{X,5} \in \mathbb{R}_{>0}$ arbitrarily small but universal.*

*Proof.* Each expectation in the proposed bound satisfies constraints of Proposition 5.6 with $t_f = 2$, so it suffices to show $\mathfrak{m}_+ \lesssim 1$ and verify the claims about $t_{av,m} \in \mathbb{R}_{\geqslant 0}$. As $\beta_m - \beta_{m-1} = \varepsilon_{X,3}$, we get the bound $\mathfrak{m}_+ \lesssim \varepsilon_{X,3}^{-1}$. This also gives $t_{av,m}t_{av,m-1}^{-1} \leqslant N^{\varepsilon_{X,3}}$ immediately in addition to monotonicity in $m \in \mathbb{Z}_{>0}$ by slightly further inspection of the definition of $t_{av,m} \in \mathbb{R}_{\geqslant 0}$. To check the $t_{av,\mathfrak{m}_+}$-bound, we note $\beta_{\mathfrak{m}_+} - \frac{1}{2}\beta_{\mathfrak{m}_+-1} = \frac{1}{2}\beta_{\mathfrak{m}_+} + \frac{1}{2}\varepsilon_{X,3} \leqslant \frac{1}{4} + 3\varepsilon_{X,3}$ and $\beta_X \geqslant \frac{1}{3}$, and thus $-\beta_X - \frac{1}{2}\beta_{\mathfrak{m}_+-1} + \beta_{\mathfrak{m}_+} \leqslant -\frac{1}{12} + 3\varepsilon_{X,3}$. $\square$

**Corollary 5.13.** *Consider the deterministic data $\{\widetilde{\beta}_m\}_{m=0}^{\infty}$ and $\{\widetilde{t}_{av,m}\}_{m=1}^{\infty}$ and $\{\widetilde{t}_{N,\varepsilon,m}\}_{m=1}^{\infty}$ and $\widetilde{\mathfrak{m}_+} \in \mathbb{Z}_{>0}$ defined as follows.*

- *Define $\widetilde{\beta_0} \stackrel{\bullet}{=} 0$ and define $\widetilde{\beta_m} \stackrel{\bullet}{=} \widetilde{\beta}_{m-1} + \varepsilon_{X,2}$ for $\varepsilon_{X,2} \in \mathbb{R}_{>0}$ arbitrarily small but universal and $m \in \mathbb{Z}_{\geqslant 1}$.*
- *Define $\widetilde{t}_{av,m} \stackrel{\bullet}{=} N^{-\frac{5}{4}-\frac{1}{2}\widetilde{\beta}_{m-1}+\widetilde{\beta}_m+\varepsilon_{X,3}}$ for $m \in \mathbb{Z}_{\geqslant 1}$ and $\varepsilon_{X,3} \in \mathbb{R}_{>0}$ arbitrarily small but uniformly bounded below and chosen so that $\widetilde{t}_{av,m+1}$ is an integer multiple of $\widetilde{t}_{av,m}$. The integer-multiple-constraint will only be useful for convenience in Sect. 7.*
- *Take $\widetilde{t}_{N,\varepsilon,m} \in \mathbb{R}_{\geqslant 0}$ to be any deterministic time satisfying $0 \leqslant \widetilde{t}_{N,\varepsilon,m} \lesssim N^\varepsilon \widetilde{t}_{av,m}$ with $\varepsilon \in \mathbb{R}_{>0}$ sufficiently small but universal.*
- *Define $\widetilde{\mathfrak{m}_+} \in \mathbb{Z}_{\geqslant 0}$ as the smallest positive integer for which $\widetilde{\beta}_{\widetilde{\mathfrak{m}_+}} > \beta_X + 3\varepsilon_{X,2}$.*

*For some universal constant $\beta_{\mathrm{univ}} \in \mathbb{R}_{>0}$ and for any deterministic $0 \leqslant \mathfrak{t}_{\mathfrak{s}} \leqslant 2$, we have the estimate*

$$\sup_{m=1,\ldots,\widetilde{\mathfrak{m}}_+}\sup_{\mathfrak{i}=1,2}\sup_{\mathfrak{l}=1,\ldots,N^{\beta_X}} \mathbf{E}\|\bar{\mathfrak{P}}_{T,x}^{N}(N^{\beta_X}|\mathscr{C}_{\widetilde{\mathfrak{t}}_{\mathrm{av},m};\beta_{m-1}^{\sim},\beta_m^{\sim}}^{\mathbf{T},+,\widetilde{\mathfrak{t}}_{N,\varepsilon,m},\mathfrak{i}}(\widetilde{\mathfrak{g}}_{S+\mathfrak{t}_{\mathfrak{s}},y}^{\mathfrak{l}})|)\|_{2;\mathbb{T}_N}$$
$$\lesssim N^{-\beta_{\mathrm{univ}}}. \tag{5.15}$$

*We have $\widetilde{\mathfrak{m}}_+ \lesssim 1$ and $1 \leqslant \widetilde{\mathfrak{t}}_{\mathrm{av},m}(\widetilde{\mathfrak{t}}_{\mathrm{av},m-1})^{-1} \leqslant N^{\varepsilon_{X,2}}$ and $\widetilde{\mathfrak{t}}_{\mathrm{av},\widetilde{\mathfrak{m}}_+} \lesssim N^{-13/12+\varepsilon_{X,4}}$ for $\varepsilon_{X,4} \in \mathbb{R}_{>0}$ arbitrarily small but universal.*

*Proof.* This follows by Proposition 5.11 how Corollary 5.12 follows by Proposition 5.6. $\qquad\square$

With Corollary 5.12, we may control time-averages of $\mathscr{C}^{\mathbf{X},-}$ with respect to the "maximal" time-scale $\mathfrak{t}_{\mathrm{av},\mathfrak{m}_+} \in \mathbb{R}_{\geqslant 0}$ defined in the statement of Corollary 5.12 by the multiscale procedure outlined in the would-be-proof of Pseudo-Proposition 2.14. Thus, we replace the original spatial-average-with-cutoff $\mathscr{C}^{\mathbf{X},-}$ with its time-average on this time-scale $\mathfrak{t}_{\mathrm{av},\mathfrak{m}_+} \in \mathbb{R}_{\geqslant 0}$ via step-by-step replacement as in the would-be-proof of Pseudo-Proposition 2.13. We then estimate the errors with the following corollary.

**Corollary 5.14.** *Consider the sequence $\{\mathfrak{t}_{\mathrm{av},\ell}\}_{\ell=0}^{\infty}$ and $\mathfrak{l}_+ \in \mathbb{Z}_{\geqslant 0}$ defined below, where $\varepsilon_{X,2} \in \mathbb{R}_{>0}$ is arbitrarily small but universal.*

- *Define $\mathfrak{t}_{\mathrm{av},\ell+1} = \lfloor N^{\varepsilon_{X,2}} \rfloor \mathfrak{t}_{\mathrm{av},\ell} \wedge \mathfrak{t}_{\mathrm{av},\mathfrak{m}_+}$ where $\mathfrak{t}_{\mathrm{av},\mathfrak{m}_+}$ is from Corollary 5.12. Define $\mathfrak{l}_+$ as the first index $\ell$ for which $\mathfrak{t}_{\mathrm{av},\ell} = \mathfrak{t}_{\mathrm{av},\mathfrak{m}_+}$.*
- *Choose $\mathfrak{t}_{\mathrm{av},0}$ such that $N^{-2} \lesssim \mathfrak{t}_{\mathrm{av},0} \lesssim N^{-2+\varepsilon_{X,2}}$ and $\mathfrak{t}_{\mathrm{av},\ell+1}\mathfrak{t}_{\mathrm{av},\ell}^{-1} \in \mathbb{Z}_{>0}$ for all $\ell \in \mathbb{Z}_{\geqslant 0}$. Our choice of $\mathfrak{t}_{\mathrm{av},0}$ depends on $\varepsilon_{X,2}$ and $\mathfrak{t}_{\mathrm{av},\mathfrak{m}_+}$.*

*There exists $\beta_{\mathrm{univ}} \in \mathbb{R}_{>0}$ universal so that if $\varepsilon_{X,2} \in \mathbb{R}_{>0}$ is sufficiently small, we have uniform boundedness $\mathfrak{l}_+ \lesssim_{\varepsilon_{X,2}} 1$ and*

$$\sup_{\ell=0,\ldots,\mathfrak{l}_+} \mathfrak{t}_{\mathrm{av},\ell+1}^{1/4} \mathbf{E}\|\bar{\mathfrak{P}}_{T,x}^{N}(N^{\frac{1}{2}}|\mathscr{A}_{\mathfrak{t}_{\mathrm{av},\ell}}^{\mathbf{T},+}\mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{S,y})|)\|_{1;\mathbb{T}_N} \lesssim N^{-\beta_{\mathrm{univ}}}. \tag{5.16}$$

*Proof.* If we pick $\varepsilon_{X,2} \in \mathbb{R}_{>0}$ sufficiently small, we may replace the prefactor $\mathfrak{t}_{\mathrm{av},\ell+1}^{1/4} \to \mathfrak{t}_{\mathrm{av},\ell}^{1/4}$ for the $\ell$-quantity in the supremum while changing $\beta_{\mathrm{univ}} \in \mathbb{R}_{>0}$ on the RHS of the proposed estimate by a factor of $1/2$. At this point, we observe that all $\mathfrak{t}_{\mathrm{av},\ell} \in \mathbb{R}_{>0}$ in question satisfy constraints assumed for $\mathfrak{t}_{\mathrm{av}} \in \mathbb{R}_{>0}$ in Proposition 5.4. Indeed, the times $\mathfrak{t}_{\mathrm{av},0}, \mathfrak{t}_{\mathrm{av},\mathfrak{l}_+} \in \mathbb{R}_{>0}$ with maximal and minimal indices do by definition and by Corollary 5.12, and the times $\mathfrak{t}_{\mathrm{av},\ell} \in \mathbb{R}_{>0}$ are monotone in the index. Thus, the estimate on the sup follows by Proposition 5.4. The bound $\mathfrak{l}_+ \lesssim_{\varepsilon_{X,2}} 1$ follows by $\lfloor N^{\varepsilon_{X,2}} \rfloor \gtrsim N^{\varepsilon_{X,2}/2}$ and elementary considerations. $\qquad\square$

*Remark 5.15.* Like Corollary 5.12, the integer-multiple-constraint for time-scales in Corollary 5.14 is for later convenience.

The next/last result follows by Proposition 5.9 how Corollary 5.14 follows by Proposition 5.4.

**Corollary 5.16.** *Consider the sequence $\{\widetilde{\mathfrak{t}}_{\mathrm{av},\ell}\}_{\ell=0}^{\infty}$ and $\widetilde{\mathfrak{l}}_+ \in \mathbb{Z}_{\geqslant 0}$ defined below, where $\varepsilon_{X,2} \in \mathbb{R}_{>0}$ is arbitrarily small but universal.*

- *Define $\widetilde{\mathfrak{t}}_{\mathrm{av},\ell+1} = \lfloor N^{\varepsilon_{X,2}} \rfloor \widetilde{\mathfrak{t}}_{\mathrm{av},\ell} \wedge \widetilde{\mathfrak{t}}_{\mathrm{av},\widetilde{\mathfrak{m}}_+}$ where $\widetilde{\mathfrak{t}}_{\mathrm{av},\widetilde{\mathfrak{m}}_+}$ is from Corollary 5.13. Define $\widetilde{\mathfrak{l}}_+$ as the first index $\ell$ for which $\widetilde{\mathfrak{t}}_{\mathrm{av},\ell} = \widetilde{\mathfrak{t}}_{\mathrm{av},\widetilde{\mathfrak{m}}_+}$.*

- *Choose* $\widetilde{\mathfrak{t}}_{\mathrm{av},0}$ *such that* $N^{-2} \lesssim \widetilde{\mathfrak{t}}_{\mathrm{av},0} \lesssim N^{-2+\varepsilon_{X,2}}$ *and* $\widetilde{\mathfrak{t}}_{\mathrm{av},\ell+1}(\widetilde{\mathfrak{t}}_{\mathrm{av},\ell})^{-1} \in \mathbb{Z}_{>0}$. *Our choice of* $\widetilde{\mathfrak{t}}_{\mathrm{av},0}$ *depends on* $\varepsilon_{X,2}$ *and* $\widetilde{\mathfrak{t}}_{\mathrm{av},\widetilde{\mathfrak{m}}_+}$.

*There exists* $\beta_{\mathrm{univ}} \in \mathbb{R}_{>0}$ *universal so that if* $\varepsilon_{X,2} \in \mathbb{R}_{>0}$ *is sufficiently small, we have uniform boundedness* $\widetilde{\mathfrak{l}}_+ \lesssim_{\varepsilon_{X,2}} 1$ *and*

$$\sup_{\ell=0,\ldots,\widetilde{\mathfrak{l}}_+} \sup_{\mathfrak{l}=1,\ldots,N^{\beta_X}} (\widetilde{\mathfrak{t}}_{\mathrm{av},\ell+1})^{1/4} \mathbf{E} \|\bar{\mathfrak{P}}_{T,x}^N (N^{\beta_X}|\mathscr{A}_{\widetilde{\mathfrak{t}}_{\mathrm{av},\ell}}^{\mathbf{T},+}(\widetilde{\mathfrak{g}}_{S,y}^{\mathfrak{l}})|)\|_{1;\mathbb{T}_N} \lesssim N^{-\beta_{\mathrm{univ}}}. \tag{5.17}$$

The rest of this section is dedicated to the proofs of Proposition 5.2, Proposition 5.4, Proposition 5.6, Proposition 5.9, and Proposition 5.11. We again invite the reader to go to Sect. 6 at least upon a first reading as the proofs of the aforementioned results are on the technical side. As for proofs themselves, we will replace mesoscopic space-time averages by "local" versions by coupling arguments, and then we analyze these "local" versions by a reduction to equilibrium estimates Lemma 3.10, Lemma 3.11, and Lemma 3.12. As for Proposition 5.2, we will do the same but instead of Lemma 3.10, Lemma 3.11, and Lemma 3.12 we will apply static LDP estimates in Lemma 3.13 and Corollary 3.14. Every result will also require some technical gymnastics we explain below; see Lemma 5.17, for example. We will *not* present their proofs in the written order. We start with proofs of Proposition 5.6 and Proposition 5.11 because these proofs contain almost all ideas necessary to get the rest of the propositions. We will then move to the proofs for both Proposition 5.4 and Proposition 5.9. We conclude with a proof of Proposition 5.2.

*5.7. Proof of Proposition 5.6.* We will adopt notation from the statement of Proposition 5.6 throughout this subsection. We delay the proofs of preliminary lemmas until the last subsection of this section to avoid obscuring the proof of Proposition 5.6.

We will depend on several preliminary ingredients starting with aforementioned technical gymnastics. Roughly speaking, it allows us to replace the heat operator within the LHS of the proposed estimate in Proposition 5.6 with averaging against the constant function on $[0, \mathfrak{t}_{\mathfrak{f}}] \times \mathbb{T}_N$ free of short-time singularities of the heat kernel. We do this by "direct" convolution bound.

**Lemma 5.17.** *Take any possibly random function* $\varphi_{S,y} : \mathbb{R}_{\geqslant 0} \times \mathbb{T}_N \to \mathbb{R}$ *and suppose* $\|\varphi\|_{\infty;\mathbb{T}_N} \lesssim N^{-\beta_-}$ *with a universal implied constant. Provided any* $\varepsilon \in \mathbb{R}_{>0}$ *and* $\mathfrak{t}_{\mathfrak{s}}, \mathfrak{t}_{\mathfrak{f}} \in \mathbb{R}_{\geqslant 0}$, *we have the following deterministic estimate with universal implied constant:*

$$\|\bar{\mathfrak{P}}_{T,x}^N (N^{\frac{1}{2}}|\varphi_{S+\mathfrak{t}_{\mathfrak{s}},y}|)\|_{\mathfrak{t}_{\mathfrak{f}};\mathbb{T}_N} \lesssim N^{1+\varepsilon_{\mathrm{cpt}}+\frac{1}{2}\varepsilon-\frac{1}{2}\beta_-} \int_0^{\mathfrak{t}_{\mathfrak{f}}} \widetilde{\sum}_{y\in\mathbb{T}_N} |\varphi_{S+\mathfrak{t}_{\mathfrak{s}},y}| \, \mathrm{d}S + N^{-\varepsilon}. \tag{5.18}$$

From this last gymnastics-estimate, we get the following. It roughly follows by taking expectation of the RHS in the bound of Lemma 5.17 with $\varphi$ the iterated cutoff $\mathscr{C}^{\mathbf{T},+}\mathscr{C}^{\mathbf{X},-}$ inside the heat operator on the LHS of the proposed estimate in Proposition 5.6. The iterated expectation on the RHS of the proposed estimate in this next result Lemma 5.18 follows by decomposing the expectation of a dynamic functional in terms of an expectation over the path-space measure and then an expectation over the initial configuration, the latter of which we then average over the integral-sum of the RHS of the estimate in Lemma 5.17.

**Lemma 5.18.** *Recall* Definition 3.7. *Provided any* $\varepsilon \in \mathbb{R}_{>0}$ *and* $\mathfrak{t}_{\mathfrak{f}} \in \mathbb{R}_{\geqslant 0}$ *and* $\mathfrak{i} \in \{1, 2\}$, *we have, with notation explained after,*

$$\mathbf{E}\|\bar{\mathfrak{P}}_{T,x}^N (N^{\frac{1}{2}} |\mathscr{C}_{\mathfrak{t}_{\mathrm{av}};\beta_-,\beta_+}^{\mathbf{T},+,\mathfrak{t}_{N,\varepsilon},\mathfrak{i}} \mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-} (\mathfrak{g}_{S+\mathfrak{t}_{\mathfrak{s}}}, y)|)\|_{\mathfrak{t}_{\mathfrak{f}};\mathbb{T}_N}$$
$$\lesssim N^{1+\varepsilon_{\mathrm{cpt}}+\frac{1}{2}\varepsilon-\frac{1}{2}\beta_-}\mathfrak{t}_{\mathfrak{f}}\mathbf{E}^{\mu_{0,\mathbb{Z}}}\bar{\mathfrak{f}}_{\mathfrak{t}_{\mathfrak{s}},\mathfrak{t}_{\mathfrak{f}},\mathbb{T}_N}\mathbf{E}_{\eta_0}^{\mathrm{path}} |\mathscr{C}_{\mathfrak{t}_{\mathrm{av}};\beta_-,\beta_+}^{\mathbf{T},+,\mathfrak{t}_{N,\varepsilon},\mathfrak{i}} \mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{0,0})| + N^{-\varepsilon}. \quad (5.19)$$

*Above the expectation* $\mathbf{E}_{\eta_0}^{\mathrm{path}}$ *is an expectation with respect to the path-space measure induced by the original* $\Omega_{\mathbb{Z}}$-*valued exclusion process with an initial configuration* $\eta_0$ *that is then being averaged via* $\mathbf{E}^{\mu_{0,\mathbb{Z}}}$ *against the space-time averaged density* $\bar{\mathfrak{f}}_{\mathfrak{t}_{\mathfrak{s}},\mathfrak{t}_{\mathfrak{f}},\mathbb{T}_N}$.

We observe the path-space expectation of the iterated cutoff $\mathscr{C}^{\mathbf{T},+}\mathscr{C}^{\mathbf{X},-}$ is a *global* statistic. It depends on the entire configuration $\eta_0 \in \Omega_{\mathbb{Z}}$. However, because we integrate the mesoscopic spatial-cutoff $\mathscr{C}^{\mathbf{X},-}$ on a mesoscopic time-scale $\mathfrak{t}_{N,\varepsilon} \in \mathbb{R}_{\geqslant 0}$, it "looks like" a mesoscopic statistic. The next ingredient localizes these space-time averages of interest in Lemma 5.18. To state it, we require some notation. First let us recall the $\mathfrak{I}$-local exclusion processes introduced in Definition 3.9.

**Definition 5.19.** For $\mathfrak{t}_{\mathrm{av}} \in \mathbb{R}_{\geqslant 0}$, set $\mathfrak{l}_{\mathfrak{t}_{\mathrm{av}}} \overset{\bullet}{=} (N(2\mathfrak{t}_{\mathrm{av}} + 2\mathfrak{t}_{N,\varepsilon})^{\frac{1}{2}} + N^{\frac{3}{2}}(2\mathfrak{t}_{\mathrm{av}} + 2\mathfrak{t}_{N,\varepsilon}) + N^{\beta_X}) \log^{100} N$ and set $\mathfrak{I}_{\mathfrak{t}_{\mathrm{av}}} \overset{\bullet}{=} [\![-\mathfrak{l}_{\mathfrak{t}_{\mathrm{av}}}, \mathfrak{l}_{\mathfrak{t}_{\mathrm{av}}}]\!] \subseteq \mathbb{Z}$.

*Remark 5.20.* Consider the particle random walk for the $\mathfrak{S}^{N,!!}$-process on $\Omega_{\mathbb{Z}}$, and suppose it starts in the "middle" $N^{\beta_X}$ of $\mathfrak{I}_{\mathfrak{t}_{\mathrm{av}}}$. The probability this random walk exits $\mathfrak{I}_{\mathfrak{t}_{\mathrm{av}}}$ before time $2\mathfrak{t}_{\mathrm{av}} + 2\mathfrak{t}_{N,\varepsilon}$ is order $\mathbf{Exp}(-\log^{100} N) \lesssim_C N^{-C}$ for any $C \in \mathbb{R}_{>0}$. To prove this, we may first remove the drift because $|\mathfrak{I}_{\mathfrak{t}_{\mathrm{av}}}|$ is much larger than the drift on time-scale $2\mathfrak{t}_{\mathrm{av}} + 2\mathfrak{t}_{N,\varepsilon}$. For the symmetric part of the particle random walk, we use its sub-Gaussian property and a Doob maximal inequality to control its maximal process.

The following construction is a localization of space-time averages evaluated on $\mathfrak{I}_{\mathfrak{t}_{\mathrm{av}}}$-exclusion processes. Everything below is basically Definition 5.1, Definition 5.3, and Definition 5.5 but for $\mathfrak{I}_{\mathfrak{t}_{\mathrm{av}}}$-local processes; it is blind to anything outside $\mathfrak{I}_{\mathfrak{t}_{\mathrm{av}}} \subseteq \mathbb{Z}$.

**Definition 5.21.** Consider any $\mathfrak{t}_{\mathrm{av}} \in \mathbb{R}_{\geqslant 0}$. We define $\eta_{\mathfrak{t}_{\mathrm{av}};\bullet}$ as the $\Omega_{\mathfrak{I}_{\mathfrak{t}_{\mathrm{av}}}}$-valued $\mathfrak{I}_{\mathfrak{t}_{\mathrm{av}}}$-local exclusion process defined in Definition 3.9 with $\mathfrak{I}_{\mathfrak{t}_{\mathrm{av}}} \subseteq \mathbb{Z}$ defined earlier and the initial configuration $\eta_{\mathfrak{t}_{\mathrm{av}};0} = \Pi_{\mathfrak{I}_{\mathfrak{t}_{\mathrm{av}}}} \eta_0 \in \Omega_{\mathfrak{I}_{\mathfrak{t}_{\mathrm{av}}}}$, namely the projection on $\mathfrak{I}_{\mathfrak{t}_{\mathrm{av}}} \subseteq \mathbb{Z}$ of the initial configuration of the *original* $\Omega_{\mathbb{Z}}$-exclusion process. We now define local versions of Definitions 5.1, 5.3, and 5.5.

- Define $\mathfrak{g}_{\mathfrak{t}_{\mathrm{av}};S,0} \overset{\bullet}{=} \mathfrak{g}_{0,0}(\eta_{\mathfrak{t}_{\mathrm{av}};S})$ as the pseudo-gradient in Proposition 2.7 but evaluated at the $\mathfrak{I}_{\mathfrak{t}_{\mathrm{av}}}$-local process $\eta_{\mathfrak{t}_{\mathrm{av}};\bullet}$ at $S \in \mathbb{R}_{\geqslant 0}$. Observe this is well-defined because $\mathfrak{I}_{\mathfrak{t}_{\mathrm{av}}}$ contains the support of $\mathfrak{g}_{0,0}$ if $N \in \mathbb{Z}_{>0}$ is sufficiently large but universal.
- Recall $\beta_X = \frac{1}{3} + \varepsilon_{X,1}$ in Proposition 2.7. We define the event $\mathbf{1}[\mathscr{E}_{\mathfrak{t}_{\mathrm{av}};S}^{\mathbf{X},-}] = \mathbf{1}[|\mathscr{A}_{N^{\beta_X}}^{\mathbf{X},-} (\mathfrak{g}_{\mathfrak{t}_{\mathrm{av}};S,0})| \lesssim N^{-\frac{1}{2}\beta_X+\frac{1}{2}\varepsilon_{X,1}}]$ where

$$\mathscr{A}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{\mathfrak{t}_{\mathrm{av}};S,0}) = \widetilde{\sum}_{\mathfrak{l}=1,\ldots,N^{\beta_X}} \tau_{-7\mathfrak{l}\mathfrak{m}}\mathfrak{g}_{\mathfrak{t}_{\mathrm{av}};S,0}. \quad (5.20)$$

We clarify that $\tau_{-7\mathfrak{l}\mathfrak{m}}\mathfrak{g}_{\mathfrak{t}_{\mathrm{av}};S,0}$ corresponds to shifting the $\eta_{\mathfrak{t}_{\mathrm{av}};S}$-configuration by $-7\mathfrak{l}\mathfrak{m}$, which yields a particle configuration in $\Omega_{\mathfrak{I}_{\mathfrak{t}_{\mathrm{av}}}-7\mathfrak{l}\mathfrak{m}}$, and then evaluating $\mathfrak{g}_{0,0}$ at the particle configuration obtained by projecting the shifted configuration in $\Omega_{\mathfrak{I}_{\mathfrak{t}_{\mathrm{av}}}-7\mathfrak{l}\mathfrak{m}}$ onto the support of $\mathfrak{g}_{0,0}$ which is a *fixed* subset in $\mathbb{Z}$. Indeed, observe the support of $\mathfrak{g}_{0,0}$

is contained properly in $\mathfrak{I}_{t_{av}} - 7\mathfrak{l}\mathfrak{m} \subseteq \mathbb{Z}$ for all $\mathfrak{l} \in [\![0, N^{\beta_X}]\!]$ for $N \in \mathbb{Z}_{>0}$ sufficiently large depending only on $\mathfrak{m} \in \mathbb{Z}_{>0}$ because of definition of $\mathfrak{I}_{t_{av}} \subseteq \mathbb{Z}$. We emphasize the supports of the summands defining the spatial-average in the last display are disjoint subsets of $\mathfrak{I}_{t_{av}} \subseteq \mathbb{Z}$ as was the case in Proposition 2.7. We additionally define the following cutoff of the previous spatial-average via CLT considerations:

$$\mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{t_{av};S,0}) \overset{\bullet}{=} \mathscr{A}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{t_{av};S,0}) \cdot \mathbf{1}[\mathscr{E}_{t_{av};S}^{\mathbf{X},-}]. \tag{5.21}$$

- We first recall the time-scale $t_{av} \in \mathbb{R}_{\geqslant 0}$. Define the following time-averages acting on both spatial-averages and their cutoffs of the pseudo-gradient functional $\mathfrak{g}_{t_{av}}$; from the first bullet point that is localized to the $\mathfrak{I}_{t_{av}}$-local process:

$$\mathscr{A}_{t_{av}}^{\mathbf{T},+}\mathscr{A}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{t_{av};S,0}) \overset{\bullet}{=} \mathbf{1}_{t_{av}\neq 0}t_{av}^{-1}\int_0^{t_{av}} \mathscr{A}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{t_{av};S+\mathfrak{r},0})d\mathfrak{r} + \mathbf{1}_{t_{av}=0}\mathscr{A}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{t_{av};S,0}) \tag{5.22a}$$

$$\mathscr{A}_{t_{av}}^{\mathbf{T},+}\mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{t_{av};S,0}) \overset{\bullet}{=} \mathbf{1}_{t_{av}\neq 0}t_{av}^{-1}\int_0^{t_{av}} \mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{t_{av};S+\mathfrak{r},0})d\mathfrak{r} + \mathbf{1}_{t_{av}=0}\mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{t_{av};S,0}). \tag{5.22b}$$

The above space-time averages only care about what happens/are functionals along the $\eta_{t_{av}}$-exclusion process. Let us define the following pair of events for maximal processes and time-averages. The same comment about maximal processes that we made in Definition 5.5 applies here upon replacing $\mathfrak{g}_{S,0}$ of the original $\Omega_{\mathbb{Z}}$-valued process by $\mathfrak{g}_{t_{av};S,0}$ of the $\mathfrak{I}_{t_{av}}$-local process:

$$\mathbf{1}[\mathscr{G}_{t_{av};\lesssim;S,0}^{\beta_-}] \overset{\bullet}{=} \mathbf{1}\left(\sup_{0\leqslant t\leqslant t_{av}} t \cdot t_{av}^{-1}|\mathscr{A}_t^{\mathbf{T},+}\mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{t_{av};S,0})| \lesssim N^{-\beta_-}\right) \tag{5.23a}$$

$$\mathbf{1}[\mathscr{G}_{t_{av};\geqslant;S,0}^{\beta_+}] \overset{\bullet}{=} \mathbf{1}\left(\sup_{0\leqslant t\leqslant t_{av}} t \cdot t_{av}^{-1}|\mathscr{A}_t^{\mathbf{T},+}\mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{t_{av};S,0})| \geqslant N^{-\beta_+}\right). \tag{5.23b}$$

where the exponents $\beta_-, \beta_+ \in \mathbb{R}_{\geqslant 0}$ are universal and satisfy $\beta_- \leqslant \beta_+$. Recalling $t_{N,\varepsilon} \in \mathbb{R}_{\geqslant 0}$ from the statement of Proposition 5.6, we then define the following two-sided cutoffs of the previous $\mathfrak{I}_{t_{av}}$-local space-time averages:

$$\mathscr{C}_{t_{av};\beta_-,\beta_+}^{\mathbf{T},+,t_{N,\varepsilon},1}\mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{t_{av};S,0}) \overset{\bullet}{=} \mathscr{A}_{t_{av}}^{\mathbf{T},+}\mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{t_{av};S+t_{N,\varepsilon},0}) \cdot \mathbf{1}[\mathscr{G}_{t_{av};\lesssim;S+t_{N,\varepsilon},0}^{\beta_-}]\mathbf{1}[\mathscr{G}_{t_{av};\geqslant;S,0}^{\beta_+}] \tag{5.24a}$$

$$\mathscr{C}_{t_{av};\beta_-,\beta_+}^{\mathbf{T},+,t_{N,\varepsilon},2}\mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{t_{av};S,0}) \overset{\bullet}{=} \mathscr{A}_{t_{av}}^{\mathbf{T},+}\mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{t_{av};S,0}) \cdot \mathbf{1}[\mathscr{G}_{t_{av};\lesssim;S,0}^{\beta_-}]\mathbf{1}[\mathscr{G}_{t_{av};\geqslant;S+t_{N,\varepsilon},0}^{\beta_+}]. \tag{5.24b}$$

The clarifying remarks we made for cutoffs of space-time averages in Definition 5.5 hold for the above $\mathfrak{I}_{t_{av}}$-local averages.

By speed-of-propagation estimates and a coupling very similar to that used in the proof of Lemma 3.15 we get the following replacement of mesoscopic dynamic functionals of the original $\Omega_{\mathbb{Z}}$-valued exclusion process with their local versions.

**Lemma 5.22.** *We have the following for $C \in \mathbb{R}_{>0}$ arbitrarily large but universal, $\mathfrak{i} \in \{1, 2\}$, and any deterministic $t_{av} \lesssim N^{10}$:*

$$\mathbf{E}_{\eta_0}^{\mathrm{path}}|\mathscr{A}_{t_{av}}^{\mathbf{T},+}\mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{0,0})|^2 = \mathbf{E}_{\Pi_{\mathfrak{I}_{t_{av}}}\eta_0}^{\mathrm{loc}}|\mathscr{A}_{t_{av}}^{\mathbf{T},+}\mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{t_{av};0,0})|^2 + \mathscr{O}_C(N^{-C}) \tag{5.25}$$

$$\mathbf{E}_{\eta_0}^{\mathrm{path}}|\mathscr{C}_{t_{av};\beta_-,\beta_+}^{\mathbf{T},+,t_{N,\varepsilon},\mathfrak{i}}\mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{0,0})| = \mathbf{E}_{\Pi_{\mathfrak{I}_{t_{av}}}\eta_0}^{\mathrm{loc}}|\mathscr{C}_{t_{av};\beta_-,\beta_+}^{\mathbf{T},+,t_{N,\varepsilon},\mathfrak{i}}\mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{t_{av};0,0})| + \mathscr{O}_C(N^{-C}). \tag{5.26}$$

- *The expectation $\mathbf{E}_{\eta_0}^{\text{path}}$ is with respect to the path-space measure induced by the $\mathfrak{S}^{N,!!}$-dynamic on $\Omega_{\mathbb{Z}}$ with initial spins $\eta_0 \in \Omega_{\mathbb{Z}}$.*
- *The expectation $\mathbf{E}_{\Pi_{\mathfrak{I}_{\mathrm{t_{av}}}}\eta_0}^{\text{loc}}$ with respect to the path-space measure under the $\mathfrak{I}_{\mathrm{t_{av}}}$-local process on $\Omega_{\mathfrak{I}_{\mathrm{t_{av}}}}$ with initial spins $\Pi_{\mathfrak{I}_{\mathrm{t_{av}}}}\eta_0$.*
- *The functionals inside expectations depend only on a finite time-interval, but the two expectations are defined for infinite time.*

We take advantage of localization via local equilibrium Lemma 3.8 and equilibrium calculations in Lemmas 3.10 and 3.11.

**Lemma 5.23.** *With a universal implied constant, we have the following estimate in which $\varepsilon_X = \max_i \varepsilon_{X,i}$:*

$$N^{1+\varepsilon_{\text{cpt}}+\frac{1}{2}\varepsilon-\frac{1}{2}\beta_-}\mathfrak{t}_{\mathfrak{f}}\mathbf{E}^{\mu_{0,\mathbb{Z}}}\bar{\mathfrak{f}}_{\mathfrak{t}_{\mathfrak{s}},\mathfrak{t}_{\mathfrak{f}},\mathbb{T}_N}^{\mathfrak{I}_{\mathrm{t_{av}}}}\mathbf{E}_{\Pi_{\mathfrak{I}_{\mathrm{t_{av}}}}\eta_0}^{\text{loc}}|\mathscr{C}_{\mathrm{t_{av}};\beta_-,\beta_+}^{\mathbf{T},+,\mathfrak{t}_{N,\varepsilon},\mathrm{i}}\mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{\mathrm{t_{av}};0,0})|$$
$$\lesssim N^{-\frac{3}{4}\beta_-+10\varepsilon_X} + N^{-\frac{1}{2}+10\varepsilon_X} + N^{-\beta_{\text{univ}}}\mathfrak{t}_{\mathfrak{f}}. \tag{5.27}$$

*Proof of Proposition 5.6.* We use Lemma 5.18 to get the first bound and the second estimate in Lemma 5.22 to get the second bound:

$$\mathbf{E}\|\bar{\mathfrak{P}}_{T,x}^{N}(N^{\frac{1}{2}}|\mathscr{C}_{\mathrm{t_{av}};\beta_-,\beta_+}^{\mathbf{T},+,\mathfrak{t}_{N,\varepsilon},\mathrm{i}}\mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{S+\mathfrak{t}_{\mathfrak{s}},y})|)\|_{\mathfrak{t}_{\mathfrak{f}};\mathbb{T}_N}$$
$$\lesssim N^{1+\varepsilon_{\text{cpt}}+\frac{1}{2}\varepsilon-\frac{1}{2}\beta_-}\mathfrak{t}_{\mathfrak{f}}\mathbf{E}^{\mu_{0,\mathbb{Z}}}\bar{\mathfrak{f}}_{\mathfrak{t}_{\mathfrak{s}},\mathfrak{t}_{\mathfrak{f}},\mathbb{T}_N}\mathbf{E}_{\eta_0}^{\text{path}}|\mathscr{C}_{\mathrm{t_{av}};\beta_-,\beta_+}^{\mathbf{T},+,\mathfrak{t}_{N,\varepsilon},\mathrm{i}}\mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{0,0})| + N^{-\varepsilon} \tag{5.28}$$
$$\lesssim N^{1+\varepsilon_{\text{cpt}}+\frac{1}{2}\varepsilon-\frac{1}{2}\beta_-}\mathfrak{t}_{\mathfrak{f}}\mathbf{E}^{\mu_{0,\mathbb{Z}}}\bar{\mathfrak{f}}_{\mathfrak{t}_{\mathfrak{s}},\mathfrak{t}_{\mathfrak{f}},\mathbb{T}_N}\mathbf{E}_{\Pi_{\mathfrak{I}_{\mathrm{t_{av}}}}\eta_0}^{\text{loc}}|\mathscr{C}_{\mathrm{t_{av}};\beta_-,\beta_+}^{\mathbf{T},+,\mathfrak{t}_{N,\varepsilon},\mathrm{i}}\mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{\mathrm{t_{av}};0,0})| + N^{-\varepsilon}$$
$$\tag{5.29}$$
$$= N^{1+\varepsilon_{\text{cpt}}+\frac{1}{2}\varepsilon-\frac{1}{2}\beta_-}\mathfrak{t}_{\mathfrak{f}}\mathbf{E}^{\mu_{0,\mathbb{Z}}}\bar{\mathfrak{f}}_{\mathfrak{t}_{\mathfrak{s}},\mathfrak{t}_{\mathfrak{f}},\mathbb{T}_N}^{\mathfrak{I}_{\mathrm{t_{av}}}}\mathbf{E}_{\Pi_{\mathfrak{I}_{\mathrm{t_{av}}}}\eta_0}^{\text{loc}}|\mathscr{C}_{\mathrm{t_{av}};\beta_-,\beta_+}^{\mathbf{T},+,\mathfrak{t}_{N,\varepsilon},\mathrm{i}}\mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{\mathrm{t_{av}};0,0})| + N^{-\varepsilon}.$$
$$\tag{5.30}$$

The last line above follows by realizing that after the replacement-by-local expectation via Lemma 5.22, the remaining inner expectation depends only on the particle configuration in the subset $\mathfrak{I}_{\mathrm{t_{av}}} \subseteq \mathbb{Z}$. In particular, we may project the density $\bar{\mathfrak{f}}$ from the line before onto this block in the outer expectation. It then suffices to apply Lemma 5.23 to the first term in (5.30). □

*5.8. Proof of Proposition 5.11.* The proof for Proposition 5.11 follows by almost identical considerations as in the proof of Proposition 5.6. We list here a few differences. First we note Proposition 5.11 is concerned with order $N^{\beta_X}$ terms as opposed to order $N^{1/2}$ terms in Proposition 5.6. This changes the relevant power-counting for the proof of Proposition 5.11 "in our favor". Second, the $\widetilde{\mathfrak{g}}^{\mathfrak{l}}$-terms in Proposition 5.11 only exhibit time-averages with cutoff; they are not being spatially-averaged via $\mathscr{A}^{\mathbf{X},-}$ operators. This works "against our favor" as we cannot leverage improved a priori estimates that we depended on in the proof of Proposition 5.6. The two changes cooperate well enough for our power-counting to give us the upper bound in the proposed estimate. We clarify that in our local equilibrium calculations for time-averages, the support of $\widetilde{\mathfrak{g}}^{\mathfrak{l}}$-terms grows too fast to only apply Lemma 3.11 to estimate Sobolev norms. We will also require Lemma 3.12 to reduce Sobolev estimates of $\widetilde{\mathfrak{g}}^{\mathfrak{l}}$ to those of a pseudo-gradient factor with uniformly bounded support. We now inherit notation of Proposition 5.11 for this subsection.

Similar to the proof for Proposition 5.6, we will present preliminary ingredients, use them to get Proposition 5.11, and then defer the proof of the preliminary ingredients to the last subsection. The first preliminary is an analog of gymnastics in Lemma 5.17 and Lemma 5.18 with an almost identical proof. Note the below recovers Lemma 5.17 if $\beta_X$ were actually equal to $\frac{1}{2}$.

**Lemma 5.24.** *Provided the setting of* Lemma *5.17, for any* $\mathfrak{i} \in \{1, 2\}$, *we have deterministic and expectation estimates:*

$$\|\bar{\mathfrak{P}}^N_{T,x}(N^{\beta_X}|\varphi_{S+\mathfrak{t}_\mathfrak{s},y}|)\|_{\mathfrak{t}_\mathfrak{f};\mathbb{T}_N}$$
$$\lesssim N^{\frac{3}{2}\beta_X+\frac{1}{4}+\varepsilon_{\mathrm{cpt}}+\frac{1}{2}\varepsilon-\frac{1}{2}\beta_-} \int_0^{\mathfrak{t}_\mathfrak{f}} \widetilde{\sum}_{y\in\mathbb{T}_N} |\varphi_{S+\mathfrak{t}_\mathfrak{s},y}|\,\mathrm{d}S \; + \; N^{-\varepsilon} \tag{5.31}$$

$$\mathbf{E}\|\bar{\mathfrak{P}}^N_{T,x}(N^{\beta_X}|\mathscr{C}^{\mathbf{T},+,\mathfrak{t}_{N,\varepsilon},\mathfrak{i}}_{\mathfrak{t}_{\mathrm{av}};\beta_-,\beta_+}(\widetilde{\mathfrak{g}}^{\mathfrak{l}}_{S+\mathfrak{t}_\mathfrak{s},y})|)\|_{\mathfrak{t}_\mathfrak{f};\mathbb{T}_N}$$
$$\lesssim N^{\frac{3}{2}\beta_X+\frac{1}{4}+\varepsilon_{\mathrm{cpt}}+\frac{1}{2}\varepsilon-\frac{1}{2}\beta_-}\mathfrak{t}_\mathfrak{f}\mathbf{E}^{\mu_{0,\mathbb{Z}}}\bar{\mathfrak{f}}_{\mathfrak{t}_\mathfrak{s},\mathfrak{t}_\mathfrak{f},\mathbb{T}_N}\mathbf{E}^{\mathrm{path}}_{\eta_0}|\mathscr{C}^{\mathbf{T},+,\mathfrak{t}_{N,\varepsilon},\mathfrak{i}}_{\mathfrak{t}_{\mathrm{av}};\beta_-,\beta_+}(\widetilde{\mathfrak{g}}^{\mathfrak{l}}_{0,0})| \; + \; N^{-\varepsilon}. \tag{5.32}$$

*Above the expectation* $\mathbf{E}^{\mathrm{path}}_{\eta_0}$ *is an expectation with respect to the path-space measure induced by the original* $\Omega_\mathbb{Z}$*-valued exclusion process with an initial configuration* $\eta_0 \in \Omega_\mathbb{Z}$ *that is then being averaged via* $\mathbf{E}^{\mu_{0,\mathbb{Z}}}$ *against the space-time averaged density* $\bar{\mathfrak{f}}_{\mathfrak{t}_\mathfrak{s},\mathfrak{t}_\mathfrak{f},\mathbb{T}_N}$.

Parallel to the proof of Proposition 5.6, we localize the inner expectation on the RHS in the second estimate in Lemma 5.24 to reduce our estimates for it to local equilibrium calculations. The localization mechanism follows similar constructions and procedure that we used in Lemma 5.22. This starts with the following local versions of Definition 5.8 and Definition 5.10.

**Definition 5.25.** We recall the $\mathfrak{I}_{\mathfrak{t}_{\mathrm{av}}}$ block in Definition 5.19 and the $\eta_{\mathfrak{t}_{\mathrm{av}};\bullet}$-process of the first paragraph of Definition 5.21.

- Define $\widetilde{\mathfrak{g}}^{\mathfrak{l}}_{\mathfrak{t}_{\mathrm{av}};S,0} \overset{\bullet}{=} \widetilde{\mathfrak{g}}^{\mathfrak{l}}_{0,0}(\eta_{\mathfrak{t}_{\mathrm{av}};S})$ with pseudo-gradient factor as in Proposition 2.7 but evaluated along the $\mathfrak{I}_{\mathfrak{t}_{\mathrm{av}}}$-local process.
- We now define the following $\mathfrak{I}_{\mathfrak{t}_{\mathrm{av}}}$-local time-average that only sees data of the $\eta_{\mathfrak{t}_{\mathrm{av}};\bullet}$-exclusion process:

$$\mathscr{A}^{\mathbf{T},+}_{\mathfrak{t}_{\mathrm{av}}}(\widetilde{\mathfrak{g}}^{\mathfrak{l}}_{\mathfrak{t}_{\mathrm{av}};S,0}) \overset{\bullet}{=} \mathbf{1}_{\mathfrak{t}_{\mathrm{av}}\neq 0}\mathfrak{t}^{-1}_{\mathrm{av}}\int_0^{\mathfrak{t}_{\mathrm{av}}} \widetilde{\mathfrak{g}}^{\mathfrak{l}}_{\mathfrak{t}_{\mathrm{av}};S+\mathfrak{r},0}\mathrm{d}\mathfrak{r} \; + \; \mathbf{1}_{\mathfrak{t}_{\mathrm{av}}=0}\widetilde{\mathfrak{g}}^{\mathfrak{l}}_{\mathfrak{t}_{\mathrm{av}};S,0}. \tag{5.33}$$

- Let us define the following events for maximal processes and time-averages. We will again adopt the same comment about maximal processes we gave in Definition 5.10 upon the replacement of the functional $\widetilde{\mathfrak{g}}^{\mathfrak{l}}$ of the original $\Omega_\mathbb{Z}$-valued exclusion process by the functional $\widetilde{\mathfrak{g}}^{\mathfrak{l}}_{\mathfrak{t}_{\mathrm{av}};S,0}$ of the $\mathfrak{I}_{\mathfrak{t}_{\mathrm{av}}}$-local exclusion process:

$$\mathbf{1}[\mathscr{F}^{\beta_-}_{\mathfrak{t}_{\mathrm{av}};\leqslant;S,0}] \overset{\bullet}{=} \mathbf{1}\left(\sup_{0\leqslant\mathfrak{t}\leqslant\mathfrak{t}_{\mathrm{av}}}\mathfrak{t}\cdot\mathfrak{t}^{-1}_{\mathrm{av}}|\mathscr{A}^{\mathbf{T},+}_{\mathfrak{t}}(\widetilde{\mathfrak{g}}^{\mathfrak{l}}_{\mathfrak{t}_{\mathrm{av}};S,0})| \lesssim N^{-\beta_-}\right) \tag{5.34a}$$

$$\mathbf{1}[\mathscr{F}^{\beta_+}_{\mathfrak{t}_{\mathrm{av}};\geqslant;S,0}] \overset{\bullet}{=} \mathbf{1}\left(\sup_{0\leqslant\mathfrak{t}\leqslant\mathfrak{t}_{\mathrm{av}}}\mathfrak{t}\cdot\mathfrak{t}^{-1}_{\mathrm{av}}|\mathscr{A}^{\mathbf{T},+}_{\mathfrak{t}}(\widetilde{\mathfrak{g}}^{\mathfrak{l}}_{\mathfrak{t}_{\mathrm{av}};S,0})| \geqslant N^{-\beta_+}\right). \tag{5.34b}$$

Above the exponents $\beta_-, \beta_+ \in \mathbb{R}_{\geqslant 0}$ are deterministic, universal, and assumed to satisfy $\beta_- \leqslant \beta_+$. Recalling $\mathfrak{t}_{N,\varepsilon} \in \mathbb{R}_{\geqslant 0}$ from the statement of Proposition 5.11, we

define the following two-sided cutoffs of the previous $\mathfrak{I}_{\mathsf{t}_{\mathrm{av}}}$-local time-average:

$$\mathscr{C}_{\mathsf{t}_{\mathrm{av}};\beta_-,\beta_+}^{\mathbf{T},+,\mathsf{t}_{N,\varepsilon},1}(\widetilde{\mathfrak{g}}_{\mathsf{t}_{\mathrm{av}};S,0}^{\mathfrak{l}}) \overset{\bullet}{=} \mathscr{A}_{\mathsf{t}_{\mathrm{av}}}^{\mathbf{T},+}(\widetilde{\mathfrak{g}}_{\mathsf{t}_{\mathrm{av}};S+\mathsf{t}_{N,\varepsilon},0}^{\mathfrak{l}}) \cdot \mathbf{1}[\mathscr{F}_{\mathsf{t}_{\mathrm{av}};\lesssim;S+\mathsf{t}_{N,\varepsilon},0}^{\beta_-}]\mathbf{1}[\mathscr{F}_{\mathsf{t}_{\mathrm{av}};\geqslant;S,0}^{\beta_+}]$$
$$(5.35a)$$

$$\mathscr{C}_{\mathsf{t}_{\mathrm{av}};\beta_-,\beta_+}^{\mathbf{T},+,\mathsf{t}_{N,\varepsilon},2}(\widetilde{\mathfrak{g}}_{\mathsf{t}_{\mathrm{av}};S,0}^{\mathfrak{l}}) \overset{\bullet}{=} \mathscr{A}_{\mathsf{t}_{\mathrm{av}}}^{\mathbf{T},+}(\widetilde{\mathfrak{g}}_{\mathsf{t}_{\mathrm{av}};S,0}^{\mathfrak{l}}) \cdot \mathbf{1}[\mathscr{F}_{\mathsf{t}_{\mathrm{av}};\lesssim;S,0}^{\beta_-}]\mathbf{1}[\mathscr{F}_{\mathsf{t}_{\mathrm{av}};\geqslant;S+\mathsf{t}_{N,\varepsilon},0}^{\beta_+}]. \quad (5.35b)$$

The following preliminary ingredient to the proof of Proposition 5.11 is the parallel to Lemma 5.22 but for time-averages of $\widetilde{\mathfrak{g}}^{\mathfrak{l}}$ and their local versions. Its proof is exactly the same as proof of Lemma 5.22 because we are still comparing a time-average on the time-scale $\mathsf{t}_{\mathrm{av}} \in \mathbb{R}_{\geqslant 0}$ of a functional with support of size $N^{\beta_X}$ times universal factors to its $\mathfrak{I}_{\mathsf{t}_{\mathrm{av}}}$-local version. We write the following localization-via-coupling bound precisely but it is just Lemma 5.22 upon replacing $\mathscr{C}^{\mathbf{X},-}(\mathfrak{g})$-content by $\widetilde{\mathfrak{g}}^{\mathfrak{l}}$-content.

**Lemma 5.26.** *We have the following for $C \in \mathbb{R}_{>0}$ arbitrarily large but universal, $\mathfrak{i} \in \{1, 2\}$, and any deterministic $\mathsf{t}_{\mathrm{av}} \lesssim N^{10}$:*

$$\mathbf{E}_{\eta_0}^{\mathrm{path}}|\mathscr{A}_{\mathsf{t}_{\mathrm{av}}}^{\mathbf{T},+}(\widetilde{\mathfrak{g}}_{0,0}^{\mathfrak{l}})|^2 = \mathbf{E}_{\Pi_{\mathfrak{I}_{\mathsf{t}_{\mathrm{av}}}}\eta_0}^{\mathrm{loc}}|\mathscr{A}_{\mathsf{t}_{\mathrm{av}}}^{\mathbf{T},+}(\widetilde{\mathfrak{g}}_{\mathsf{t}_{\mathrm{av}};0,0}^{\mathfrak{l}})|^2 + \mathscr{O}_C(N^{-C}) \quad (5.36)$$

$$\mathbf{E}_{\eta_0}^{\mathrm{path}}|\mathscr{C}_{\mathsf{t}_{\mathrm{av}};\beta_-,\beta_+}^{\mathbf{T},+,\mathsf{t}_{N,\varepsilon},\mathfrak{i}}(\widetilde{\mathfrak{g}}_{0,0}^{\mathfrak{l}})| = \mathbf{E}_{\Pi_{\mathfrak{I}_{\mathsf{t}_{\mathrm{av}}}}\eta_0}^{\mathrm{loc}}|\mathscr{C}_{\mathsf{t}_{\mathrm{av}};\beta_-,\beta_+}^{\mathbf{T},+,\mathsf{t}_{N,\varepsilon},\mathfrak{i}}(\widetilde{\mathfrak{g}}_{\mathsf{t}_{\mathrm{av}};0,0}^{\mathfrak{l}})| + \mathscr{O}_C(N^{-C}). \quad (5.37)$$

- *The expectation $\mathbf{E}_{\eta_0}^{\mathrm{path}}$ is with respect to the path-space measure induced by the $\mathfrak{S}^{N,!!}$-dynamic on $\Omega_{\mathbb{Z}}$ with initial spins $\eta_0 \in \Omega_{\mathbb{Z}}$.*
- *The expectation $\mathbf{E}_{\Pi_{\mathfrak{I}_{\mathsf{t}_{\mathrm{av}}}}\eta_0}^{\mathrm{loc}}$ with respect to the path-space measure under the $\mathfrak{I}_{\mathsf{t}_{\mathrm{av}}}$-local process on $\Omega_{\mathfrak{I}_{\mathsf{t}_{\mathrm{av}}}}$ with initial spins $\Pi_{\mathfrak{I}_{\mathsf{t}_{\mathrm{av}}}}\eta_0$.*
- *The functionals inside expectations depend only on a finite time-interval, but the two expectations are defined for infinite time.*

Parallel to Lemma 5.23 we used in the proof of Proposition 5.6, we take advantage of localization in Lemma 5.26 with the following which ultimately follows by reduction to local equilibrium via Lemma 3.8 and local equilibrium calculations.

**Lemma 5.27.** *With a universal implied constant, we have the following uniformly in $\mathfrak{l} \in [\![1, N^{\beta_X}]\!]$ in which $\varepsilon_X = \max_i \varepsilon_{X,i}$:*

$$N^{\frac{3}{2}\beta_X+\frac{1}{4}+\varepsilon_{\mathrm{cpt}}+\frac{1}{2}\varepsilon-\frac{1}{2}\beta_-}\mathsf{t}_{\mathfrak{f}}\mathbf{E}^{\mu_{0,\mathbb{Z}}}\bar{\mathfrak{f}}_{\mathsf{t}_{\mathfrak{s}},\mathsf{t}_{\mathfrak{f}},\mathbb{T}_N}^{\mathfrak{I}_{\mathsf{t}_{\mathrm{av}}}}\mathbf{E}_{\Pi_{\mathfrak{I}_{\mathsf{t}_{\mathrm{av}}}}\eta_0}^{\mathrm{loc}}|\mathscr{C}_{\mathsf{t}_{\mathrm{av}};\beta_-,\beta_+}^{\mathbf{T},+,\mathsf{t}_{N,\varepsilon},\mathfrak{i}}(\widetilde{\mathfrak{g}}_{\mathsf{t}_{\mathrm{av}};0,0}^{\mathfrak{l}})|$$
$$\lesssim N^{-\frac{1}{8}-\frac{3}{4}\beta_-+10\varepsilon_X} + N^{-\frac{1}{2}+10\varepsilon_X} + \mathsf{t}_{\mathfrak{f}}N^{-\beta_{\mathrm{univ}}}. \quad (5.38)$$

*Proof of Proposition 5.11.* Following the proof of Proposition 5.6, we use Lemma 5.24 and the second bound in Lemma 5.26:

$$\mathbf{E}\|\bar{\mathfrak{P}}_{T,x}^N(N^{\frac{1}{2}}|\mathscr{C}_{\mathsf{t}_{\mathrm{av}};\beta_-,\beta_+}^{\mathbf{T},+,\mathsf{t}_{N,\varepsilon},\mathfrak{i}}(\widetilde{\mathfrak{g}}_{S+\mathsf{t}_{\mathfrak{s}},y}^{\mathfrak{l}})|)\|_{\mathsf{t}_{\mathfrak{f}};\mathbb{T}_N}$$
$$\lesssim N^{\frac{3}{2}\beta_X+\frac{1}{4}+\varepsilon_{\mathrm{cpt}}+\frac{1}{2}\varepsilon-\frac{1}{2}\beta_-}\mathsf{t}_{\mathfrak{f}}\mathbf{E}^{\mu_{0,\mathbb{Z}}}\bar{\mathfrak{f}}_{\mathsf{t}_{\mathfrak{s}},\mathsf{t}_{\mathfrak{f}},\mathbb{T}_N}\mathbf{E}_{\eta_0}^{\mathrm{path}}|\mathscr{C}_{\mathsf{t}_{\mathrm{av}};\beta_-,\beta_+}^{\mathbf{T},+,\mathsf{t}_{N,\varepsilon},\mathfrak{i}}(\widetilde{\mathfrak{g}}_{0,0}^{\mathfrak{l}})| + N^{-\varepsilon} \quad (5.39)$$
$$\lesssim N^{\frac{3}{2}\beta_X+\frac{1}{4}+\varepsilon_{\mathrm{cpt}}+\frac{1}{2}\varepsilon-\frac{1}{2}\beta_-}\mathsf{t}_{\mathfrak{f}}\mathbf{E}^{\mu_{0,\mathbb{Z}}}\bar{\mathfrak{f}}_{\mathsf{t}_{\mathfrak{s}},\mathsf{t}_{\mathfrak{f}},\mathbb{T}_N}^{\mathfrak{I}_{\mathsf{t}_{\mathrm{av}}}}\mathbf{E}_{\Pi_{\mathfrak{I}_{\mathsf{t}_{\mathrm{av}}}}\eta_0}^{\mathrm{loc}}|\mathscr{C}_{\mathsf{t}_{\mathrm{av}};\beta_-,\beta_+}^{\mathbf{T},+,\mathsf{t}_{N,\varepsilon},\mathfrak{i}}(\widetilde{\mathfrak{g}}_{\mathsf{t}_{\mathrm{av}};0,0}^{\mathfrak{l}})| + N^{-\varepsilon}.$$
$$(5.40)$$

We clarify (5.40) also requires the observation that the inner expectation therein depends only on the initial particle configuration on the block $\mathfrak{I}_{\mathsf{t}_{\mathrm{av}}}$. This allows us to project the probability density $\bar{\mathfrak{f}}$ in (5.40) on its marginal distribution on the block $\mathfrak{I}_{\mathsf{t}_{\mathrm{av}}}$. It now suffices to apply the estimate in Lemma 5.27 to the first term in (5.40) to finish the proof.                                                                                       $\square$

*5.9. Proof of Proposition 5.4.* First we declare that we adopt all notation within the statement of Proposition 5.4. The proof of Proposition 5.4 is easier because of the short-time factor $\mathfrak{t}_{\mathrm{av}}^{1/4}$ on the LHS of the proposed bound. However, it follows the same general scheme of the proof of Proposition 5.6, for example. Precisely, we use technical gymnastics to remove the singularity in the heat kernel inside the heat operator on the LHS of the proposed estimate in Proposition 5.4. We then localize mesoscopic space-time averages, apply local equilibrium via Lemma 3.8, and perform equilibrium calculations via Lemmas 3.10 and 3.11.

The following first ingredient is slightly different than its analog Lemma 5.17. For example, it squares the test function.

**Lemma 5.28.** *Take any $\varphi : \mathbb{R}_{\geq 0} \times \mathbb{T}_N \to \mathbb{R}$. We have a deterministic and expectation bound with universal implied constants:*

$$\left( \| \bar{\mathfrak{P}}_{T,x}^N (N^{\frac{1}{2}} |\varphi_{S,y}|) \|_{1;\mathbb{T}_N} \right)^2 \lesssim N^{\frac{5}{4}+\varepsilon_{\mathrm{cpt}}} \int_0^1 \widetilde{\sum_{y \in \mathbb{T}_N}} |\varphi_{S,y}|^2 \, \mathrm{d}S \qquad (5.41)$$

$$\left( \mathbf{E} \| \bar{\mathfrak{P}}_{T,x}^N (N^{\frac{1}{2}} |\mathscr{A}_{\mathfrak{t}_{\mathrm{av}}}^{\mathbf{T},+} \mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-} (\mathfrak{g}_{S,y})|) \|_{1;\mathbb{T}_N} \right)^2 \lesssim N^{\frac{5}{4}+\varepsilon_{\mathrm{cpt}}} \mathbf{E}^{\mu_{0,\mathbb{Z}}} \bar{\mathfrak{f}}_{1,\mathbb{T}_N} \mathbf{E}_{\eta_0}^{\mathrm{path}} |\mathscr{A}_{\mathfrak{t}_{\mathrm{av}}}^{\mathbf{T},+} \mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-} (\mathfrak{g}_{0,0})|^2. \qquad (5.42)$$

Following the proof of Proposition 5.6, we localize the path-space expectation on the RHS of the second estimate in Lemma 5.28. The error is controlled already in Lemma 5.22. Thus, parallel to Lemma 5.23, we take advantage of localization below.

**Lemma 5.29.** *Recall the $\mathfrak{I}_{\mathfrak{t}_{\mathrm{av}}}$ block in Definition 5.19. Provided any $\varepsilon \in \mathbb{R}_{>0}$, we have*

$$N^{\frac{5}{4}+\varepsilon_{\mathrm{cpt}}} \mathfrak{t}_{\mathrm{av}}^{\frac{1}{2}} \mathbf{E}^{\mu_{0,\mathbb{Z}}} \bar{\mathfrak{f}}_{1,\mathbb{T}_N}^{\mathfrak{I}_{\mathfrak{t}_{\mathrm{av}}}} \mathbf{E}_{\Pi_{\mathfrak{I}_{\mathfrak{t}_{\mathrm{av}}}} \eta_0}^{\mathrm{loc}} |\mathscr{A}_{\mathfrak{t}_{\mathrm{av}}}^{\mathbf{T},+} \mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-} (\mathfrak{g}_{\mathfrak{t}_{\mathrm{av}};0,0})|^2 \qquad (5.43)$$

$$\lesssim_\varepsilon N^{\frac{5}{4}+\varepsilon_{\mathrm{cpt}}} \mathfrak{t}_{\mathrm{av}}^{\frac{1}{2}} N^{-\frac{3}{4}-\beta_X+\varepsilon_{X,1}+\varepsilon} |\mathbb{T}_N|^{-1} |\mathfrak{I}_{\mathfrak{t}_{\mathrm{av}}}|^3 \; + \; N^{\frac{5}{4}+\varepsilon_{\mathrm{cpt}}} N^{-2-\beta_X} \mathfrak{t}_{\mathrm{av}}^{-\frac{1}{2}}. \qquad (5.44)$$

The $\mathfrak{t}_{\mathrm{av}}^{1/2}$-factor on the LHS comes by squaring the "time-regularity factor" $\mathfrak{t}_{\mathrm{av}}^{1/4}$ in the proposed bound of Proposition 5.4.

*Proof of Proposition 5.4.* Following the proof of Proposition 5.6, we employ the second estimate in Lemma 5.28 along with the first estimate in Lemma 5.22 that we introduced in the proof of Proposition 5.6. As $\mathfrak{t}_{\mathrm{av}} \in \mathbb{R}_{\geq 0}$ is uniformly bounded, we get

$$\left( \mathfrak{t}_{\mathrm{av}}^{\frac{1}{4}} \mathbf{E} \| \bar{\mathfrak{P}}_{T,x}^N (N^{\frac{1}{2}} |\mathscr{A}_{\mathfrak{t}_{\mathrm{av}}}^{\mathbf{T},+} \mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-} (\mathfrak{g}_{S,y})|) \|_{1;\mathbb{T}_N} \right)^2$$

$$\lesssim N^{\frac{5}{4}+\varepsilon_{\mathrm{cpt}}} \mathfrak{t}_{\mathrm{av}}^{\frac{1}{2}} \mathbf{E}^{\mu_{0,\mathbb{Z}}} \bar{\mathfrak{f}}_{1,\mathbb{T}_N} \mathbf{E}_{\eta_0}^{\mathrm{path}} |\mathscr{A}_{\mathfrak{t}_{\mathrm{av}}}^{\mathbf{T},+} \mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-} (\mathfrak{g}_{0,0})|^2 \qquad (5.45)$$

$$\lesssim N^{\frac{5}{4}+\varepsilon_{\mathrm{cpt}}} \mathfrak{t}_{\mathrm{av}}^{\frac{1}{2}} \mathbf{E}^{\mu_{0,\mathbb{Z}}} \bar{\mathfrak{f}}_{1,\mathbb{T}_N}^{\mathfrak{I}_{\mathfrak{t}_{\mathrm{av}}}} \mathbf{E}_{\Pi_{\mathfrak{I}_{\mathfrak{t}_{\mathrm{av}}}} \eta_0}^{\mathrm{loc}} |\mathscr{A}_{\mathfrak{t}_{\mathrm{av}}}^{\mathbf{T},+} \mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-} (\mathfrak{g}_{\mathfrak{t}_{\mathrm{av}};0,0})|^2 \; + \; N^{-100}. \qquad (5.46)$$

The same clarifying remarks we made after the estimate of (5.30) from the proof for Proposition 5.6 are in order here as well. Thus, it remains to estimate the first expectation from the estimate (5.46). We do this by applying Lemma 5.29 to this first term in (5.46). The proof for Proposition 5.4 then follows by power-counting upon recalling the constraint $N^{-2} \lesssim \mathfrak{t}_{\mathrm{av}} \lesssim N^{-1}$ as well as the length-scale $|\mathfrak{I}_{\mathfrak{t}_{\mathrm{av}}}| \lesssim_\varepsilon N^{1+\varepsilon} \mathfrak{t}_{\mathrm{av}}^{1/2} + N^{3/2+\varepsilon} \mathfrak{t}_{\mathrm{av}} + N^{\beta_X+\varepsilon}$ given any $\varepsilon \in \mathbb{R}_{>0}$ from Definition 5.19 and $|\mathbb{T}_N| \gtrsim N^{5/4+\varepsilon_{\mathrm{cpt}}}$ from the beginning of Sect. 4. This power-counting is ultimately elementary and thus omitted. $\qquad \square$

*5.10. Proof of Proposition 5.9.* Let us again inherit all notation and framework of Proposition 5.9. The proof of Proposition 5.9 resembles strongly the proof of Proposition 5.4. We employ the following preliminary gymnastics parallel to Lemma 5.28 to avoid the heat kernel singularity. Again, if $\beta_X$ were actually $\frac{1}{2}$, the first estimate below would recover that of Lemma 5.28.

**Lemma 5.30.** *Take any $\varphi : \mathbb{R}_{\geqslant 0} \times \mathbb{T}_N \to \mathbb{R}$. We have a deterministic and expectation bound with universal implied constants:*

$$\|\bar{\mathfrak{P}}_{T,x}^N(N^{\beta_X}|\varphi_{S,y}|)\|_{1;\mathbb{T}_N} \lesssim \left( N^{2\beta_X + \frac{1}{4} + \varepsilon_{\mathrm{cpt}}} \int_0^1 \widetilde{\sum}_{y \in \mathbb{T}_N} |\varphi_{S,y}|^2 \, \mathrm{d}S \right)^{\frac{1}{2}} \tag{5.47}$$

$$\left( \mathbf{E} \|\bar{\mathfrak{P}}_{T,x}^N(N^{\beta_X}|\mathscr{A}_{\mathfrak{t}_{\mathrm{av}}}^{\mathbf{T},+}(\widetilde{\mathfrak{g}}_{S,y}^{\mathfrak{l}})|)\|_{1;\mathbb{T}_N} \right)^2 \lesssim N^{2\beta_X + \frac{1}{4} + \varepsilon_{\mathrm{cpt}}} \mathbf{E}^{\mu_{0,\mathbb{Z}}} \bar{\mathfrak{f}}_{1,\mathbb{T}_N} \mathbf{E}_{\eta_0}^{\mathrm{path}} |\mathscr{A}_{\mathfrak{t}_{\mathrm{av}}}^{\mathbf{T},+}(\widetilde{\mathfrak{g}}_{0,0}^{\mathfrak{l}})|^2. \tag{5.48}$$

From Lemma 5.30 by following the proof of Proposition 5.11, we localize the inner path-space expectation on the RHS of the second bound in Lemma 5.30. Similar to the proof of Proposition 5.4, this is actually already done in Lemma 5.26. We now take advantage of localization via the following parallel to Lemma 5.27 and Lemma 5.29.

**Lemma 5.31.** *Provided any $\varepsilon \in \mathbb{R}_{>0}$, uniformly in $\mathfrak{l} \in [\![ 1, N^{\beta_X} ]\!]$ we have*

$$N^{2\beta_X + \frac{1}{4} + \varepsilon_{\mathrm{cpt}}} \mathfrak{t}_{\mathrm{av}}^{\frac{1}{2}} \mathbf{E}^{\mu_{0,\mathbb{Z}}} \bar{\mathfrak{f}}_{1,\mathbb{T}_N}^{\mathfrak{J}_{\mathfrak{t}_{\mathrm{av}}}} \mathbf{E}_{\Pi_{\mathfrak{J}_{\mathfrak{t}_{\mathrm{av}}}}\eta_0}^{\mathrm{loc}} |\mathscr{A}_{\mathfrak{t}_{\mathrm{av}}}^{\mathbf{T},+}(\widetilde{\mathfrak{g}}_{\mathfrak{t}_{\mathrm{av}};0,0}^{\mathfrak{l}})|^2$$

$$\lesssim_\varepsilon N^{2\beta_X + \frac{1}{4} + \varepsilon_{\mathrm{cpt}}} \mathfrak{t}_{\mathrm{av}}^{\frac{1}{2}} N^{-\frac{3}{4} + \varepsilon} |\mathbb{T}_N|^{-1} |\mathfrak{J}_{\mathfrak{t}_{\mathrm{av}}}|^3 + N^{2\beta_X + \frac{1}{4} + \varepsilon_{\mathrm{cpt}}} N^{-2} \mathfrak{t}_{\mathrm{av}}^{-\frac{1}{2}}. \tag{5.49}$$

*Proof of Proposition 5.9.* We follow somewhat the proof of Proposition 5.4. To start, we employ the gymnastics of Lemma 5.30 and the localization estimate (5.36) of Lemma 5.26. This is in parallel to the first step in the proof of Proposition 5.4:

$$\left( \mathfrak{t}_{\mathrm{av}}^{\frac{1}{4}} \mathbf{E} \|\bar{\mathfrak{P}}_{T,x}^N(N^{\beta_X}|\mathscr{A}_{\mathfrak{t}_{\mathrm{av}}}^{\mathbf{T},+}(\widetilde{\mathfrak{g}}_{S,y}^{\mathfrak{l}})|)\|_{1;\mathbb{T}_N} \right)^2$$

$$\lesssim N^{2\beta_X + \frac{1}{4} + \varepsilon_{\mathrm{cpt}}} \mathfrak{t}_{\mathrm{av}}^{\frac{1}{2}} \mathbf{E}^{\mu_{0,\mathbb{Z}}} \bar{\mathfrak{f}}_{1,\mathbb{T}_N} \mathbf{E}_{\eta_0}^{\mathrm{path}} |\mathscr{A}_{\mathfrak{t}_{\mathrm{av}}}^{\mathbf{T},+}(\widetilde{\mathfrak{g}}_{0,0}^{\mathfrak{l}})|^2 \tag{5.50}$$

$$\lesssim N^{2\beta_X + \frac{1}{4} + \varepsilon_{\mathrm{cpt}}} \mathfrak{t}_{\mathrm{av}}^{\frac{1}{2}} \mathbf{E}^{\mu_{0,\mathbb{Z}}} \bar{\mathfrak{f}}_{1,\mathbb{T}_N}^{\mathfrak{J}_{\mathfrak{t}_{\mathrm{av}}}} \mathbf{E}_{\Pi_{\mathfrak{J}_{\mathfrak{t}_{\mathrm{av}}}}\eta_0}^{\mathrm{loc}} |\mathscr{A}_{\mathfrak{t}_{\mathrm{av}}}^{\mathbf{T},+}(\widetilde{\mathfrak{g}}_{\mathfrak{t}_{\mathrm{av}};0,0}^{\mathfrak{l}})|^2 + N^{-100}. \tag{5.51}$$

The same clarifying remarks we made after the estimate of (5.40) from the proof of Proposition 5.11 are in order here as well. Like the proof of Proposition 5.4, we apply the estimate in Lemma 5.31 to the first term in (5.51). The proof for Proposition 5.9 then follows by power-counting in $N$ and recalling $N^{-2} \lesssim \mathfrak{t}_{\mathrm{av}} \lesssim N^{-1}$, again like the end of the proof of Proposition 5.4. □

*5.11. Proof of Proposition 5.2.* We turn to the proof for the static estimate in Proposition 5.2. The idea behind the proof for this estimate in Proposition 5.2 is the same as every argument we have written thus far. By this we mean gymnastics to remove the heat kernel singularity in the heat operator in Proposition 5.2 followed by a reduction to local equilibrium and equilibrium calculations. However, there are no time-averages

in Proposition 5.2, so we will apply large-deviations estimates for canonical ensembles in Lemma 3.13 and Corollary 3.14. We reemphasize that intuitively, Proposition 5.2 follows from the negligibility of the replacement $\mathscr{A}^{\mathbf{X},-} \to \mathscr{C}^{\mathbf{X},-}$ for canonical ensembles *at large-deviations scale*, from which we reduce from non-stationary measures to stationary measures by local equilibrium in Lemma 3.8. We begin with "gymnastics" that resembles Lemma 5.28.

**Lemma 5.32.** *Consider the block* $\mathfrak{I} = [\![-100\mathfrak{m}N^{\beta_X}, 100\mathfrak{m}N^{\beta_X}]\!]$. *We have the following with universal implied constant:*

$$\left(\mathbf{E}\|\tilde{\mathfrak{P}}^N_{T,x}(N^{\frac{1}{2}}|\mathscr{A}^{\mathbf{X},-}_{N^{\beta_X}}(\mathfrak{g}_{S,y}) - \mathscr{C}^{\mathbf{X},-}_{N^{\beta_X}}(\mathfrak{g}_{S,y})|)\|_{1;\mathbb{T}_N}\right)^2$$
$$\lesssim N^{\frac{5}{4}+\varepsilon_{\mathrm{cpt}}}\mathbf{E}^{\mu_{0,\mathbb{Z}}}\bar{\mathfrak{f}}^{\mathfrak{I}}_{1,\mathbb{T}_N}|\mathscr{A}^{\mathbf{X},-}_{N^{\beta_X}}(\mathfrak{g}_{0,0}) - \mathscr{C}^{\mathbf{X},-}_{N^{\beta_X}}(\mathfrak{g}_{0,0})|^2. \qquad (5.52)$$

We observe the RHS of the estimate in Lemma 5.32 is the expectation of a $\mathfrak{I}$-local statistic against the averaged law $\bar{\mathfrak{f}}$ of the particle system. The block $\mathfrak{I}$ is that in Lemma 5.32. Thus, we may again employ Lemma 3.8 to reduce estimation on the RHS of the estimate in Lemma 5.32 to a problem of equilibrium calculations. Ultimately, we get the following.

**Lemma 5.33.** *With a universal implied constant, we have the following for the same block* $\mathfrak{I} = [\![-100\mathfrak{m}N^{\beta_X}, 100\mathfrak{m}N^{\beta_X}]\!]$:

$$N^{\frac{5}{4}+\varepsilon_{\mathrm{cpt}}}\mathbf{E}^{\mu_{0,\mathbb{Z}}}\bar{\mathfrak{f}}^{\mathfrak{I}}_{1,\mathbb{T}_N}|\mathscr{A}^{\mathbf{X},-}_{N^{\beta_X}}(\mathfrak{g}_{0,0}) - \mathscr{C}^{\mathbf{X},-}_{N^{\beta_X}}(\mathfrak{g}_{0,0})|^2 \lesssim N^{-\beta_{\mathrm{univ}}}. \qquad (5.53)$$

*Proof of Proposition 5.2.* It suffices to combine the estimates of Lemma 5.32 and Lemma 5.33. $\qquad\square$

*5.12. Proofs of Technical Estimates.* We spend the rest of this section on proofs of all preliminary technical ingredients.

*Proof of Lemma 5.17.* Consider any $(T, x) \in [0, \mathfrak{t}_{\mathfrak{f}}] \times \mathbb{T}_N$ and define $\mathfrak{t}_N \overset{\bullet}{=} T - N^{-\frac{1}{2}+\beta_- -\varepsilon}$ and $\mathfrak{t}_{N,+} = \mathfrak{t}_N \vee 0$. Upon redefining $\varphi$ by a $\mathfrak{t}_{\mathfrak{s}}$ time-shift, it suffices to assume $\mathfrak{t}_{\mathfrak{s}} = 0$. We first write

$$\tilde{\mathfrak{P}}^N_{T,x}(N^{\frac{1}{2}}|\varphi_{S,y}|) = \tilde{\mathfrak{P}}^N_{T,x}(N^{\frac{1}{2}}|\varphi_{S,y}|\mathbf{1}_{\mathfrak{t}_{N,+}\leqslant S\leqslant T}) + \tilde{\mathfrak{P}}^N_{T,x}(N^{\frac{1}{2}}|\varphi_{S,y}|\mathbf{1}_{0\leqslant S\leqslant\mathfrak{t}_{N,+}})$$
$$\overset{\bullet}{=} \Phi_1 + \Phi_2. \qquad (5.54)$$

We will control $\Phi_1$ and $\Phi_2$ uniformly on the space-time set $[0, \mathfrak{t}_{\mathfrak{f}}] \times \mathbb{T}_N$. We bound $\Phi_1$ by the convolution estimate (A.3) for the heat operator which controls the heat operator by the supremum of whatever test function that it acts on times the length of the time-interval that the integral in the heat operator is defined on:

$$\tilde{\mathfrak{P}}^N_{T,x}(N^{\frac{1}{2}}|\varphi_{S,y}|\mathbf{1}_{\mathfrak{t}_{N,+}\leqslant S\leqslant T}) \lesssim N^{\frac{1}{2}}\|\varphi\|_{\infty;\mathbb{T}_N}|T - \mathfrak{t}_{N,+}|$$
$$\lesssim N^{\frac{1}{2}}\|\varphi\|_{\infty;\mathbb{T}_N}N^{-\frac{1}{2}+\beta_- -\varepsilon} \lesssim N^{-\varepsilon}. \qquad (5.55)$$

The final estimate on the far RHS of (5.55) follows from recalling the a priori estimate $|\varphi| \lesssim N^{-\beta_-}$. To estimate $\Phi_2$, we use the on-diagonal heat kernel estimate implied by (A.1). Observe that on the set $[0, \mathfrak{t}_{N,+}]$ we have the bound $\mathfrak{s}_{S,T} \gtrsim N^{-1/2+\beta_- -\varepsilon}$. This

gives a deterministic bound that uses the aforementioned lower bound to control the short-time singularity in the heat kernel:

$$\bar{\mathfrak{P}}^N_{T,x}(N^{\frac{1}{2}}|\varphi_{S,y}|\mathbf{1}_{0\leqslant S\leqslant t_{N,+}})$$

$$\lesssim N^{-1}\int_0^{t_{N,+}}\mathfrak{s}_{S,T}^{-1/2}\sum_{y\in\mathbb{T}_N}N^{\frac{1}{2}}|\varphi_{S,y}|\,dS$$

$$\lesssim N^{-1}N^{\frac{1}{4}-\frac{1}{2}\beta_-+\frac{1}{2}\varepsilon}N^{\frac{1}{2}}\int_0^{t_{N,+}}\sum_{y\in\mathbb{T}_N}|\varphi_{S,y}|\,dS. \tag{5.56}$$

From the previous display, we multiply and divide by $|\mathbb{T}_N|$ on the far RHS to get a spatial average over $\mathbb{T}_N$; recall $|\mathbb{T}_N|\lesssim N^{5/4+\varepsilon_{\mathrm{cpt}}}$ from the beginning of Sect. 4. Moreover, we may additionally extend the domain of integration $[0,t_{N,+}]\to[0,t_{\mathrm{f}}]$ after we multiply and divide by $|\mathbb{T}_N|$. This gives the following estimate which identifies the sum over $\mathbb{T}_N$ by $|\mathbb{T}_N|$ times the average over $\mathbb{T}_N$ and then performs power-counting for the host of $N$-dependent factors that appear in the display below:

$$N^{-1}N^{\frac{1}{4}-\frac{1}{2}\beta_-+\frac{1}{2}\varepsilon}N^{\frac{1}{2}}\int_0^{t_{N,+}}\sum_{y\in\mathbb{T}_N}|\varphi_{S,y}|\,dS$$

$$\lesssim N^{-1}N^{\frac{1}{4}-\frac{1}{2}\beta_-+\frac{1}{2}\varepsilon}N^{\frac{1}{2}}|\mathbb{T}_N|\int_0^{t_{N,+}}\widetilde{\sum}_{y\in\mathbb{T}_N}|\varphi_{S,y}|\,dS \tag{5.57}$$

$$\lesssim N^{1-\frac{1}{2}\beta_-+\frac{1}{2}\varepsilon+\varepsilon_{\mathrm{cpt}}}\int_0^{t_{\mathrm{f}}}\widetilde{\sum}_{y\in\mathbb{T}_N}|\varphi_{S,y}|\,dS. \tag{5.58}$$

We reiterate that the final bound follows from power-counting on the RHS of the previous line and then extending the domain of integration $[0,t_{N,+}]\to[0,t_{\mathrm{f}}]$. We also note that (5.55) and (5.56) and (5.58) provide estimates for $\Phi_1$ and $\Phi_2$ uniformly on $[0,t_{\mathrm{f}}]\times\mathbb{T}_N$ so by the triangle inequality for $\|\|_{t_{\mathrm{f}};\mathbb{T}_N}$ and by (5.54), (5.55), (5.56) and (5.58), we finish the proof. $\qquad\square$

*Proof of Lemma 5.18.* Observe the $\mathscr{C}^{\mathbf{T},+,t_{N,\varepsilon},\mathbf{i}}$ term in the heat operator on the LHS of the proposed estimate has an upper bound of order $N^{-\beta_-}$ by construction in Definition 5.5. Thus, we may use Lemma 5.17 for $\varphi$ equal to this $\mathscr{C}^{\mathbf{T},+,t_{N,\varepsilon},\mathbf{i}}$-term to get

$$\mathbf{E}\|\bar{\mathfrak{P}}^N_{T,x}(N^{\frac{1}{2}}|\mathscr{C}^{\mathbf{T},+,t_{N,\varepsilon},\mathbf{i}}_{t_{\mathrm{av}};\beta_-,\beta_+}\mathscr{C}^{\mathbf{X},-}_{N^{\beta_X}}(\mathfrak{g}_{S+t_{\mathfrak{s}}},y)|)\|_{t_{\mathrm{f}};\mathbb{T}_N}$$

$$\lesssim N^{1+\varepsilon_{\mathrm{cpt}}+\frac{1}{2}\varepsilon-\frac{1}{2}\beta_-}\int_0^{t_{\mathrm{f}}}\widetilde{\sum}_{y\in\mathbb{T}_N}\mathbf{E}|\mathscr{C}^{\mathbf{T},+,t_{N,\varepsilon},\mathbf{i}}_{t_{\mathrm{av}};\beta_-,\beta_+}\mathscr{C}^{\mathbf{X},-}_{N^{\beta_X}}(\mathfrak{g}_{S+t_{\mathfrak{s}}},y)|\,dS + N^{-\varepsilon}. \tag{5.59}$$

We now make the next observation with explanation/justification after in which we use notation to be explained as well:

$$\mathbf{E}|\mathscr{C}^{\mathbf{T},+,t_{N,\varepsilon},\mathbf{i}}_{t_{\mathrm{av}};\beta_-,\beta_+}\mathscr{C}^{\mathbf{X},-}_{N^{\beta_X}}(\mathfrak{g}_{S+t_{\mathfrak{s}}},y)| = \mathbf{E}^{\mu_{0,\mathbb{Z}}}\mathfrak{f}_{S+t_{\mathfrak{s}},N}\mathbf{E}^{\mathrm{path}}_{\eta_0}|\mathscr{C}^{\mathbf{T},+,t_{N,\varepsilon},\mathbf{i}}_{t_{\mathrm{av}};\beta_-,\beta_+}\mathscr{C}^{\mathbf{X},-}_{N^{\beta_X}}(\mathfrak{g}_0,y)|. \tag{5.60}$$

The previous statement (5.60) follows by decomposing the expectation of the *path-space functional* given by the $\mathscr{C}^{\mathbf{T},+,t_{N,\varepsilon},\mathbf{i}}$-term into the expectation with respect to the path-space measure corresponding to the original $\Omega_{\mathbb{Z}}$-valued process after time $S+t_{\mathfrak{s}}$ with an initial particle configuration given by the particle configuration at time $S+t_{\mathfrak{s}}$ and afterwards taking expectation over this initial configuration for the path-space expectation with respect to the law of the particle system $\mathfrak{f}_{S+t_{\mathfrak{s}},N}d\mu_{0,\mathbb{Z}}$ at time $S+t_{\mathfrak{s}}$.

However, this is the same thing as sampling the initial configuration for the path-space configuration according to the time-0 configuration if we pick the initial measure of $\int_{S+t_s,N} d\mu_{0,\mathbb{Z}}$. Here, we emphasize that the path-space expectation after we shift the initial measure to $\int_{S+t_s,N} d\mu_{0,\mathbb{Z}}$ is with respect to the path-space measure of the exclusion process starting at time 0 rather than time $S+t_s$. However, the law of the particle system dynamic is invariant under any time-shift as long as we pick the same initial configuration. We also get the following, which we explain afterwards, where $(\tau_y \eta)_Z = \eta_{y+Z}$ is the shift-map on $\Omega_{\mathbb{Z}}$:

$$\mathbf{E}^{\mu_{0,\mathbb{Z}}} \int_{S+t_s,N} \mathbf{E}^{\mathrm{path}}_{\eta_0} |\mathscr{C}^{\mathbf{T},+,t_{N,\varepsilon},\mathbf{i}}_{t_{av};\beta_-,\beta_+} \mathscr{C}^{\mathbf{X},-}_{N^\beta X}(\mathfrak{g}_0,y)| = \mathbf{E}^{\mu_{0,\mathbb{Z}}} \int_{S+t_s,N} \mathbf{E}^{\mathrm{path}}_{\tau_{-y}\eta_0} |\mathscr{C}^{\mathbf{T},+,t_{N,\varepsilon},\mathbf{i}}_{t_{av};\beta_-,\beta_+} \mathscr{C}^{\mathbf{X},-}_{N^\beta X}(\mathfrak{g}_0,0)|$$

$$= \mathbf{E}^{\mu_{0,\mathbb{Z}}} \tau_y \int_{S+t_s,N} \mathbf{E}^{\mathrm{path}}_{\eta_0} |\mathscr{C}^{\mathbf{T},+,t_{N,\varepsilon},\mathbf{i}}_{t_{av};\beta_-,\beta_+} \mathscr{C}^{\mathbf{X},-}_{N^\beta X}(\mathfrak{g}_0,0)|.$$

The first identity follows by the shift-invariance of the path-space expectation. In particular, on the far LHS we take the path-space expectation of particle system data supported near $y \in \mathbb{Z}$. This is identical to path-space expectation with the same data but shifted to be centered at $0 \in \mathbb{Z}$, because the path-space measure is invariant under spatial shifts if the initial configuration is kept the same. Such a "shifted" path-space expectation is what middle of the above line encodes. The second identity follows from the invariance of the measure $\mu_{0,\mathbb{Z}}$ under shift-maps $\tau_y : \Omega_{\mathbb{Z}} \to \Omega_{\mathbb{Z}}$ and that the composition $\tau_y \tau_{-y}$ is the identity on $\Omega_{\mathbb{Z}}$. We clarify the inner expectation in the middle of the last display is a function of the shifted configuration $\tau_{-y}\eta_0$ which we take expectation of against $\int_{S+t_s,N} d\mu_{0,\mathbb{Z}}$. Via the last display, (5.60), linearity of expectation/integration, and Fubini's theorem, as in the one-block step in [14] we get the following where the fourth line follows by the path-space expectation in the line before being independent of all integration/summation variables:

$$\int_0^{t_f} \widetilde{\sum}_{y \in \mathbb{T}_N} \mathbf{E} |\mathscr{C}^{\mathbf{T},+,t_{N,\varepsilon},\mathbf{i}}_{t_{av};\beta_-,\beta_+} \mathscr{C}^{\mathbf{X},-}_{N^\beta X}(\mathfrak{g}_{S+t_s},y)| \, dS \tag{5.61}$$

$$= \int_0^{t_f} \widetilde{\sum}_{y \in \mathbb{T}_N} \mathbf{E}^{\mu_{0,\mathbb{Z}}} \int_{S+t_s,N} \mathbf{E}^{\mathrm{path}}_{\eta_0} |\mathscr{C}^{\mathbf{T},+,t_{N,\varepsilon},\mathbf{i}}_{t_{av};\beta_-,\beta_+} \mathscr{C}^{\mathbf{X},-}_{N^\beta X}(\mathfrak{g}_0,y)| \, dS$$

$$= \int_0^{t_f} \widetilde{\sum}_{y \in \mathbb{T}_N} \mathbf{E}^{\mu_{0,\mathbb{Z}}} \tau_y \int_{S+t_s,N} \mathbf{E}^{\mathrm{path}}_{\eta_0} |\mathscr{C}^{\mathbf{T},+,t_{N,\varepsilon},\mathbf{i}}_{t_{av};\beta_-,\beta_+} \mathscr{C}^{\mathbf{X},-}_{N^\beta X}(\mathfrak{g}_0,0)| \, dS \tag{5.62}$$

$$= \mathbf{E}^{\mu_{0,\mathbb{Z}}} \left( \int_0^{t_f} \widetilde{\sum}_{y \in \mathbb{T}_N} \tau_y \int_{S+t_s,N} \mathbf{E}^{\mathrm{path}}_{\eta_0} |\mathscr{C}^{\mathbf{T},+,t_{N,\varepsilon},\mathbf{i}}_{t_{av};\beta_-,\beta_+} \mathscr{C}^{\mathbf{X},-}_{N^\beta X}(\mathfrak{g}_0,0)| \, dS \right) \tag{5.63}$$

$$= \mathbf{E}^{\mu_{0,\mathbb{Z}}} \left( \left( \int_0^{t_f} \widetilde{\sum}_{y \in \mathbb{T}_N} \tau_y \int_{S+t_s,N} \, dS \right) \mathbf{E}^{\mathrm{path}}_{\eta_0} |\mathscr{C}^{\mathbf{T},+,t_{N,\varepsilon},\mathbf{i}}_{t_{av};\beta_-,\beta_+} \mathscr{C}^{\mathbf{X},-}_{N^\beta X}(\mathfrak{g}_0,0)| \right). \tag{5.64}$$

The average in parentheses in (5.64) is $t_f \bar{\int}_{t_s,t_f,\mathbb{T}_N}$ given prior to Lemma 3.8, so (5.59) and (5.64) finish the proof. $\qquad \square$

*Proof of Lemma 5.22.* We first construct a two-species process, one evolving on $\Omega_{\mathbb{Z}}$ and another on $\Omega_{\mathfrak{I}_{t_{av}}}$.

- Consider the original process $\eta_\bullet$ with the infinitesimal generator $\mathfrak{S}^{N,!!}$ evolving on $\Omega_{\mathbb{Z}}$. This is the first species, and we specify the symmetric component of the particle random walks is given pathwise by Poisson clocks attached to *bonds*. In particular, a ringing of a bond indicates swapping spins at the two points along that bond. All the other dynamics are pieces of the totally asymmetric part of $\mathfrak{S}^{N,!!}$ and use any Poisson clocks. This construction was used in the proof of Lemma 3.15 for example.

- Consider the $\mathfrak{I}_{\mathfrak{t}_{av}}$-local process $\eta_{\mathfrak{I}_{\mathfrak{t}_{av}};\bullet,\bullet}$ evolving on $\Omega_{\mathfrak{I}_{\mathfrak{t}_{av}}}$. This is the second species. We will similarly specify the symmetric component of the particle random walks is given pathwise by Poisson clocks attached to bonds *in* $\mathfrak{I}_{\mathfrak{t}_{av}} \subseteq \mathbb{Z}$; note in particular inf $\mathfrak{I}_{\mathfrak{t}_{av}}$ and sup $\mathfrak{I}_{\mathfrak{t}_{av}}$ are connected with the associated speed $\alpha_1 \in \mathbb{R}_{>0}$. We now couple symmetric dynamics of the two species by coupling any spin-swap dynamics along any shared bonds. In particular, we will use the same Poisson clock for bonds that appear in both of $\mathbb{Z}$ and $\mathfrak{I}_{\mathfrak{t}_{av}}$. Moreover, all Poisson clocks corresponding to totally asymmetric jumps across bonds in $\mathfrak{I}_{\mathfrak{t}_{av}}$ are coupled to Poisson clocks from the first species above that are associated to the same totally asymmetric jump across the same bond in $\mathbb{Z}$ if such a bond is in $\mathbb{Z}$. In other words, we employ the basic coupling for all totally asymmetric jumps that appear in both dynamics. All other Poisson clocks, such as symmetric spin-swap clocks attached to bonds not in the first $\mathbb{Z}$-process and totally asymmetric jump clocks also not in the first $\mathbb{Z}$-process, are constructed in any arbitrary fashion.

Assume these two species above exhibit the same spins in $\mathfrak{I}_{\mathfrak{t}_{av}} \subseteq \mathbb{Z}$ initially. We will show the probability of seeing a discrepancy between these two species inside the neighborhood of radius $N^{\beta_X} \log^{50} N$ around $0 \in \mathbb{Z}$ before the time $2\mathfrak{t}_{av} + 2\mathfrak{t}_{N,\varepsilon}$ is bounded above by $N^{-C}$ up to $C$-dependent factors for any $C \in \mathbb{R}_{>0}$. Before we actually show the previous claim, observe that the claim would provide both of (5.25) and (5.26). Indeed, the path-space functionals of interest in both (5.25) and (5.26) only depend on the trajectories of spins inside the support of $\mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g})$, which is contained in the aforementioned neighborhood around $0 \in \mathbb{Z}$, until time $2\mathfrak{t}_{av} + 2\mathfrak{t}_{N,\varepsilon}$. We clarify the paths taken by these spins in the aforementioned support certainly depend upon the global configuration of spins. Regardless, if the two species agree on these spins until time $2\mathfrak{t}_{av} + 2\mathfrak{t}_{N,\varepsilon}$, then the two functionals agree on this path of the two-species process. If not, we bound the expectations via $|\mathfrak{g}| \lesssim 1$ and the $N^{-C}$-probability of such an event.

We now prove the claim about discrepancies between the two species appearing in the neighborhood of radius $N^{\beta_X} \log^{50} N$ around $0 \in \mathbb{Z}$ before time $2\mathfrak{t}_{av} + 2\mathfrak{t}_{N,\varepsilon}$. We will argue via large deviations for random walks as in the proof of Lemma 3.15. First, for convenience we will let $\mathfrak{r}_N = N^{\beta_X} \log^{50} N$ denote the radius of the aforementioned neighborhood. Observe $|\mathfrak{I}_{\mathfrak{t}_{av}}| \gg \mathfrak{r}_N$.

- Initially, there are no discrepancies in $\mathfrak{I}_{\mathfrak{t}_{av}} \subseteq \mathbb{Z}$ between the two species by construction. Note there certainly might be some outside this block. By construction of the coupling between the two species, discrepancies cannot be made by shared Poisson clocks, but only moved or killed. In particular, any discrepancy in $\mathfrak{I}_{\mathfrak{t}_{av}}$ must be created by ringing some Poisson clock in at least one of the two species that is located within $\mathfrak{m}$ of the boundary of $\mathfrak{I}_{\mathfrak{t}_{av}}$ as each particle performs range-$\mathfrak{m}$ random walks. First, we conclude any discrepancy that appears inside the neighborhood of radius $\mathfrak{r}_N$ around $0 \in \mathbb{Z}$ at any time must originally be a discrepancy that is within $\mathfrak{m}$ of the boundary of $\mathfrak{I}_{\mathfrak{t}_{av}}$ that propagates into $\mathfrak{r}_N$ of $0 \in \mathbb{Z}$. Second, the number of discrepancies in $\mathfrak{I}_{\mathfrak{t}_{av}}$ that can be made until time $2\mathfrak{t}_{av} + 2\mathfrak{t}_{N,\varepsilon}$ is bounded by the number of times Poisson clocks at sites within $\mathfrak{m}$ of the boundary of $\mathfrak{I}_{\mathfrak{t}_{av}}$ can ring before the time $2\mathfrak{t}_{av} + 2\mathfrak{t}_{N,\varepsilon} \lesssim N^{20}$. As these Poisson clocks have speed $N^2$, this is at most $N^{C_1}$ outside an event of exponentially low probability $\mathscr{O}(\mathbf{Exp}(-N^{\kappa_{C_1}}))$ for any $C_1 \in \mathbb{R}_{>0}$ arbitrarily large but universal and for $\kappa_{C_1} \in \mathbb{R}_{>0}$.

- Take any discrepancy in the previous bullet point, let it start within $\mathfrak{m}$ of the boundary of $\mathfrak{I}_{\mathfrak{t}_{av}}$, and wait until it is within $\frac{1}{5}|\mathfrak{I}_{\mathfrak{t}_{av}}|$ of a radius-$\mathfrak{r}_N$ neighborhood of $0 \in \mathbb{Z}$. As in the proof of Lemma 3.15, the dynamics of such discrepancy is a random walk with a

free unsuppressed symmetric component of speed $N^2$ and the maximal jump-length $\mathfrak{m} \lesssim 1$, with an environment-dependent drift of order $N^{3/2}$ of maximal jump-length $\mathfrak{m} \lesssim 1$, and a random killing mechanism, *at least* as long as such discrepancy stays within $\frac{1}{5}|\mathfrak{I}_{\mathfrak{t}_{av}}|$ of the radius-$\mathfrak{r}_N$ neighborhood of $0 \in \mathbb{Z}$. We emphasize that this free and unsuppressed nature of the symmetric component of the discrepancy random walks comes from coupling bonds and symmetric spin-swaps as opposed to the basic coupling. As $|\mathfrak{I}_{\mathfrak{t}_{av}}| \gg \mathfrak{r}_N$, the distance between the edge of the $\frac{1}{5}|\mathfrak{I}_{\mathfrak{t}_{av}}|$-neighborhood and the $\mathfrak{r}_N$-neighborhood is order $|\mathfrak{I}_{\mathfrak{t}_{av}}|$, so the probability the discrepancy arrives within $\mathfrak{r}_N$ of $0 \in \mathbb{Z}$ is controlled by the probability that the aforementioned random walk exits a ball of radius order $|\mathfrak{I}_{\mathfrak{t}_{av}}|$, both these statements restricted to the time-interval $[0, 2\mathfrak{t}_{av} + 2\mathfrak{t}_{N,\varepsilon}]$. Indeed, we wait for such discrepancy to land within $\mathfrak{r}_N$ of $0 \in \mathbb{Z}$ through the previous random walk dynamics, or we wait for it to get beyond $\frac{1}{5}|\mathfrak{I}_{\mathfrak{t}_{av}}|$ of this neighborhood. If it does the latter, we wait for it to get within $\frac{1}{5}|\mathfrak{I}_{\mathfrak{t}_{av}}|$ of this neighborhood again and repeat. But if it does the former, then it must have traveled distance at least order $|\mathfrak{I}_{\mathfrak{t}_{av}}|$ via the aforementioned random walk dynamics.

- Continuing with the above bullet point, the probability that this discrepancy lands in the radius-$\mathfrak{r}_N$ neighborhood of $0 \in \mathbb{Z}$ is then order $\kappa_C N^{-2C}$ for any $C \in \mathbb{R}_{>0}$ courtesy of standard random walk/large deviations estimates because of the additional log-factors in $|\mathfrak{I}_{\mathfrak{t}_{av}}|$; see Remark 5.20. We take union bound over all $N^{C_1}$ possible discrepancies and choose $C \geqslant C_1$ arbitrarily large but universal to complete the proof of the high-probability discrepancy claim.

This completes the proof.                                                    □

*Proof of Lemma 5.23.* The inner expectation on the LHS of the proposed estimate is a functional on $\Omega_{\mathfrak{I}_{\mathfrak{t}_{av}}}$ because dependence on the $\mathfrak{I}_{\mathfrak{t}_{av}}$-stochastic process is only through its initial configuration; we have taken expectation over the associated $\mathfrak{I}_{\mathfrak{t}_{av}}$-local path-space measure conditioning on its initial configuration. Thus, we may apply the entropy inequality from Lemma 3.8 with $\varphi$ equal to the inner expectation on the LHS of the proposed estimate with support $\mathfrak{I}_{\mathfrak{t}_{av}}$ and with the choice of constant $\kappa = N^{\beta_-}$. We reemphasize that we choose $\mathfrak{I}_N = \mathbb{T}_N$ in Lemma 3.8 which is okay because $\mathbb{T}_N$ satisfies the constraints we assumed for the block $\mathfrak{I}_N$ stated in Proposition 3.6. In our application of Lemma 3.8, we also choose $\mathfrak{t}_0 = \mathfrak{t}_\mathfrak{s}$ here and $T = \mathfrak{t}_\mathfrak{f}$. This all gives

$$\mathfrak{t}_\mathfrak{f} \mathbf{E}^{\mu_{0,\mathbb{Z}}} \bar{\mathfrak{f}}_{\mathfrak{t}_\mathfrak{s},\mathfrak{t}_\mathfrak{f},\mathbb{T}_N}^{\mathfrak{I}_{\mathfrak{t}_{av}}} \mathbf{E}_{\Pi_{\mathfrak{I}_{\mathfrak{t}_{av}}} \eta_0}^{\mathrm{loc}} |\mathscr{C}_{\mathfrak{t}_{av};\beta_-,\beta_+}^{\mathbf{T},+,\mathfrak{t}_{N,\varepsilon},\mathfrak{i}} \mathscr{C}_{N^{\beta}X}^{\mathbf{X},-}(\mathfrak{g}_{\mathfrak{t}_{av}};0,0)|$$
$$\lesssim N^{-\beta_- - \frac{3}{4} + \varepsilon} |\mathbb{T}_N|^{-1} |\mathfrak{I}_{\mathfrak{t}_{av}}|^3 + \mathfrak{t}_\mathfrak{f} \sup_{\varrho \in \mathbb{R}} \Psi_\varrho \qquad (5.65)$$

where $\varepsilon \in \mathbb{R}_{>0}$ is arbitrarily small but universal and for convenience we defined the following equilibrium log-exp term:

$$\Psi_\varrho \overset{\bullet}{=} N^{-\beta_-} \log \mathbf{E}^{\mu_{\varrho,\mathfrak{I}_{\mathfrak{t}_{av}}}^{\mathrm{can}}} \mathbf{Exp}\left(N^{\beta_-} \mathbf{E}_{\Pi_{\mathfrak{I}_{\mathfrak{t}_{av}}} \eta_0}^{\mathrm{loc}} |\mathscr{C}_{\mathfrak{t}_{av};\beta_-,\beta_+}^{\mathbf{T},+,\mathfrak{t}_{N,\varepsilon},\mathfrak{i}} \mathscr{C}_{N^{\beta}X}^{\mathbf{X},-}(\mathfrak{g}_{\mathfrak{t}_{av}};0,0)|\right). \qquad (5.66)$$

The first term on the RHS of (5.65) is bounded by power-counting. We do this later and focus on the more interesting term $\Psi_\varrho$. First observe that by definition of $\mathscr{C}^{\mathbf{T},+}$, in the exponential defining $\Psi_\varrho$, the $\mathbf{E}^{\mathrm{loc}}$-term is at most $N^{-\beta_-}$ times universal factors; see Definition 5.21 for the definition of the $\mathscr{C}^{\mathbf{T},+}$-term. Thus the quantity inside the exponential is uniformly bounded. Employing the inequalities $\mathbf{Exp}(|x|) \leqslant 1 + \kappa|x|$ for

$\kappa \in \mathbb{R}_{\geqslant 0}$ and $x \in \mathbb{R}$ both uniformly bounded and $\log(1 + |x|) \leqslant |x|$, we get

$$\Psi_{\varrho} \leqslant N^{-\beta_-} \log \mathbf{E}^{\mu_{\varrho,\mathfrak{I}}^{\mathrm{can}} \mathfrak{t}_{\mathrm{av}}} \left( 1 + \kappa \cdot N^{\beta_-} \mathbf{E}^{\mathrm{loc}}_{\Pi_{\mathfrak{I}_{\mathfrak{t}_{\mathrm{av}}}} \eta_0} |\mathscr{C}^{\mathbf{T},+,\mathfrak{t}_{N,\varepsilon},\mathfrak{i}}_{\mathfrak{t}_{\mathrm{av}};\beta_-,\beta_+} \mathscr{C}^{\mathbf{X},-}_{N^{\beta_X}}(\mathfrak{g}_{\mathfrak{t}_{\mathrm{av}};0,0})| \right) \quad (5.67)$$

$$= N^{-\beta_-} \log \left( 1 + \kappa \cdot N^{\beta_-} \mathbf{E}^{\mu_{\varrho,\mathfrak{I}}^{\mathrm{can}} \mathfrak{t}_{\mathrm{av}}} \mathbf{E}^{\mathrm{loc}}_{\Pi_{\mathfrak{I}_{\mathfrak{t}_{\mathrm{av}}}} \eta_0} |\mathscr{C}^{\mathbf{T},+,\mathfrak{t}_{N,\varepsilon},\mathfrak{i}}_{\mathfrak{t}_{\mathrm{av}};\beta_-,\beta_+} \mathscr{C}^{\mathbf{X},-}_{N^{\beta_X}}(\mathfrak{g}_{\mathfrak{t}_{\mathrm{av}};0,0})| \right) \quad (5.68)$$

$$\lesssim \mathbf{E}^{\mu_{\varrho,\mathfrak{I}}^{\mathrm{can}} \mathfrak{t}_{\mathrm{av}}} \mathbf{E}^{\mathrm{loc}}_{\Pi_{\mathfrak{I}_{\mathfrak{t}_{\mathrm{av}}}} \eta_0} |\mathscr{C}^{\mathbf{T},+,\mathfrak{t}_{N,\varepsilon},\mathfrak{i}}_{\mathfrak{t}_{\mathrm{av}};\beta_-,\beta_+} \mathscr{C}^{\mathbf{X},-}_{N^{\beta_X}}(\mathfrak{g}_{\mathfrak{t}_{\mathrm{av}};0,0})|. \quad (5.69)$$

To estimate the double expectation in (5.69), let us first ease notation and let $\widetilde{\mathbf{E}}$ denote the iterated expectation in (5.69). We now recall definition of the $\mathscr{C}^{\mathbf{T},+}$-term in (5.69) within Definition 5.21 to be the time-average $\mathscr{A}^{\mathbf{T},+}$ multiplied by two indicator functions. For the sake of getting an upper bound, we drop the upper-bound indicator function and apply the Cauchy-Schwarz inequality to get the following. Below, we emphasize that because the canonical ensemble initial measure we take in definition of $\widetilde{\mathbf{E}}$ is an invariant measure for the $\mathfrak{I}_{\mathfrak{t}_{\mathrm{av}}}$-local process we are taking expectation with respect to in $\widetilde{\mathbf{E}}$, this tells us that time-shifts do not change the law of functionals of the $\mathfrak{I}_{\mathfrak{t}_{\mathrm{av}}}$-local process as long as the $\mathfrak{I}_{\mathfrak{t}_{\mathrm{av}}}$-local process starts with any canonical ensemble initial measure. We deduce this "invariance-under-time-shift" upon writing the $\widetilde{\mathbf{E}}$-expectation of any time-shifted functional as an expectation with respect to the path-space measure induced by the $\mathfrak{I}_{\mathfrak{t}_{\mathrm{av}}}$-local process with an initial measure given by the law of the particle system after evolving for a time equal to the time-shift. We can do this with the Markov property for the $\mathfrak{I}_{\mathfrak{t}_{\mathrm{av}}}$-local process. But canonical ensembles are invariant measures, so the "time-shifted" initial measure is the same canonical ensemble. This lets us forget the time-shift $\mathfrak{t}_{N,\varepsilon} \in \mathbb{R}_{\geqslant 0}$ in $\widetilde{\mathbf{E}}$-expectations below until (5.72). More precisely, we note the following.

- The laws of $\mathbf{1}[\mathscr{G}^{\beta_+}_{\mathfrak{t}_{\mathrm{av}};\geqslant;S,0}]$ and $\mathscr{A}^{\mathbf{T},+}_{\mathfrak{t}_{\mathrm{av}}} \mathscr{C}^{\mathbf{X},-}_{N^{\beta_X}}(\mathfrak{g}_{\mathfrak{t}_{\mathrm{av}};S,0})$ are $S$-independent under $\widetilde{\mathbf{E}}$. See Definition 5.21 for their definitions.

Thus at the level of $\widetilde{\mathbf{E}}$, the time-shift $\mathfrak{t}_{N,\varepsilon}$ is irrelevant. For any $\mathfrak{i} \in \{1, 2\}$, we get the following with some justification given after:

$$\widetilde{\mathbf{E}}|\mathscr{C}^{\mathbf{T},+,\mathfrak{t}_{N,\varepsilon},\mathfrak{i}}_{\mathfrak{t}_{\mathrm{av}};\beta_-,\beta_+} \mathscr{C}^{\mathbf{X},-}_{N^{\beta_X}}(\mathfrak{g}_{\mathfrak{t}_{\mathrm{av}};0,0})|$$

$$\lesssim \left( \widetilde{\mathbf{E}}|\mathscr{A}^{\mathbf{T},+}_{\mathfrak{t}_{\mathrm{av}}} \mathscr{C}^{\mathbf{X},-}_{N^{\beta_X}}(\mathfrak{g}_{\mathfrak{t}_{\mathrm{av}};0,0})|^2 \right)^{1/2} \left( \widetilde{\mathbf{E}}\mathbf{1}[\mathscr{G}^{\beta_+}_{\mathfrak{t}_{\mathrm{av}};\geqslant;0,0}] \right)^{1/2} \quad (5.70)$$

$$\lesssim \left( \widetilde{\mathbf{E}}|\mathscr{A}^{\mathbf{T},+}_{\mathfrak{t}_{\mathrm{av}}} \mathscr{C}^{\mathbf{X},-}_{N^{\beta_X}}(\mathfrak{g}_{\mathfrak{t}_{\mathrm{av}};0,0})|^2 \right)^{1/2} \left( N^{2\beta_+} \widetilde{\mathbf{E}}\sup_{0 \leqslant \mathfrak{t} \leqslant \mathfrak{t}_{\mathrm{av}}} \mathfrak{t}^2 \mathfrak{t}_{\mathrm{av}}^{-2} |\mathscr{A}^{\mathbf{T},+}_{\mathfrak{t}} \mathscr{C}^{\mathbf{X},-}_{N^{\beta_X}}(\mathfrak{g}_{\mathfrak{t}_{\mathrm{av}};0,0})|^2 \right)^{1/2} \quad (5.71)$$

$$\lesssim N^{\beta_+} \widetilde{\mathbf{E}}\sup_{0 \leqslant \mathfrak{t} \leqslant \mathfrak{t}_{\mathrm{av}}} \mathfrak{t}^2 \mathfrak{t}_{\mathrm{av}}^{-2} |\mathscr{A}^{\mathbf{T},+}_{\mathfrak{t}} \mathscr{C}^{\mathbf{X},-}_{N^{\beta_X}}(\mathfrak{g}_{\mathfrak{t}_{\mathrm{av}};0,0})|^2. \quad (5.72)$$

To get (5.72) from the previous line, we control a process by its running supremum (take $\mathfrak{t} = \mathfrak{t}_{\mathrm{av}}$ on the RHS of this next bound):

$$|\mathscr{A}^{\mathbf{T},+}_{\mathfrak{t}_{\mathrm{av}}} \mathscr{C}^{\mathbf{X},-}_{N^{\beta_X}}(\mathfrak{g}_{\mathfrak{t}_{\mathrm{av}};0,0})|^2 \leqslant \sup_{0 \leqslant \mathfrak{t} \leqslant \mathfrak{t}_{\mathrm{av}}} \mathfrak{t}^2 \mathfrak{t}_{\mathrm{av}}^{-2} |\mathscr{A}^{\mathbf{T},+}_{\mathfrak{t}} \mathscr{C}^{\mathbf{X},-}_{N^{\beta_X}}(\mathfrak{g}_{\mathfrak{t}_{\mathrm{av}};0,0})|^2 \quad (5.73)$$

The bound preceding (5.72) follows from the Chebyshev inequality for second moments applied to the probability of $\mathscr{G}^{\beta_+}_{\mathfrak{t}_{\mathrm{av}};\geqslant;0,0}$. To estimate the last supremum we would like to apply the Kipnis-Varadhan time-average estimate in Lemma 3.10. Indeed, this would be a major advantage of reducing our estimates to equilibrium calculations. However, the functional that we are time-averaging via $\mathscr{A}^{\mathbf{T},+}$, which we emphasize has no indicator functions attached to it anymore in (5.72), is not a spatial-average of pseudo-gradients

with disjoint support. The obstruction to this being true is the cutoff present in $\mathscr{C}^{\mathbf{X},-}$. We will remove this cutoff. First, observe the following courtesy of definition of $\mathscr{C}^{\mathbf{X},-}$ from Definition 5.21 and the uniform boundedness of $\mathfrak{g}$:

$$|\mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{t_{av};\mathfrak{r},0}) - \mathscr{A}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{t_{av};\mathfrak{r},0})| \lesssim \mathbf{1}[(\mathscr{E}_{t_{av};\mathfrak{r}}^{\mathbf{X},-})^C]. \tag{5.74}$$

We will use (5.74) to replace the $\mathscr{C}^{\mathbf{X},-}$ term (5.72) by the spatial-average $\mathscr{A}^{\mathbf{X},-}$ without cutoff. In particular, we will employ the following estimate with universal implied constant uniformly over $0 \leqslant t \leqslant t_{av}$. We will use notation we define afterwards:

$$\begin{aligned}
&t^2 t_{av}^{-2} |\mathscr{A}_t^{\mathbf{T},+} \mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{t_{av};0,0})|^2 \\
&= t^2 t_{av}^{-2} |\mathscr{A}_t^{\mathbf{T},+} \mathscr{A}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{t_{av};0,0}) + \mathscr{A}_t^{\mathbf{T},+} \mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{t_{av};0,0}) - \mathscr{A}_t^{\mathbf{T},+} \mathscr{A}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{t_{av};0,0})|^2 \\
&\lesssim t^2 t_{av}^{-2} |\mathscr{A}_t^{\mathbf{T},+} \mathscr{A}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{t_{av};0,0})|^2 + t^2 t_{av}^{-2} |\mathscr{A}_t^{\mathbf{T},+} \mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{t_{av};0,0}) - \mathscr{A}_t^{\mathbf{T},+} \mathscr{A}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{t_{av};0,0})|^2 \\
&\lesssim t^2 t_{av}^{-2} |\mathscr{A}_t^{\mathbf{T},+} \mathscr{A}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{t_{av};0,0})|^2 + t^2 t_{av}^{-2} |\mathscr{A}_t^{\mathbf{T},+}(\mathbf{1}[(\mathscr{E}_{t_{av};0}^{\mathbf{X},-})^C])|^2.
\end{aligned}$$

The second bound follows by the elementary inequality $|a+b|^2 \lesssim |a|^2 + |b|^2$ for all $a, b \in \mathbb{R}$. The last bound follows by (5.74), where in the last bound the term $\mathscr{A}_t^{\mathbf{T},+}(\mathbf{1}[(\mathscr{E}_{t_{av};\mathfrak{r}}^{\mathbf{X},-})^C])$ is the time-average over $[0, t]$ of the random indicator function $\mathbf{1}[(\mathscr{E}_{t_{av};\mathfrak{r}}^{\mathbf{X},-})^C]$. To bound the second term in the last line, we use the Cauchy-Schwarz inequality with respect to the time *integral* $t\mathscr{A}_t^{\mathbf{T},+}$:

$$\begin{aligned}
t^2 t_{av}^{-2} |\mathscr{A}_t^{\mathbf{T},+}(\mathbf{1}[(\mathscr{E}_{t_{av};0}^{\mathbf{X},-})^C])|^2 &= t_{av}^{-2} |\left(\int_0^t \mathbf{1}[(\mathscr{E}_{t_{av};\mathfrak{r}}^{\mathbf{X},-})^C] \, d\mathfrak{r}\right)|^2 \\
&\leqslant t_{av}^{-2} t \int_0^t \mathbf{1}[(\mathscr{E}_{t_{av};\mathfrak{r}}^{\mathbf{X},-})^C] \, d\mathfrak{r} \leqslant t_{av}^{-1} \int_0^{t_{av}} \mathbf{1}[(\mathscr{E}_{t_{av};\mathfrak{r}}^{\mathbf{X},-})^C] \, d\mathfrak{r}.
\end{aligned}$$

The first identity follows by cancelling the $t^{-2}$ factor in the square of the time-average on the far LHS with the $t^2$-factor therein. The second estimate follows from the Cauchy-Schwarz inequality with respect to the time-integral. The third estimate follows by extending the integration domain $[0, t] \to [0, t_{av}]$ as $0 \leqslant t \leqslant t_{av}$ and thus noting $t_{av}^{-1} t \leqslant 1$. The last two displays give

$$\begin{aligned}
&\widetilde{\mathbf{E}} \sup_{0 \leqslant t \leqslant t_{av}} t^2 t_{av}^{-2} |\mathscr{A}_t^{\mathbf{T},+} \mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{t_{av};0,0})|^2 \\
&\lesssim \widetilde{\mathbf{E}} \sup_{0 \leqslant t \leqslant t_{av}} t^2 t_{av}^{-2} |\mathscr{A}_t^{\mathbf{T},+} \mathscr{A}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{t_{av};0,0})|^2 + t_{av}^{-1} \int_0^{t_{av}} \widetilde{\mathbf{E}} \mathbf{1}[(\mathscr{E}_{t_{av};\mathfrak{r}}^{\mathbf{X},-})^C] d\mathfrak{r}. \quad (5.75)
\end{aligned}$$

We observe the final expectation in the second term of (5.75) is the probability of an event in $\Omega_{\mathfrak{I}_{t_{av}}}$ under a canonical ensemble. Indeed, the event therein depends on only statistics at time $\mathfrak{r}$ under the $\mathfrak{I}_{t_{av}}$-local exclusion process, and the canonical ensemble initial measure is invariant for this exclusion process. In particular, it is the probability under canonical ensemble measure that the spatial-average of $N^{\beta_X}$-many pseudo-gradients with disjoint support exceeds its natural "CLT" bound. Combining (5.75) with this observation and the canonical ensemble estimate of Lemma 3.13 with $\varphi_j = \tau_{-7j\mathfrak{m}}\mathfrak{g}_{0,0}$ and $J = N^{\beta_X}$ gives the bound

$$\begin{aligned}
&\widetilde{\mathbf{E}} \sup_{0 \leqslant t \leqslant t_{av}} t^2 t_{av}^{-2} |\mathscr{A}_t^{\mathbf{T},+} \mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{t_{av};0,0})|^2 \\
&\lesssim \widetilde{\mathbf{E}} \sup_{0 \leqslant t \leqslant t_{av}} t^2 t_{av}^{-2} |\mathscr{A}_t^{\mathbf{T},+} \mathscr{A}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{t_{av};0,0})|^2 + N^{-100}. \quad (5.76)
\end{aligned}$$

The error term $N^{-100}$ in (5.76) can be replaced by something exponentially small if desired. Anyway, the first term on the RHS of (5.76) is now a space-time average of pseudo-gradients with disjoint support free of any indicator functions or cutoffs. Thus may use Lemma 3.10 with $\varphi$ the $\mathscr{A}^{\mathbf{X},-}$ spatial average and Lemma 3.11 with $J = N^{\beta_X}$ and $\varphi_j = \tau_{-7jm}\mathfrak{g}_{0,0}$ with support $\mathfrak{I}_j$; note that in the following estimate, we use the uniform boundedness of $|\mathfrak{I}_j|$-cardinalities and of the shifted pseudo-gradients $\varphi_j$:

$$
\begin{aligned}
\widetilde{\mathbf{E}} \sup_{0 \leqslant t \leqslant t_{av}} t^2 t_{av}^{-2} |\mathscr{A}_t^{\mathbf{T},+}\mathscr{A}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{t_{av};0,0})|^2 &\lesssim t_{av}^{-1} \|\mathscr{A}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{t_{av};0,0})\|^2_{\dot{\mathbf{H}}^{-1}_{\varrho,\mathfrak{I}_{t_{av}}}} \\
&\lesssim t_{av}^{-1} N^{-\beta_X} \sup_{j=1,\ldots,J} \sup_{\varrho_j \in \mathbb{R}} \|\varphi_j\|^2_{\dot{\mathbf{H}}^{-1}_{\varrho_j,\mathfrak{I}_j}} \\
&\lesssim N^{-2-\beta_X} t_{av}^{-1}.
\end{aligned}
\tag{5.77}
$$

We combine (5.65), (5.69), (5.72), (5.76), and (5.77) to get the following in which we use $t_f N^{\beta_+} N^{-100} \lesssim N^{-97}$:

$$
\begin{aligned}
t_f \mathbf{E}^{\mu_{0,\mathbb{Z}}} \bar{f}_{t_s,t_f,\mathbb{T}_N}^{-\mathfrak{I}_{t_{av}}} \mathbf{E}^{loc}_{\Pi_{\mathfrak{I}_{t_{av}}} \eta_0} &|\mathscr{C}_{t_{av};\beta_-,\beta_+}^{\mathbf{T},+,t_{N,\varepsilon},i} \mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{t_{av};0,0})| \\
&\lesssim N^{-\beta_- -\frac{3}{4}+\varepsilon}|\mathbb{T}_N|^{-1}|\mathfrak{I}_{t_{av}}|^3 + t_f N^{-2-\beta_X+\beta_+} t_{av}^{-1} + N^{-97}.
\end{aligned}
\tag{5.78}
$$

The LHS of this last estimate is the LHS of the proposed estimate in the current lemma but without the $N$-dependent prefactor in the proposed estimate. Putting this $N$-dependent prefactor back in the above and power-counting completes the proof upon recalling that $|\mathfrak{I}_{t_{av}}| \lesssim_\varepsilon N^{1+\varepsilon} t_{av}^{1/2} + N^{3/2+\varepsilon} t_{av} + N^{\beta_X+\varepsilon}$ for any $\varepsilon \in \mathbb{R}_{>0}$ and recalling that $|\mathbb{T}_N| \gtrsim N^{5/4+\varepsilon_{cpt}}$. We appeal to Definition 5.19 and the beginning of Sect. 4 for these length-scale estimates, respectively. $\quad\square$

*Proof of Lemma 5.24.* The first estimate follows from the proof for Lemma 5.17 but choosing $t_N \overset{\bullet}{=} T - N^{-\beta_X+\beta_- -\varepsilon}$ therein. To prove the second estimate, we follow the proof of Lemma 5.18 upon replacing $\mathscr{C}^{\mathbf{X},-}(\mathfrak{g}) \to \widetilde{\mathfrak{g}}^{\mathfrak{l}}$ and using the same argument. We note that in following the argument of Lemma 5.18, we start with the $\|\|_{t_f;\mathbb{T}_N}$-bound obtained in the first estimate of the current lemma with $\varphi$ equal to the $\mathscr{C}^{\mathbf{T},+,t_{N,\varepsilon},i}$-term, that certainly satisfies the constraints for $\varphi$ in the first estimate of the current lemma, instead of the estimate of Lemma 5.17. Precisely, after doing this and then taking expectation we get, for any $\varepsilon \in \mathbb{R}_{>0}$,

$$
\begin{aligned}
\mathbf{E}\|\bar{\mathfrak{P}}_{T,x}^N (N^{\beta_X}|\mathscr{C}_{t_{av};\beta_-,\beta_+}^{\mathbf{T},+,t_{N,\varepsilon},i}(\widetilde{\mathfrak{g}}_{S+t_s,y}^{\mathfrak{l}})|)\|_{t_f;\mathbb{T}_N} \\
\lesssim N^{\frac{3}{2}\beta_X+\frac{1}{4}+\varepsilon_{cpt}+\frac{1}{2}\varepsilon-\frac{1}{2}\beta_-} \int_0^{t_f} \widetilde{\sum_{y \in \mathbb{T}_N}} \mathbf{E}|\mathscr{C}_{t_{av};\beta_-,\beta_+}^{\mathbf{T},+,t_{N,\varepsilon},i}(\widetilde{\mathfrak{g}}_{S+t_s,y}^{\mathfrak{l}})| \, dS + N^{-\varepsilon}.
\end{aligned}
\tag{5.79}
$$

At this point, we use the same procedure of decomposing expectations via path-space measure expectations and initial configuration expectations along with the space-time-shift invariance of the law of the path-space measure. $\quad\square$

*Proof of Lemma 5.26.* This follows by the same coupling construction as was used to establish Lemma 5.22. Indeed, note the support of $\widetilde{\mathfrak{g}}^{\mathfrak{l}}$ is also bounded above in size by $N^{\beta_X}$ times universal factors, so in the proof of Lemma 5.26 we are also interested here in looking at when discrepancies in the coupling from the proof of Lemma 5.22 get within $N^{\beta_X} \log^{50} N$ of $x \in \mathbb{Z}$. $\quad\square$

*Proof of Lemma 5.27.* Observe the inner expectation on the LHS of the proposed estimate is a functional on the configuration space $\Omega_{\mathfrak{I}_{\mathfrak{t}_{av}}}$ since dependence on the path-space measure induced by the $\mathfrak{I}_{\mathfrak{t}_{av}}$-local process goes away given we take expectation over this path-space measure *conditioning* on the initial particle configuration. This is actually the same observation we made at the beginning of the proof of Lemma 5.23. The outer expectation on the LHS of the proposed estimate is thus an expectation of a $\mathfrak{I}_{\mathfrak{t}_{av}}$-statistic with respect to the $\bar{\mathfrak{f}}$ probability density which we recall is a space-time averaged probability density. We may therefore apply the local equilibrium estimate/entropy inequality from Lemma 3.8 with $\kappa = N^{\beta_-}$, with $\mathfrak{I}_N = \mathbb{T}_N$, and with $\varphi$ the inner expectation on the LHS of the proposed estimate. We emphasize that choosing $\mathfrak{I}_N = \mathbb{T}_N$ in this application of Lemma 3.8 is okay as $\mathbb{T}_N$ satisfies assumptions of Proposition 3.6. We also emphasize our choice of $\varphi$ has support $\mathfrak{I}_{\mathfrak{t}_{av}}$. We finally note that in applying Lemma 3.8 we take $\mathfrak{t}_0$ therein to be $\mathfrak{t}_{\mathfrak{s}}$ here and $T$ therein to be $\mathfrak{t}_{\mathfrak{f}}$ here. We ultimately get

$$\mathfrak{t}_{\mathfrak{f}} \mathbf{E}^{\mu_{0,\mathbb{Z}}} \bar{\mathfrak{f}}^{\mathfrak{I}_{\mathfrak{t}_{av}}}_{\mathfrak{t}_{\mathfrak{s}},\mathfrak{t}_{\mathfrak{f}},\mathbb{T}_N} \mathbf{E}^{\mathrm{loc}}_{\Pi_{\mathfrak{I}_{\mathfrak{t}_{av}}} \eta_0} |\mathscr{C}^{\mathbf{T},+,\mathfrak{t}_{N,\varepsilon},\mathfrak{i}}_{\mathfrak{t}_{av};\beta_-,\beta_+}(\widetilde{\mathfrak{g}}^{\mathfrak{l}}_{\mathfrak{t}_{av};0,0})| \;\lesssim\; N^{-\beta_- - \frac{3}{4}+\varepsilon} |\mathbb{T}_N|^{-1} |\mathfrak{I}_{\mathfrak{t}_{av}}|^3 \;+\; \mathfrak{t}_{\mathfrak{f}} \sup_{\varrho \in \mathbb{R}} \widetilde{\Psi}_\varrho \tag{5.80}$$

where again $\varepsilon \in \mathbb{R}_{>0}$ is arbitrarily small but universal and we defined the equilibrium log-exp term

$$\widetilde{\Psi}_\varrho \;\overset{\bullet}{=}\; N^{-\beta_-} \log \mathbf{E}^{\mu^{\mathrm{can}}_{\varrho,\mathfrak{I}_{\mathfrak{t}_{av}}}} \mathbf{Exp}\left( N^{\beta_-} \mathbf{E}^{\mathrm{loc}}_{\Pi_{\mathfrak{I}_{\mathfrak{t}_{av}}} \eta_0} |\mathscr{C}^{\mathbf{T},+,\mathfrak{t}_{N,\varepsilon},\mathfrak{i}}_{\mathfrak{t}_{av};\beta_-,\beta_+}(\widetilde{\mathfrak{g}}^{\mathfrak{l}}_{\mathfrak{t}_{av};0,0})| \right). \tag{5.81}$$

The first term on the RHS of (5.80) is handled again by power-counting which we do at the end of the proof. The analysis of the second term on the RHS of (5.80) is addressed in similar fashion as the $\Psi_\varrho$-terms from the proof of Proposition 5.6 except now it is actually a little easier as we illustrate. We recall from Definition 5.25 that the expectation inside the exponential defining $\widetilde{\Psi}$-quantities is bounded by $N^{-\beta_-}$ times universal factors by definition. In particular, the term inside of the exponential defining $\widetilde{\Psi}_\varrho$ is uniformly bounded. This is identical to what we observed for $\Psi_\varrho$-terms in the proof for Proposition 5.6. Thus, we use the exponential and logarithm bounds given prior to (5.69) in the proof of Proposition 5.6 to get, with universal implied constant,

$$\widetilde{\Psi}_\varrho \;\lesssim\; \mathbf{E}^{\mu^{\mathrm{can}}_{\varrho,\mathfrak{I}_{\mathfrak{t}_{av}}}} \mathbf{E}^{\mathrm{loc}}_{\Pi_{\mathfrak{I}_{\mathfrak{t}_{av}}} \eta_0} |\mathscr{C}^{\mathbf{T},+,\mathfrak{t}_{N,\varepsilon},\mathfrak{i}}_{\mathfrak{t}_{av};\beta_-,\beta_+}(\widetilde{\mathfrak{g}}^{\mathfrak{l}}_{\mathfrak{t}_{av};0,0})|. \tag{5.82}$$

At this we point we will perform equilibrium calculations like with the proof for Lemma 5.23 after (5.69). First, we follow the proof of (5.72) via the Cauchy-Schwarz inequality to estimate the last expectation from the RHS of (5.82) in terms of a second moment of the time-average $\mathscr{A}^{\mathbf{T},+}$ without any cutoff and the probability of the lower-bound-cutoff-event defining $\mathscr{F}^{\beta_+}$ from Definition 5.25. Indeed observe that $\mathscr{C}^{\mathbf{T},+,\mathfrak{t}_{N,\varepsilon},\mathfrak{i}}$ time-averaging is, in absolute value, controlled by $\mathscr{A}^{\mathbf{T},+}$ time-averaging times the indicator function of this last $\mathscr{F}^{\beta_+}$-event. To estimate the probability of this lower-bound-cutoff-event $\mathscr{F}^{\beta_+}$ we employ the Chebyshev inequality for a second moment and observe that at the level of expectations the time-shift $\mathfrak{t}_{N,\varepsilon} \in \mathbb{R}_{\geq 0}$ is irrelevant. This is because the canonical ensemble initial measure in our iterated expectation within the RHS of (5.82) is invariant for the $\mathfrak{I}_{\mathfrak{t}_{av}}$-local process, so the time-shift $\mathfrak{t}_{N,\varepsilon}$ does not affect the law

of $\mathscr{A}^{\mathbf{T},+}$-terms. If $\widetilde{\mathbf{E}}$ is the double expectation in (5.82), we obtain the following analog of the estimate (5.72) from the proof of Lemma 5.23:

$$\widetilde{\mathbf{E}}|\mathscr{C}^{\mathbf{T},+,t_{N,\varepsilon},\mathfrak{i}}_{t_{\mathrm{av}};\beta_-,\beta_+}(\widetilde{\mathfrak{g}}^{\mathfrak{l}}_{t_{\mathrm{av}};0,0})| \lesssim \left(\widetilde{\mathbf{E}}|\mathscr{A}^{\mathbf{T},+}_{t_{\mathrm{av}}}(\widetilde{\mathfrak{g}}^{\mathfrak{l}}_{t_{\mathrm{av}};0,0})|^2\right)^{\frac{1}{2}}\left(\widetilde{\mathbf{E}}\mathbf{1}[\mathscr{F}^{\beta_+}_{t_{\mathrm{av}};\geqslant;0,0}]\right)^{\frac{1}{2}} \tag{5.83}$$

$$\lesssim \left(\widetilde{\mathbf{E}}|\mathscr{A}^{\mathbf{T},+}_{t_{\mathrm{av}}}(\widetilde{\mathfrak{g}}^{\mathfrak{l}}_{t_{\mathrm{av}};0,0})|^2\right)^{\frac{1}{2}}\left(N^{2\beta_+}\widetilde{\mathbf{E}}\sup_{0\leqslant t\leqslant t_{\mathrm{av}}}t^2 t_{\mathrm{av}}^{-2}|\mathscr{A}^{\mathbf{T},+}_t(\widetilde{\mathfrak{g}}^{\mathfrak{l}}_{t_{\mathrm{av}};0,0})|^2\right)^{\frac{1}{2}} \tag{5.84}$$

$$\lesssim N^{\beta_+}\widetilde{\mathbf{E}}\sup_{0\leqslant t\leqslant t_{\mathrm{av}}}t^2 t_{\mathrm{av}}^{-2}|\mathscr{A}^{\mathbf{T},+}_t(\widetilde{\mathfrak{g}}^{\mathfrak{l}}_{t_{\mathrm{av}};0,0})|^2. \tag{5.85}$$

To get (5.85) from the the preceding line, as with the proof of (5.72) we used a bound for a process by a running supremum:

$$|\mathscr{A}^{\mathbf{T},+}_{t_{\mathrm{av}}}(\widetilde{\mathfrak{g}}^{\mathfrak{l}}_{t_{\mathrm{av}};0,0})|^2 \lesssim \sup_{0\leqslant t\leqslant t_{\mathrm{av}}}t^2 t_{\mathrm{av}}^{-2}|\mathscr{A}^{\mathbf{T},+}_t(\widetilde{\mathfrak{g}}^{\mathfrak{l}}_{t_{\mathrm{av}};0,0})|^2. \tag{5.86}$$

We now estimate the remaining $\widetilde{\mathbf{E}}$-expectation in (5.85). First we apply the Kipnis-Varadhan inequality of Lemma 3.10 with $\varphi$ therein equal to $\widetilde{\mathfrak{g}}^{\mathfrak{l}}$ here and $\mathfrak{I}$ therein equal to $\mathfrak{I}_{t_{\mathrm{av}}}$ here. This controls the last $\widetilde{\mathbf{E}}$-expectation in terms of the Sobolev norm of $\widetilde{\mathfrak{g}}^{\mathfrak{l}}$. To control this Sobolev norm, we employ Lemma 3.12 to bound it by the Sobolev norm of the pseudo-gradient factor inside $\widetilde{\mathfrak{g}}^{\mathfrak{l}}$. This last Sobolev norm of the pseudo-gradient factor in $\widetilde{\mathfrak{g}}^{\mathfrak{l}}$, which is uniformly bounded and has uniformly bounded support, is then controlled by its uniform bound and a speed factor $N^{-2}$ by Lemma 3.11. We ultimately get

$$\widetilde{\mathbf{E}}\sup_{0\leqslant t\leqslant t_{\mathrm{av}}}t^2 t_{\mathrm{av}}^{-2}|\mathscr{A}^{\mathbf{T},+}_t(\widetilde{\mathfrak{g}}^{\mathfrak{l}}_{t_{\mathrm{av}};0,0})|^2 \lesssim t_{\mathrm{av}}^{-1}\|\widetilde{\mathfrak{g}}^{\mathfrak{l}}\|^2_{\dot{\mathbf{H}}^{-1}_{\varrho,\mathfrak{I}_{t_{\mathrm{av}}}}} \lesssim N^{-2}t_{\mathrm{av}}^{-1}. \tag{5.87}$$

We combine (5.82), (5.85), and (5.87) to bound $\widetilde{\Psi}$-terms. Combining the resulting $\widetilde{\Psi}$-bound with (5.80) then gives

$$t_{\mathfrak{f}}\mathbf{E}^{\mu_{0,\mathbb{Z}}}\bar{\mathfrak{f}}^{\mathfrak{I}_{t_{\mathrm{av}}}}_{t_{\mathfrak{s}},t_{\mathfrak{f}},\mathbb{T}_N}\mathbf{E}^{\mathrm{loc}}_{\Pi_{\mathfrak{I}_{t_{\mathrm{av}}}}\eta_0}|\mathscr{C}^{\mathbf{T},+,t_{N,\varepsilon},\mathfrak{i}}_{t_{\mathrm{av}};\beta_-,\beta_+}(\widetilde{\mathfrak{g}}^{\mathfrak{l}}_{t_{\mathrm{av}};0,0})| \lesssim N^{-\beta_--\frac{3}{4}+\varepsilon}|\mathbb{T}_N|^{-1}|\mathfrak{I}_{t_{\mathrm{av}}}|^3 + t_{\mathfrak{f}}N^{-2+\beta_+}t_{\mathrm{av}}^{-1}. \tag{5.88}$$

We recall $|\mathbb{T}_N| \gtrsim N^{5/4+\varepsilon_{\mathrm{cpt}}}$ from Sect. 4 and the length-scale $|\mathfrak{I}_{t_{\mathrm{av}}}| \lesssim_\varepsilon N^{1+\varepsilon}t_{\mathrm{av}}^{1/2} + N^{3/2+\varepsilon}t_{\mathrm{av}} + N^{\beta_X+\varepsilon}$; power-counting gives the following bound for the RHS of (5.88) in which we recall $\varepsilon_X = \max_i \varepsilon_{X,i}$ from the statement of Lemma 5.27 for example:

$$N^{\frac{3}{2}\beta_X+\frac{1}{4}+\varepsilon_{\mathrm{cpt}}+\frac{1}{2}\varepsilon-\frac{1}{2}\beta_-}t_{\mathfrak{f}}\mathbf{E}^{\mu_{0,\mathbb{Z}}}\bar{\mathfrak{f}}^{\mathfrak{I}_{t_{\mathrm{av}}}}_{t_{\mathfrak{s}},t_{\mathfrak{f}},\mathbb{T}_N}\mathbf{E}^{\mathrm{loc}}_{\Pi_{\mathfrak{I}_{t_{\mathrm{av}}}}\eta_0}|\mathscr{C}^{\mathbf{T},+,t_{N,\varepsilon},\mathfrak{i}}_{t_{\mathrm{av}};\beta_-,\beta_+}(\widetilde{\mathfrak{g}}^{\mathfrak{l}}_{t_{\mathrm{av}};0,0})|$$
$$\lesssim N^{-\frac{1}{8}-\frac{3}{4}\beta_-+10\varepsilon_X} + N^{-\frac{1}{2}+10\varepsilon_X} + t_{\mathfrak{f}}N^{-\beta_{\mathrm{univ}}}. \tag{5.89}$$

This last estimate is the proposed estimate of the current lemma as it holds uniformly in $\mathfrak{l} \in [\![1, N^{\beta_X}]\!]$, so we are done. $\qquad\square$

*Proof of Lemma 5.28.* Take any $T \in [0, 1]$ and $x \in \mathbb{T}_N$. To get the first bound, we employ the Cauchy-Schwarz inequality with respect to the space-time "integral" against the heat operator and the fact that the heat kernel is a probability measure on $\mathbb{T}_N$:

$$\bar{\mathfrak{P}}_{T,x}^N(N^{\frac{1}{2}}|\varphi_{S,y}|) = \int_0^T \mathfrak{s}_{S,T}^{-\frac{1}{4}} \mathfrak{s}_{S,T}^{\frac{1}{4}} \sum_{y \in \mathbb{T}_N} \bar{\mathfrak{P}}_{S,T,x,y}^N \cdot N^{\frac{1}{2}}|\varphi_{S,y}| \, dS$$

$$\lesssim \left( \int_0^T \mathfrak{s}_{S,T}^{-\frac{1}{2}} dS \right)^{\frac{1}{2}} \left( \int_0^T \mathfrak{s}_{S,T}^{\frac{1}{2}} \sum_{y \in \mathbb{T}_N} \bar{\mathfrak{P}}_{S,T,x,y}^N \cdot N|\varphi_{S,y}|^2 \, dS \right)^{\frac{1}{2}} \tag{5.90}$$

$$\lesssim \left( \int_0^T \mathfrak{s}_{S,T}^{\frac{1}{2}} \sum_{y \in \mathbb{T}_N} \bar{\mathfrak{P}}_{S,T,x,y}^N \cdot N|\varphi_{S,y}|^2 \, dS \right)^{\frac{1}{2}}. \tag{5.91}$$

The final bound (5.91) follows by integrating the first integral in the preceding bound. We now use the on-diagonal heat kernel estimate implied by (A.1). As $|\mathbb{T}_N| \lesssim N^{5/4 + \varepsilon_{\mathrm{cpt}}}$, we get what follows; the last line follows by extending the integration-domain:

$$\left( \int_0^T \mathfrak{s}_{S,T}^{\frac{1}{2}} \sum_{y \in \mathbb{T}_N} \bar{\mathfrak{P}}_{S,T,x,y}^N \cdot N|\varphi_{S,y}|^2 \, dS \right)^{\frac{1}{2}} \lesssim \left( N^{-1} \int_0^T \mathfrak{s}_{S,T}^{\frac{1}{2}} \mathfrak{s}_{S,T}^{-\frac{1}{2}} \sum_{y \in \mathbb{T}_N} N|\varphi_{S,y}|^2 \, dS \right)^{\frac{1}{2}} \tag{5.92}$$

$$\lesssim \left( |\mathbb{T}_N| \int_0^T \widetilde{\sum}_{y \in \mathbb{T}_N} |\varphi_{S,y}|^2 \, dS \right)^{\frac{1}{2}} \tag{5.93}$$

$$\lesssim \left( N^{\frac{5}{4} + \varepsilon_{\mathrm{cpt}}} \int_0^T \widetilde{\sum}_{y \in \mathbb{T}_N} |\varphi_{S,y}|^2 \, dS \right)^{\frac{1}{2}} \tag{5.94}$$

$$\lesssim \left( N^{\frac{5}{4} + \varepsilon_{\mathrm{cpt}}} \int_0^1 \widetilde{\sum}_{y \in \mathbb{T}_N} |\varphi_{S,y}|^2 \, dS \right)^{\frac{1}{2}}. \tag{5.95}$$

We combine (5.91) and (5.95) to get a bound for the LHS of (5.91) that is uniform on $[0, 1] \times \mathbb{T}_N$ and thus we deduce the first estimate in the current lemma. To get the second estimate in the current lemma, we employ the first estimate therein to get

$$\left( \mathbf{E} \| \bar{\mathfrak{P}}_{T,x}^N(N^{\frac{1}{2}} | \mathscr{A}_{\mathrm{tav}}^{\mathbf{T},+} \mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{S,y})|) \|_{1;\mathbb{T}_N} \right)^2$$

$$\lesssim \left( \mathbf{E} \left( N^{\frac{5}{4} + \varepsilon_{\mathrm{cpt}}} \int_0^1 \widetilde{\sum}_{y \in \mathbb{T}_N} | \mathscr{A}_{\mathrm{tav}}^{\mathbf{T},+} \mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{S,y})|^2 \, dS \right)^{1/2} \right)^2 \tag{5.96}$$

$$\lesssim N^{\frac{5}{4} + \varepsilon_{\mathrm{cpt}}} \int_0^1 \widetilde{\sum}_{y \in \mathbb{T}_N} \mathbf{E} | \mathscr{A}_{\mathrm{tav}}^{\mathbf{T},+} \mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{S,y})|^2 \, dS. \tag{5.97}$$

The last estimate (5.97) follows by the Cauchy-Schwarz inequality with respect to the expectation $\mathbf{E}$. At this point, we follow the proof of Lemma 5.18 via decomposing the last expectation in (5.97) in terms of an expectation with respect to a path-space measure with an initial configuration that is then sampled according to the space-time averaged law of the particle system. In particular, the replacement of $\mathscr{C}^{\mathbf{T},+} \to \mathscr{A}^{\mathbf{T},+}$ and the presence of the square does not change the property of the term within the expectation in

(5.97) being a functional of the original exclusion process on $\Omega_{\mathbb{Z}}$ with generator $\mathfrak{S}^{N,!!}$ which is all we needed to decompose this last expectation in (5.97) in terms of path-space expectation and expectation over initial configuration.                                                                 $\square$

*Proof of Lemma 5.29.* We follow the proof of Lemma 5.23. Observe the inner expectation on the LHS of the proposed estimate is a $\mathfrak{I}_{\mathfrak{t}_{\mathrm{av}}}$-local functional that is integrated against the space-time averaged law $\bar{\mathfrak{f}}$ of the particle system. We may therefore again apply local equilibrium to the $\mathfrak{I}_{\mathfrak{t}_{\mathrm{av}}}$-local inner expectation on the LHS of the proposed estimate of the current lemma via Lemma 3.8 and perform equilibrium calculations. To be precise, we employ Lemma 3.8 with $\mathfrak{I}_N = \mathbb{T}_N$, which we recall is okay because $\mathbb{T}_N$ satisfies constraints of Proposition 3.6, with $\kappa = N^{\beta_X - \varepsilon_{X,1}}$, and with $\varphi$ equal to the inner expectation within the LHS of the proposed estimate of the current lemma with support $\mathfrak{I}_{\mathfrak{t}_{\mathrm{av}}}$. In this application of Lemma 3.8, we will take $\mathfrak{t}_{\mathfrak{s}}$ therein to be 0 and we take $\mathfrak{t}_{\mathfrak{f}}$ to be 1. This gives the following where $\varepsilon \in \mathbb{R}_{>0}$ is arbitrarily small but universal:

$$\mathbf{E}^{\mu_{0,\mathbb{Z}}} \bar{\mathfrak{f}}_{1,\mathbb{T}_N}^{\mathfrak{I}_{\mathfrak{t}_{\mathrm{av}}}} \mathbf{E}_{\Pi_{\mathfrak{I}_{\mathfrak{t}_{\mathrm{av}}}}\eta_0}^{\mathrm{loc}} |\mathscr{A}_{\mathfrak{t}_{\mathrm{av}}}^{\mathbf{T},+} \mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{\mathfrak{t}_{\mathrm{av}};0,0})|^2 \lesssim N^{-\frac{3}{4} - \beta_X + \varepsilon_{X,1} + \varepsilon} |\mathbb{T}_N|^{-1} |\mathfrak{I}_{\mathfrak{t}_{\mathrm{av}}}|^3 + \sup_{\varrho \in \mathbb{R}} \Upsilon_\varrho,$$
(5.98)

where we have introduced the following equilibrium log-exp quantities:

$$\Upsilon_\varrho \overset{\bullet}{=} N^{-\beta_X + \varepsilon_{X,1}} \log \mathbf{E}^{\mu_{\varrho,\mathfrak{I}_{\mathfrak{t}_{\mathrm{av}}}}^{\mathrm{can}}} \mathbf{Exp}\left(N^{\beta_X - \varepsilon_{X,1}} \mathbf{E}_{\Pi_{\mathfrak{I}_{\mathfrak{t}_{\mathrm{av}}}}\eta_0}^{\mathrm{loc}} |\mathscr{A}_{\mathfrak{t}_{\mathrm{av}}}^{\mathbf{T},+} \mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{\mathfrak{t}_{\mathrm{av}};0,0})|^2\right).$$
(5.99)

For the $\Upsilon_\varrho$-terms, recall the $\mathscr{C}^{\mathbf{X},-}$-term in the exponential is bounded above by $N^{-\beta_X/2 + \varepsilon_{X,1}/2}$ deterministically by construction in Definition 5.21. Its time-average $\mathscr{A}^{\mathbf{T},+}\mathscr{C}^{\mathbf{X},-}$, therefore, satisfies the same deterministic a priori estimate, and thus the quantity in the exponential defining $\Upsilon_\varrho$ is uniformly bounded deterministically. Thus, the exponential and logarithm inequalities given prior to (5.69) in the proof of Proposition 5.6 yield the following in similar fashion as how we obtained (5.69):

$$\Upsilon_\varrho \lesssim \mathbf{E}^{\mu_{\varrho,\mathfrak{I}_{\mathfrak{t}_{\mathrm{av}}}}^{\mathrm{can}}} \mathbf{E}_{\Pi_{\mathfrak{I}_{\mathfrak{t}_{\mathrm{av}}}}\eta_0}^{\mathrm{loc}} |\mathscr{A}_{\mathfrak{t}_{\mathrm{av}}}^{\mathbf{T},+} \mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{\mathfrak{t}_{\mathrm{av}};0,0})|^2.$$
(5.100)

To bound the iterated expectation in (5.100), we again employ equilibrium calculations. We have actually done this in the proof of Lemma 5.23. Precisely, we follow (5.73), (5.76), and (5.77) in the proof for Lemma 5.23 to get the following in which $N^{-100}$ is ultimately negligible for the conclusions of the current lemma:

$$\mathbf{E}^{\mu_{\varrho,\mathfrak{I}_{\mathfrak{t}_{\mathrm{av}}}}^{\mathrm{can}}} \mathbf{E}_{\Pi_{\mathfrak{I}_{\mathfrak{t}_{\mathrm{av}}}}\eta_0}^{\mathrm{loc}} |\mathscr{A}_{\mathfrak{t}_{\mathrm{av}}}^{\mathbf{T},+} \mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{\mathfrak{t}_{\mathrm{av}};0,0})|^2 \lesssim N^{-2 - \beta_X} \mathfrak{t}_{\mathrm{av}}^{-1} + N^{-100}.$$
(5.101)

We now combine (5.98), (5.100), and (5.101) to deduce the desired estimate in Lemma 5.29 and thus complete its proof.                                                                                                    $\square$

*Proof of Lemma 5.30.* The first bound follows by the proof of the first estimate in Lemma 5.28 by replacing $N^{1/2}$ with $N^{\beta_X}$ and accounting for the resulting elementary changes in power-counting. This may be screened without a calculation if we pretend $\beta_X = \frac{1}{2}$ in which case we match the exponent $2\beta_X + \frac{1}{4} + \varepsilon_{\mathrm{cpt}}$ in the first estimate of Lemma 5.30 with the exponent $\frac{5}{4} + \varepsilon_{\mathrm{cpt}}$ in the first estimate of Lemma 5.28. To prove the second estimate in Lemma 5.30, we similarly follow the proof of the second estimate in Lemma 5.28 after the following replacement and the same minor and elementary

adjustments in power-counting which we alluded to in the proof of the first estimate of the current Lemma 5.30:

$$\left(N^{\frac{1}{2}}, \mathscr{A}_{\mathfrak{t}_{av}}^{\mathbf{T},+} \mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{S,y})\right) \;\to\; \left(N^{\beta_X}, \mathscr{A}_{\mathfrak{t}_{av}}^{\mathbf{T},+}(\widetilde{\mathfrak{g}}_{S,y}^{\mathfrak{l}})\right). \qquad (5.102)$$

This completes the proof. □

*Proof of Lemma 5.31.* Looking at the beginning of the proof of Lemma 5.27, we first observe the inner expectation on the LHS of the proposed estimate is a $\mathfrak{I}_{\mathfrak{t}_{av}}$-local statistic, so that the iterated expectation within the LHS of the proposed estimate is that of a $\mathfrak{I}_{\mathfrak{t}_{av}}$-local statistic against the space-time averaged Radon-Nikodym derivative $\bar{\mathfrak{f}}$. We may thus employ the local equilibrium estimate of Lemma 3.8 to reduce the problem of estimating the iterated expectation within the LHS of the proposed estimate in terms of equilibrium calculations. In particular, we employ Lemma 3.8 with $\mathfrak{I}_N = \mathbb{T}_N$, with $\varphi$ the inner path-space expectation on the LHS of the proposed estimate which has support $\mathfrak{I}_{\mathfrak{t}_{av}}$, and with $\kappa = 1$. We clarify that this choice of $\kappa = 1$ means we will not depend on an a priori large-deviations estimate for the inner expectation that we are taking to be our local statistic $\varphi$ in this application of Lemma 3.8. We also emphasize this choice $\mathfrak{I}_N = \mathbb{T}_N$ in using Lemma 3.8 is okay because $\mathbb{T}_N$ satisfies constraints on the block $\mathfrak{I}_N$ in Proposition 3.6. Lastly, we clarify that we apply Lemma 3.8 with $\mathfrak{t}_{\mathfrak{s}}$ therein to be 0 and $\mathfrak{t}_{\mathfrak{f}}$ therein to be 1:

$$\mathbf{E}^{\mu_{0,\mathbb{Z}}}\bar{\mathfrak{f}}_{1,\mathbb{T}_N}^{\mathfrak{I}_{\mathfrak{t}_{av}}}\mathbf{E}_{\Pi_{\mathfrak{I}_{\mathfrak{t}_{av}}}\eta_0}^{\mathrm{loc}}|\mathscr{A}_{\mathfrak{t}_{av}}^{\mathbf{T},+}(\widetilde{\mathfrak{g}}_{\mathfrak{t}_{av};0,0}^{\mathfrak{l}})|^2 \;\lesssim_\varepsilon\; N^{-\frac{3}{4}+\varepsilon}|\mathbb{T}_N|^{-1}|\mathfrak{I}_{\mathfrak{t}_{av}}|^3 \;+\; \sup_{\varrho\in\mathbb{R}}\widetilde{\Upsilon}_\varrho. \quad (5.103)$$

Above $\varepsilon \in \mathbb{R}_{>0}$ is arbitrarily small but universal. We have also introduced the following equilibrium quantities:

$$\widetilde{\Upsilon}_\varrho \;\overset{\bullet}{=}\; \log \mathbf{E}^{\mu_{\varrho,\mathfrak{I}_{\mathfrak{t}_{av}}}^{\mathrm{can}}} \mathbf{Exp}\left(\mathbf{E}_{\Pi_{\mathfrak{I}_{\mathfrak{t}_{av}}}\eta_0}^{\mathrm{loc}}|\mathscr{A}_{\mathfrak{t}_{av}}^{\mathbf{T},+}(\widetilde{\mathfrak{g}}_{\mathfrak{t}_{av};0,0}^{\mathfrak{l}})|^2\right). \qquad (5.104)$$

We focus on local equilibrium calculations relevant to estimating $\widetilde{\Upsilon}$-terms. Recall in Proposition 2.7 that the $\widetilde{\mathfrak{g}}^{\mathfrak{l}}$-terms are terms that admit pseudo-gradient factors, so they are uniformly bounded. The exponential and logarithm inequalities we wrote prior to the estimate (5.69) from the proof of Lemma 5.23 then give

$$\widetilde{\Upsilon}_\varrho \;\lesssim\; \mathbf{E}^{\mu_{\varrho,\mathfrak{I}_{\mathfrak{t}_{av}}}^{\mathrm{can}}} \mathbf{E}_{\Pi_{\mathfrak{I}_{\mathfrak{t}_{av}}}\eta_0}^{\mathrm{loc}}|\mathscr{A}_{\mathfrak{t}_{av}}^{\mathbf{T},+}(\widetilde{\mathfrak{g}}_{\mathfrak{t}_{av};0,0}^{\mathfrak{l}})|^2 \;\lesssim\; N^{-2}\mathfrak{t}_{av}^{-1}. \qquad (5.105)$$

The final estimate (5.105) follows from the bound that is immediately to its left via (5.86) and (5.87) from the proof of Lemma 5.27. We now combine (5.103) and (5.105) to deduce the desired estimate of Lemma 5.31. □

*Proof of Lemma 5.32.* We may follow the proof of the second estimate in the statement of Lemma 5.28 using the first estimate therein but with the function $\varphi_{S,y} = \mathscr{A}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{S,y}) - \mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{S,y})$. Technically this gives the desired identity but with $\bar{\mathfrak{f}}$ on the RHS of the proposed identity, not its projection onto $\mathfrak{I}$. But the inner expectation on the RHS depends only on spins in $\mathfrak{I}$. □

*Proof of Lemma 5.33.* Let us observe the expectation on the LHS of the proposed bound is expectation is of a local statistic given by the squared difference against the space-time averaged probability density $\bar{\mathfrak{f}}$ of the law of the particle system. We may thus employ local equilibrium Lemma 3.8 with $\mathfrak{I}_N = \mathbb{T}_N$, which again is okay as $\mathbb{T}_N$ satisfies

constraints of Proposition 3.6, with $\varphi$ equal to the square within the LHS of the estimate in Lemma 5.33 with support contained in $\mathfrak{I}_\varphi = \mathfrak{I} = [\![-100\mathrm{m}N^{\beta_X}, 100\mathrm{m}N^{\beta_X}]\!]$ from Lemma 5.32, and with the choice of constant $\kappa = N^{\beta_X - 3\varepsilon_{X,1}}$. Again, in this application of Lemma 3.8 we also take $\mathfrak{t}_{\mathfrak{s}}$ therein to be 0 and $\mathfrak{t}_{\mathfrak{f}}$ therein to be 1. For $\varepsilon \in \mathbb{R}_{>0}$ arbitrarily small but universal, this gives

$$\mathbf{E}^{\mu_{0,\mathbb{Z}}}\bar{\mathfrak{f}}^{\mathfrak{I}}_{1,\mathbb{T}_N} |\mathscr{A}^{\mathbf{X},-}_{N^{\beta_X}}(\mathfrak{g}_{0,0}) - \mathscr{C}^{\mathbf{X},-}_{N^{\beta_X}}(\mathfrak{g}_{0,0})|^2 \lesssim N^{-\frac{3}{4}+\varepsilon} N^{-\beta_X + 3\varepsilon_{X,1}} |\mathbb{T}_N|^{-1} |\mathfrak{I}|^3 + \sup_{\varrho \in \mathbb{R}} \Gamma_\varrho.$$
(5.106)

We have introduced the following equilibrium log-exp expectations in (5.106) above:

$$\Gamma_\varrho \overset{\bullet}{=} N^{-\beta_X + 3\varepsilon_{X,1}} \log \mathbf{E}^{\mu^{\mathrm{can}}_{\varrho,\mathfrak{I}}} \mathbf{Exp} \left( N^{\beta_X - 3\varepsilon_{X,1}} |\mathscr{A}^{\mathbf{X},-}_{N^{\beta_X}}(\mathfrak{g}_{0,0}) - \mathscr{C}^{\mathbf{X},-}_{N^{\beta_X}}(\mathfrak{g}_{0,0})|^2 \right). \quad (5.107)$$

The first term on the RHS of (5.106) is again addressed by elementary power-counting in $N \in \mathbb{Z}_{>0}$ that we come back to at the end of this proof. We focus on $\Gamma$-terms on the RHS of (5.106) first. To this end, by definitions in Definition 5.1, we get

$$\mathbf{Exp} \left( N^{\beta_X - 3\varepsilon_{X,1}} |\mathscr{A}^{\mathbf{X},-}_{N^{\beta_X}}(\mathfrak{g}_{0,0}) - \mathscr{C}^{\mathbf{X},-}_{N^{\beta_X}}(\mathfrak{g}_{0,0})|^2 \right)$$
$$= \mathbf{1}[\mathscr{E}^{\mathbf{X},-}_{0,0}] + \mathbf{Exp} \left( N^{\beta_X - 3\varepsilon_{X,1}} |\mathscr{A}^{\mathbf{X},-}_{N^{\beta_X}}(\mathfrak{g}_{0,0})|^2 \right) \mathbf{1}[(\mathscr{E}^{\mathbf{X},-}_{0,0})^C]. \quad (5.108)$$

The superscript $C$ on the RHS of (5.108) denotes taking the complement of the $\mathscr{E}^{\mathbf{X},-}$ event we defined in Definition 5.1. Indeed, on this $\mathscr{E}^{\mathbf{X},-}$-event, the $\mathscr{A}^{\mathbf{X},-}$ and $\mathscr{C}^{\mathbf{X},-}$ operators/terms agree by definition, and thus the exponential of the difference is equal to 1, while outside this $\mathscr{E}^{\mathbf{X},-}$-event we note that $\mathscr{C}^{\mathbf{X},-}$ vanishes. Thus, taking expectations of (5.108) we get the consequence

$$\mathbf{E}^{\mu^{\mathrm{can}}_{\varrho,\mathfrak{I}}} \mathbf{Exp} \left( N^{\beta_X - 3\varepsilon_{X,1}} |\mathscr{A}^{\mathbf{X},-}_{N^{\beta_X}}(\mathfrak{g}_{0,0}) - \mathscr{C}^{\mathbf{X},-}_{N^{\beta_X}}(\mathfrak{g}_{0,0})|^2 \right)$$
$$= \mathbf{E}^{\mu^{\mathrm{can}}_{\varrho,\mathfrak{I}}} \mathbf{1}[\mathscr{E}^{\mathbf{X},-}_{0,0}] + \mathbf{E}^{\mu^{\mathrm{can}}_{\varrho,\mathfrak{I}}} \mathbf{Exp} \left( N^{\beta_X - 3\varepsilon_{X,1}} |\mathscr{A}^{\mathbf{X},-}_{N^{\beta_X}}(\mathfrak{g}_{0,0})|^2 \right) \mathbf{1}[(\mathscr{E}^{\mathbf{X},-}_{0,0})^C]$$
$$= 1 - \mathbf{E}^{\mu^{\mathrm{can}}_{\varrho,\mathfrak{I}}} \mathbf{1}[(\mathscr{E}^{\mathbf{X},-}_{0,0})^C] + \mathbf{E}^{\mu^{\mathrm{can}}_{\varrho,\mathfrak{I}}} \mathbf{Exp} \left( N^{\beta_X - 3\varepsilon_{X,1}} |\mathscr{A}^{\mathbf{X},-}_{N^{\beta_X}}(\mathfrak{g}_{0,0})|^2 \right) \mathbf{1}[(\mathscr{E}^{\mathbf{X},-}_{0,0})^C]$$
$$\leqslant 1 + \mathbf{E}^{\mu^{\mathrm{can}}_{\varrho,\mathfrak{I}}} \mathbf{Exp} \left( N^{\beta_X - 3\varepsilon_{X,1}} |\mathscr{A}^{\mathbf{X},-}_{N^{\beta_X}}(\mathfrak{g}_{0,0})|^2 \right) \mathbf{1}[(\mathscr{E}^{\mathbf{X},-}_{0,0})^C]. \quad (5.109)$$

The second line in the last calculation follows by recalling $(\mathscr{E}^{\mathbf{X},-})^C$ is the complement of the $\mathscr{E}^{\mathbf{X},-}$-event. The estimate (5.109) is almost trivial and follows via non-negativity of probabilities. We proceed by using the inequality $\log(1 + |x|) \leqslant |x|$, with $|x|$ equal to the expectation in (5.109), with (5.109) and definition of $\Gamma_\varrho$; we get the following with universal implied constant:

$$\Gamma_\varrho \lesssim N^{-\beta_X + 3\varepsilon_{X,1}} \mathbf{E}^{\mu^{\mathrm{can}}_{\varrho,\mathfrak{I}}} \mathbf{Exp} \left( N^{\beta_X - 3\varepsilon_{X,1}} |\mathscr{A}^{\mathbf{X},-}_{N^{\beta_X}}(\mathfrak{g}_{0,0})|^2 \right) \mathbf{1}[(\mathscr{E}^{\mathbf{X},-}_{0,0})^C]. \quad (5.110)$$

To estimate this exponential expectation, we recall by definition that $\mathscr{A}^{\mathbf{X},-}(\mathfrak{g}_{0,0})$ is the average of spatial-shifts of the pseudo-gradient $\mathfrak{g}_{0,0}$ which all have mutually disjoint support; see Proposition 2.7 for the disjoint property of supports and Definition 5.1 for $\mathscr{A}^{\mathbf{X},-}$. Thus, we use Corollary 3.14 with $J = N^{\beta_X}$, with $\varphi_j = \tau_{-7j\mathrm{m}}\mathfrak{g}_{0,0}$, with $\varepsilon_1 = 3\varepsilon_{X,1}$, and with $\varepsilon_2 = \frac{1}{2}\varepsilon_{X,1}$ to get

$$\mathbf{E}^{\mu^{\mathrm{can}}_{\varrho,\mathfrak{I}}} \mathbf{Exp} \left( N^{\beta_X - 3\varepsilon_{X,1}} |\mathscr{A}^{\mathbf{X},-}_{N^{\beta_X}}(\mathfrak{g}_{0,0})|^2 \right) \mathbf{1}[(\mathscr{E}^{\mathbf{X},-}_{0,0})^C] \lesssim_C N^{-C}. \quad (5.111)$$

The constant $C \in \mathbb{R}_{>0}$ in (5.111) is arbitrarily large but universal. We combine this with (5.110) to control $\Gamma_\varrho$ by an arbitrarily large but universal *negative* power of $N \in \mathbb{Z}_{>0}$. We combine (5.106), (5.110), and (5.111) to get

$$N^{\frac{5}{4}+\varepsilon_{\mathrm{cpt}}} \mathbf{E}^{\mu_{0,\mathbb{Z}}} \bar{\mathfrak{f}}_{1,\mathbb{T}_N}^{\mathfrak{I}} |\mathscr{A}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{0,0}) - \mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{0,0})|^2$$
$$\lesssim_C N^{\frac{5}{4}+\varepsilon_{\mathrm{cpt}}} N^{-\frac{3}{4}+\varepsilon} N^{-\beta_X+3\varepsilon_{X,1}} |\mathbb{T}_N|^{-1} |\mathfrak{I}|^3 + N^{\frac{5}{4}+\varepsilon_{\mathrm{cpt}}} N^{-C}. \tag{5.112}$$

We recall $\mathfrak{I} \subseteq \mathbb{Z}$ within the statement of Lemma 5.33. Choosing $C \in \mathbb{R}_{>0}$ sufficiently large and elementary power-counting for the first term on the RHS of (5.112), recalling $|\mathbb{T}_N| \gtrsim N^{5/4+\varepsilon_{\mathrm{cpt}}}$ from the beginning of Sect. 4, completes the proof. $\qquad\square$

## 6. Stochastic Time-Regularity Estimates

As pointed out in the would-be-proof of Pseudo-Proposition 2.13, to perform time-replacement in the statement of Pseudo-Proposition 2.13 we need time-regularity of the heat kernel $\bar{\mathfrak{P}}^N$ and of the compactified microscopic Cole-Hopf transform $\bar{\mathfrak{Z}}^N$. The former is given in Lemma A.1 and is deterministic. The latter is the current focus. Let us first introduce some notation.

**Definition 6.1.** Provided any $\mathfrak{t} \in \mathbb{R}$, let us define $\nabla_\mathfrak{t}^{\mathbf{T}} \varphi_\bullet = \varphi_{\bullet+\mathfrak{t}} - \varphi_\bullet$ for functions $\varphi : \mathbb{R}_{\geqslant 0} \to \mathbb{R}$. Like with data of the particle system, if $\varphi$ is evaluated at negative times we instead take its value at time 0.

**Proposition 6.2.** *Consider any time-scale $\mathfrak{t}_{\mathrm{reg}} \in \mathbb{R}_{\geqslant 0}$ satisfying $N^{-100} \lesssim \mathfrak{t}_{\mathrm{reg}} \lesssim N^{-1}$ and any $\varepsilon, C \in \mathbb{R}_{>0}$. Provided a random time $\mathfrak{t}_{\mathrm{st}} \in \mathbb{R}_{\geqslant 0}$ satisfying $\mathfrak{t}_{\mathrm{st}} \leqslant 1$ with probability 1, we have the following estimate outside an event of probability at most $\kappa_{\varepsilon,C} N^{-C}$:*

$$\|\nabla_{-\mathfrak{t}_{\mathrm{reg}}}^{\mathbf{T}} \bar{\mathfrak{Z}}^N\|_{\mathfrak{t}_{\mathrm{st}};\mathbb{T}_N} \lesssim_\varepsilon N^{-\frac{1}{2}+\varepsilon} + N^\varepsilon \mathfrak{t}_{\mathrm{reg}}^{\frac{1}{4}} + N^{-\frac{1}{2}+\varepsilon} \|\bar{\mathfrak{Z}}^N\|_{\mathfrak{t}_{\mathrm{st}};\mathbb{T}_N}^2 + N^\varepsilon \mathfrak{t}_{\mathrm{reg}}^{\frac{1}{4}} \|\bar{\mathfrak{Z}}^N\|_{\mathfrak{t}_{\mathrm{st}};\mathbb{T}_N}^2. \tag{6.1}$$

We explain Proposition 6.2. Assuming Theorem 1.7, the space-time process $\mathfrak{Z}^N$ looks like the SHE solution. The comparison in Proposition 4.6 says the same is true for $\bar{\mathfrak{Z}}^N$. The time-regularity of $\bar{\mathfrak{Z}}^N$ is then that of the SHE. This is Holder regularity with exponent $(\frac{1}{4})^-$. Forgetting $N^{-1/2+\varepsilon}$-terms and $N^\varepsilon$-factors on the RHS of the bound in Proposition 6.2 then gives the result.

To prove Proposition 6.2, we use the defining stochastic integral equation for $\bar{\mathfrak{Z}}^N$ introduced in Sect. 4 and then estimate time-regularity in the same way that we would estimate time-regularity of the solution of SHE through its mild Duhamel form. Let us clarify that the $N^\varepsilon$-factors, which are sub-optimal but sufficient for our purposes, are here to guarantee that such a strong time-regularity estimate proposed in Proposition 6.2 holds with the required high-probability. The $N^{-1/2+\varepsilon}$-errors that are not present in time-regularity estimates of the SHE solution come from the observation that the jumps in $\bar{\mathfrak{Z}}^N$ are order $N^{-1/2}$. We lastly remark that the $N^\varepsilon$-factors in the proposed upper bound in the statement of Proposition 6.2 will be harmless because we will only use Proposition 6.2 when we compute power-saving estimates for $\bar{\Phi}^{N,2}$-related terms, so we will only use Proposition 6.2 when we have a factor of $N^{-\beta_{\mathrm{univ}}}$ for $\beta_{\mathrm{univ}} \in \mathbb{R}_{>0}$ universal. Choosing $\varepsilon \in \mathbb{R}_{>0}$ in these $N^\varepsilon$-factors small then dwarfs them.

We additionally comment on the quadratic dependence of the $\|\|\|$-norm on the RHS of the proposed estimate in Proposition 6.2. This is mostly a technical consequence of obtaining time-regularity estimates for the stochastic-integral-type quantity in the stochastic equation for $\bar{\mathfrak{Z}}^N$ introduced in Sect. 4. In particular, we will need to apply martingale estimates to control the time-regularity of this quantity; a direct bound on $\bar{\mathfrak{Z}}^N$ in this stochastic integral type term would not let us employ martingale inequalities, so we need to take a little more care, and this gives quadratic dependence of the $\|\|\|$-norm.

Given that the proof of Proposition 6.2 looks like the proof of time-regularity estimates for the solution of SHE as we noted above, the reader is invited to skip or skim the remainder of the section. All we do is follow the proof of Proposition 3.2 in [10], namely the proof of (3.14) therein, with some technical adjustments to account for the $\|\|\|_{\mathfrak{t}_{st};\mathbb{T}_N}$-norms.

### 6.1. Proof of Proposition 6.2.

The first ingredient we use in the proof of Proposition 6.2 is simplification of Proposition 6.2. In particular, we will control the time-gradient of $\bar{\mathfrak{Z}}^N$ first on a discretization of the semi-discrete set $[0, \mathfrak{t}_{st}] \times \mathbb{T}_N$. This first step is easier since we can use a union bound as in the proof of Proposition 4.6. We also specialize to deterministic times $\mathfrak{t}_{st} \in \mathbb{R}_{\geqslant 0}$.

**Lemma 6.3.** *We admit the setting in* Proposition *6.2 and assume that* $\mathfrak{t}_{st} \in \mathbb{R}_{\geqslant 0}$ *is deterministic. Outside an event with probability at most* $\kappa_{\varepsilon,C} N^{-C}$, *we have the following estimate for the fully discretized norm defined in the statement of* Proposition *4.6:*

$$[\nabla^{\mathbf{T}}_{-\mathfrak{t}_{reg}} \bar{\mathfrak{Z}}^N]_{\mathfrak{t}_{st};\mathbb{T}_N} \lesssim_\varepsilon N^{-\frac{1}{2}+\varepsilon} + N^\varepsilon \mathfrak{t}_{reg}^{\frac{1}{4}} + N^{-\frac{1}{2}+\varepsilon} \|\bar{\mathfrak{Z}}^N\|^2_{\mathfrak{t}_{st};\mathbb{T}_N}$$
$$+ N^\varepsilon \mathfrak{t}_{reg}^{\frac{1}{4}} \|\bar{\mathfrak{Z}}^N\|^2_{\mathfrak{t}_{st};\mathbb{T}_N}. \tag{6.2}$$

We defer proof of Lemma 6.3 and other ingredients in the proof of Proposition 6.2 to future subsections. We now bootstrap to the semi-discrete norm $\|\|\|_{\mathfrak{t}_{st};\mathbb{T}_N}$ and to possibly random times $\mathfrak{t}_{st} \in \mathbb{R}_{\geqslant 0}$ through a stochastic continuity estimate.

**Lemma 6.4.** *Consider any possibly random time* $\mathfrak{t}_{st} \in \mathbb{R}_{\geqslant 0}$ *satisfying* $\mathfrak{t}_{st} \leqslant 1$ *with probability 1. Outside some event of probability at most* $\kappa_{\varepsilon,C} N^{-C}$ *with* $C \in \mathbb{R}_{>0}$ *arbitrarily large but universal and* $\varepsilon \in \mathbb{R}_{>0}$ *arbitrarily small but universal, we have*

$$\|\nabla^{\mathbf{T}}_{-\mathfrak{t}_{reg}} \bar{\mathfrak{Z}}^N\|_{\mathfrak{t}_{st};\mathbb{T}_N} \lesssim [\nabla^{\mathbf{T}}_{-\mathfrak{t}_{reg}} \bar{\mathfrak{Z}}^N]_{\mathfrak{t}_{st};\mathbb{T}_N} + N^{-\frac{1}{2}+\varepsilon} + N^{-\frac{1}{2}+\varepsilon} \|\bar{\mathfrak{Z}}^N\|^2_{\mathfrak{t}_{st};\mathbb{T}_N}. \tag{6.3}$$

*We define the random time* $\widetilde{\mathfrak{t}}_{st} \in \mathbb{R}_{\geqslant 0}$ *by mapping* $\mathfrak{t}_{st} \in \mathbb{R}_{\geqslant 0}$ *to the random closest element of the set* $\{\mathfrak{t}_j\}_{j=0}^{3N^{100}}$, *where* $\mathfrak{t}_j \stackrel{\bullet}{=} jN^{-100}$. *Outside an event of probability at most* $\kappa_{\varepsilon,C} N^{-C}$ *with* $C \in \mathbb{R}_{>0}$ *arbitrarily large and* $\varepsilon \in \mathbb{R}_{>0}$ *arbitrarily small but both universal,*

$$\|\nabla^{\mathbf{T}}_{-\mathfrak{t}_{reg}} \bar{\mathfrak{Z}}^N\|_{\mathfrak{t}_{st};\mathbb{T}_N} \lesssim \|\nabla^{\mathbf{T}}_{-\mathfrak{t}_{reg}} \bar{\mathfrak{Z}}^N\|_{\widetilde{\mathfrak{t}}_{st};\mathbb{T}_N} + N^{-\frac{1}{2}+\varepsilon} + N^{-\frac{1}{2}+\varepsilon} \|\bar{\mathfrak{Z}}^N\|^2_{\mathfrak{t}_{st};\mathbb{T}_N}.. \tag{6.4}$$

*Proof of Proposition 6.2.* If $\mathfrak{t}_{st} \in \mathbb{R}_{\geqslant 0}$ is deterministic, we apply the first estimate (6.3) of Lemma 6.4 followed by Lemma 6.3. It remains to study possibly random times $\mathfrak{t}_{st} \in \mathbb{R}_{\geqslant 0}$. Adopting notation of Lemma 6.4, by (6.4) it suffices to prove the claim for the random time $\widetilde{\mathfrak{t}}_{st} \in \mathbb{R}_{\geqslant 0}$ associated to $\mathfrak{t}_{st} \in \mathbb{R}_{\geqslant 0}$. Note $\widetilde{\mathfrak{t}}_{st} \in \mathbb{R}_{\geqslant 0}$ can take at most $5N^{100}$-many possible values. Using a union bound over such values, we may assume

$\widetilde{t}_{st}$ is deterministic and multiply the complement probability $\kappa_{\varepsilon,C} N^{-C}$ by $5N^{100}$. Thus, we get the proposed estimate with probability at least $1 - 5\kappa_{\varepsilon,C} N^{-C+100}$, so the proof is complete upon redefining $C \in \mathbb{R}_{\geqslant 0}$. □

*6.2. Proof of Lemma 6.3.* We start with time-regularity for the stochastic-integral-type term in the defining equation for $\bar{\mathfrak{z}}^N$.

**Lemma 6.5.** *We admit the setting of* Proposition 6.2 *and we assume* $t_{st} \in \mathbb{R}_{\geqslant 0}$ *is deterministic. Consider any space-time process* $\widetilde{\mathfrak{V}}^N$ *adapted to the canonical filtration of the particle system. Outside an event of probability at most $\kappa_{\varepsilon,C} N^{-C}$, we have*

$$[\nabla^{\mathbf{T}}_{-t_{reg}} \bar{\mathfrak{P}}^N (\widetilde{\mathfrak{V}}^N d\xi^N)]_{t_{st};\mathbb{T}_N} \lesssim_{\varepsilon} N^{-\frac{1}{2}+\varepsilon} + N^{\varepsilon} t_{reg}^{\frac{1}{4}} + N^{-\frac{1}{2}+\varepsilon} \|\widetilde{\mathfrak{V}}^N\|^2_{t_{st};\mathbb{T}_N} + N^{\varepsilon} t_{reg}^{\frac{1}{4}} \|\widetilde{\mathfrak{V}}^N\|^2_{t_{st};\mathbb{T}_N}. \tag{6.5}$$

*Proof of Lemma 6.3.* For convenience we recall the following stochastic equation for $\bar{\mathfrak{z}}^N$ with notation from Definition 4.3:

$$\bar{\mathfrak{z}}^N_{T,x} \overset{\bullet}{=} \bar{\mathfrak{P}}^{N,\mathbf{X}}_{T,x}(\chi_{\bullet}\mathfrak{z}^N_{0,\bullet}) + \bar{\mathfrak{P}}^N_{T,x}(\bar{\mathfrak{z}}^N d\xi^N) + \bar{\mathfrak{P}}^N_{T,x}(\bar{\Phi}^{N,2}) + \bar{\mathfrak{P}}^N_{T,x}(\bar{\Phi}^{N,3}). \tag{6.6}$$

We follow the proof of (3.14) in Proposition 3.2 in [10] with minor adjustments. Let us also assume $T \geqslant t_{reg}$. For times $T \in \mathbb{R}_{\geqslant 0}$ that do not satisfy this lower bound the time-gradient of $\bar{\mathfrak{z}}^N$ on scale $t_{reg}$ can be written as the time-gradient on a scale $t \leqslant t_{reg} \wedge T$ because we extend $\bar{\mathfrak{z}}^N$ to negative times by its value at time 0. Thus we get $t$-versions of $t_{reg}$-estimates below which are certainly controlled by such $t_{reg}$-estimates below as $t \leqslant t_{reg}$. First, we use a Chapman-Kolmogorov equation and a priori near-stationary moment bounds as in proof of (3.14) in Proposition 3.2 in [10] to get, for $\delta \in \mathbb{R}_{>0}$ arbitrarily small but universal and $p \in \mathbb{R}_{\geqslant 1}$, the following bound where we remark $\bar{\mathfrak{z}}^N_{0,\bullet} = \chi_{\bullet}\mathfrak{z}^N_{0,\bullet}$ admits near-stationary-type regularity estimates with respect to geodesic distance on $\mathbb{T}_N$ since $\chi$ itself has regularity and lets us ignore the boundary condition on the torus $\mathbb{T}_N$; see Remark 4.5:

$$\|\nabla^{\mathbf{T}}_{-t_{reg}} \bar{\mathfrak{P}}^{N,\mathbf{X}}_{T,x} \bar{\mathfrak{z}}^N_{0,\bullet}\|_{\omega;2p} \lesssim_{p,\delta} N^{-\frac{1}{2}+\delta} + N^{2\delta} t_{reg}^{1/4-\delta} \lesssim N^{-\frac{1}{2}+\delta} + N^{200\delta} t_{reg}^{1/4}. \tag{6.7}$$

Indeed, the last estimate comes from $t_{reg} \gtrsim N^{-100}$. As we later explain in the proof of Lemma 6.5, a union bound and Chebyshev inequality give, with required high probability, the following discretized bound that is $N^{100\delta}$-worse than the previous bound:

$$[\mathbf{1}_{T \geqslant t_{reg}} \nabla^{\mathbf{T}}_{-t_{reg}} \bar{\mathfrak{P}}^{N,\mathbf{X}}_{T,x} \bar{\mathfrak{z}}^N_{0,\bullet}]_{t_{st};\mathbb{T}_N} \lesssim_{\delta} N^{-\frac{1}{2}+101\delta} + N^{300\delta} t_{reg}^{1/4}. \tag{6.8}$$

We already treated the second term in the equation for $\bar{\mathfrak{z}}^N$ so we move to the $\bar{\Phi}^{N,2}$-term. Note $\|\bar{\Phi}^{N,2}\|_{t_{st};\mathbb{T}_N} \lesssim N^{1/2} \|\bar{\mathfrak{z}}^N\|_{t_{st};\mathbb{T}_N}$; we apply the time-regularity estimate (A.5) for the heat operator to get the following deterministic estimate for any $\delta \in \mathbb{R}_{>0}$:

$$\|\mathbf{1}_{T \geqslant t_{reg}} \nabla^{\mathbf{T}}_{-t_{reg}} \bar{\mathfrak{P}}^N_{T,x}(\bar{\Phi}^{N,2})\|_{t_{st};\mathbb{T}_N} \lesssim_{\delta} N^{2\delta} \|\bar{\Phi}^{N,2}\|_{t_{st};\mathbb{T}_N} t_{reg}$$
$$\lesssim N^{\frac{1}{2}+2\delta} t_{reg} \|\bar{\mathfrak{z}}^N\|_{t_{st};\mathbb{T}_N} \lesssim t_{reg}^{1/4} \|\bar{\mathfrak{z}}^N\|_{t_{st};\mathbb{T}_N}. \tag{6.9}$$

The final estimate giving the far RHS of (6.9) follows by recalling that $t_{reg} \lesssim N^{-1}$, and thus $t_{reg} \lesssim t_{reg}^{3/4} t_{reg}^{1/4} \lesssim N^{-3/4} t_{reg}^{1/4}$.

We now move to estimating the time-regularity of the heat operator term in the fixed-point equation for $\bar{3}^N$ corresponding to $\bar{\Phi}^{N,3}$. For the first term in $\bar{\Phi}^{N,3}$ we can use (A.5) as we did to establish (6.9) for $\bar{\Phi}^{N,2}$. However, the second term in $\bar{\Phi}^{N,3}$ has a gradient that we need to take advantage of as a naive bound of order $N$ would not suffice. To this end, we will instead use the space-time gradient bound (A.6). Recalling all coefficients hitting $\bar{3}^N$ in $\bar{\Phi}^{N,3}$ are uniformly bounded, for any $\delta \in \mathbb{R}_{>0}$ we have

$$\|\mathbf{1}_{T \geqslant \mathfrak{t}_{\mathrm{reg}}} \nabla^{\mathbf{T}}_{-\mathfrak{t}_{\mathrm{reg}}} \bar{\mathfrak{P}}^N_{T,x}(\bar{\Phi}^{N,3})\|_{\mathfrak{t}_{\mathrm{st}}; \mathbb{T}_N} \lesssim_\delta \mathfrak{t}^{\frac{1}{4}-\delta}_{\mathrm{reg}} \|\bar{3}^N\|_{\mathfrak{t}_{\mathrm{st}}; \mathbb{T}_N} \lesssim N^{200\delta} \mathfrak{t}^{\frac{1}{4}}_{\mathrm{reg}} \|\bar{3}^N\|_{\mathfrak{t}_{\mathrm{st}}; \mathbb{T}_N}.$$
(6.10)

We now combine (6.8), (6.9), (6.10), and Lemma 6.5 along with the fixed-point equation for $\bar{3}^N$ to complete the proof upon allowing an additional term on the RHS of the proposed estimate from the statement of Lemma 6.3 of type $N^{200\delta} \mathfrak{t}^{1/4}_{\mathrm{reg}} \|\bar{3}^N\|_{\mathfrak{t}_{\mathrm{st}}; \mathbb{T}_N}$. In particular, we get something with the correct $\mathfrak{t}_{\mathrm{reg}}$-dependence but that is linear in $\|\bar{3}^N\|_{\mathfrak{t}_{\mathrm{st}}; \mathbb{T}_N}$. However, to resolve such minor issue we use $\|\bar{3}^N\|_{\mathfrak{t}_{\mathrm{st}}; \mathbb{T}_N} \lesssim 1 + \|\bar{3}^N\|^2_{\mathfrak{t}_{\mathrm{st}}; \mathbb{T}_N}$ to get rid of this cosmetic difference, and this completes the proof. □

*Proof of Lemma 6.5.* Define $\widetilde{\mathfrak{V}} \overset{\bullet}{=} 1 + \|\bar{\mathfrak{V}}^N\|^2_{\mathfrak{t}_{\mathrm{st}}; \mathbb{T}_N}$. We claim that it suffices to establish the following for all $p \in \mathbb{R}_{\geqslant 1}$ and $\varepsilon \in \mathbb{R}_{>0}$ for reasons we explain afterwards that are related to the moment estimates/union bounds in the proof of Proposition 4.6:

$$\sup_{0 \leqslant T \leqslant \mathfrak{t}_{\mathrm{st}}} \sup_{x \in \mathbb{T}_N} \left\| \widetilde{\mathfrak{V}}^{-1} |\nabla^{\mathbf{T}}_{-\mathfrak{t}_{\mathrm{reg}}} \bar{\mathfrak{P}}^N_{T,x}(\bar{\mathfrak{V}}^N \mathrm{d}\xi^N)| \right\|_{\omega; 2p} \lesssim_{p,\varepsilon} N^{\frac{1}{2}\varepsilon} \mathfrak{t}^{\frac{1}{4}}_{\mathrm{reg}} + N^{-\frac{1}{2}+\frac{1}{2}\varepsilon}. \quad (6.11)$$

We have an additional $N^{-\varepsilon/2}$-factor in (6.11), so by Chebyshev's inequality for $p$-th moments we deduce the desired bound in Lemma 6.5 *for any fixed* $0 \leqslant T \leqslant \mathfrak{t}_{\mathrm{st}}$ and $x \in \mathbb{T}_N$ outside an event of probability $\kappa_{p,\varepsilon} N^{-p\varepsilon/2}$. A union bound over all points in the discretization $\mathfrak{I}_{\mathfrak{t}_{\mathrm{st}}} \times \mathbb{T}_N$ then gives the proposed estimate uniformly over $\mathfrak{I}_{\mathfrak{t}_{\mathrm{st}}} \times \mathbb{T}_N$ outside an event of probability $\kappa_{p,\varepsilon} N^{-p\varepsilon/2+200}$ as in the proof for Proposition 4.6. Choosing $p \gtrsim_\varepsilon 1$ sufficiently large depending only on $\varepsilon \in \mathbb{R}_{>0}$ completes the proof. To prove (6.11), we first take $T \geqslant \mathfrak{t}_{\mathrm{reg}}$ and then we take $\mathfrak{t}_{\mathrm{reg}} \geqslant N^{-2}$. We start with the following decomposition of the time-gradient from the LHS of (6.11) into a short-time/length-$\mathfrak{t}_{\mathrm{reg}}$-integral and an integral of a time-gradient of the heat kernel. First:

- Decompose the time-integration domain $[0, T]$ into $I_1 \cup I_2$ where $I_1$ looks at times until $T - \mathfrak{t}_{\mathrm{reg}}$ and $I_2$ looks at the rest.

An elementary calculation with the time-gradient of the heat operator, like with the proof for time-regularity in Proposition 3.2 in [10], gives control on the time-gradient of $\bar{\mathfrak{P}}^N$ by that of the integration-domain in $\bar{\mathfrak{P}}^N$ and that of the heat kernel therein:

$$|\nabla^{\mathbf{T}}_{-\mathfrak{t}_{\mathrm{reg}}} \bar{\mathfrak{P}}^N_{T,x}(\bar{\mathfrak{V}}^N \mathrm{d}\xi^N)| \leqslant |\bar{\mathfrak{P}}^N_{T,x}(\bar{\mathfrak{V}}^N_{S,y} \mathrm{d}\xi^N_{S,y} \mathbf{1}_{S \in I_2})|$$
$$+ |\bar{\mathfrak{P}}^N_{T,x}(\bar{\mathfrak{V}}^N_{S,y} \mathrm{d}\xi^N_{S,y} \mathbf{1}_{S \in I_1}) - \bar{\mathfrak{P}}^N_{T-\mathfrak{t}_{\mathrm{reg}},x}(\bar{\mathfrak{V}}^N_{S,y} \mathrm{d}\xi^N_{S,y} \mathbf{1}_{S \in I_1})|$$
$$\overset{\bullet}{=} \Phi_1 + \Phi_2. \quad (6.12)$$

We now consider the following decomposition for both $\Phi_1$, $\Phi_2$ in terms of level sets of the "squared-norm"-type quantity $\widetilde{\mathfrak{V}} \geqslant 1$:

$$|\Phi_{\mathfrak{i}}| \leqslant \sum_{\mathfrak{l}=1}^{\infty} |\Phi_{\mathfrak{i}}| \cdot \mathbf{1}_{\mathfrak{l} \leqslant (\widetilde{\mathfrak{V}})^{\frac{1}{2}} \leqslant \mathfrak{l}+1}. \quad (6.13)$$

We observe the product of any indicator function in the summation on the RHS of the above inequality with any other of these indicator functions can only avoid vanishing when multiplied by at most 2 other of these indicator functions. Thus, binomial expansion and this "almost-orthogonality" of the indicator functions give the following multiscale estimate for $p \in \mathbb{R}_{\geqslant 1}$:

$$\|\widetilde{\mathfrak{V}}^{-1}|\Phi_i|\|_{\omega;2p}^{2p} \lesssim_p \sum_{\mathfrak{l}=1}^{\infty} \|\widetilde{\mathfrak{V}}^{-1}|\Phi_i|\mathbf{1}_{\mathfrak{l}\leqslant(\mathfrak{V})^{\frac{1}{2}}\leqslant\mathfrak{l}+1}\|_{\omega;2p}^{2p}$$

$$\lesssim \sum_{\mathfrak{l}=1}^{\infty} \mathfrak{l}^{-4p}\|\Phi_i\mathbf{1}_{\mathfrak{l}\leqslant(\mathfrak{V})^{\frac{1}{2}}\leqslant\mathfrak{l}+1}\|_{\omega;2p}^{2p}$$

$$\lesssim \sup_{\mathfrak{l}\in\mathbb{Z}_{\geqslant 1}} \mathfrak{l}^{-2p}\|\Phi_i\mathbf{1}_{(\mathfrak{V})^{\frac{1}{2}}\leqslant\mathfrak{l}+1}\|_{\omega;2p}^{2p} \sum_{\mathfrak{l}=1}^{\infty} \mathfrak{l}^{-2p}$$

$$\lesssim \sup_{\mathfrak{l}\in\mathbb{Z}_{\geqslant 1}} \mathfrak{l}^{-2p}\|\Phi_i\mathbf{1}_{(\mathfrak{V})^{\frac{1}{2}}\leqslant\mathfrak{l}+1}\|_{\omega;2p}^{2p}. \tag{6.14}$$

At this point we will proceed with moment estimates almost exactly as in the proof of Proposition 3.2 in [10]. We will only use the indicator function hitting $\Phi_i$ in (6.14) to get an a priori estimate; we afterwards quickly remove this indicator function. We now specialize the index $i \in \{1, 2\}$; we start with $i = 1$. In particular, we will estimate $\Phi_1$ starting with the estimate (6.14). To this end, we first define the uniformly bounded process $\mathfrak{R}^N = \mathfrak{l}^{-1}\widetilde{\mathfrak{V}}^N\mathbf{1}[|\mathfrak{l}^{-1}\widetilde{\mathfrak{V}}^N| \lesssim 1]$. Concluding with the martingale bound Lemma B.1, which is a version of Lemma 3.1 in [10] and additionally needs uniform boundedness/adaptedness of $\mathfrak{R}^N$, we have

$$\mathfrak{l}^{-2p}\|\Phi_1\mathbf{1}_{(\mathfrak{V})^{\frac{1}{2}}\leqslant\mathfrak{l}+1}\|_{\omega;2p}^{2p} \leqslant \mathbf{E}|\int_{T-\mathfrak{t}_{reg}}^{T}\sum_{y\in\mathbb{T}_N}\bar{\mathfrak{P}}_{S,T,x,y}^N\cdot\mathfrak{l}^{-1}\bar{\mathfrak{V}}_{S,y}^N\mathrm{d}\xi_{S,y}^N|^{2p}\mathbf{1}_{\|\bar{\mathfrak{V}}\|_{\mathfrak{t}_{st};\mathbb{T}_N}\lesssim\mathfrak{l}+1}$$

$$\tag{6.15}$$

$$\leqslant \mathbf{E}|\int_{T-\mathfrak{t}_{reg}}^{T}\sum_{y\in\mathbb{T}_N}\bar{\mathfrak{P}}_{S,T,x,y}^N\cdot\mathfrak{R}_{S,y}^N\mathrm{d}\xi_{S,y}^N|^{2p} \tag{6.16}$$

$$\lesssim_p \left(\int_{T-\mathfrak{t}_{reg}}^{T} N\sum_{y\in\mathbb{T}_N}|\bar{\mathfrak{P}}_{S,T,x,y}^N|^2\,\mathrm{d}S\right)^p \overset{\bullet}{=} \Psi_1. \tag{6.17}$$

Indeed if $\|\bar{\mathfrak{V}}\|_{\mathfrak{t}_{st};\mathbb{T}_N} \lesssim \mathfrak{l}+1 \leqslant 2\mathfrak{l}$, then $\mathfrak{l}^{-1}\bar{\mathfrak{V}}^N$ is uniformly bounded above so in this case the cutoff $\mathfrak{R}^N$ does nothing and we get $\mathfrak{R}^N = \mathfrak{l}^{-1}\bar{\mathfrak{V}}^N$. We note that $\mathfrak{R}^N$ is adapted because $\widetilde{\mathfrak{V}}^N$ is adapted and therefore so is the indicator function in $\mathfrak{R}^N$ while $\mathfrak{l} \in \mathbb{Z}_{\geqslant 1}$ is deterministic. We now estimate $\Phi_2$ and we start with (6.14) with the index $i = 2$. By Lemma B.1, we similarly get

$$\mathfrak{l}^{-2p}\|\Phi_2\mathbf{1}_{(\mathfrak{V})^{\frac{1}{2}}\leqslant\mathfrak{l}+1}\|_{\omega;2p}^{2p} \leqslant \mathbf{E}|\int_{0}^{T-\mathfrak{t}_{reg}}\sum_{y\in\mathbb{T}_N}\nabla_{-\mathfrak{t}_{reg}}^{\mathbf{T}}\bar{\mathfrak{P}}_{S,T,x,y}^N\cdot\mathfrak{R}_{S,y}^N\mathrm{d}\xi_{S,y}^N|^{2p}\mathbf{1}_{\|\bar{\mathfrak{V}}\|_{\mathfrak{t}_{st};\mathbb{T}_N}\lesssim\mathfrak{l}+1}$$

$$\tag{6.18}$$

$$\lesssim_p \left(\int_{0}^{T-\mathfrak{t}_{reg}} N\sum_{y\in\mathbb{T}_N}|\nabla_{-\mathfrak{t}_{reg}}^{\mathbf{T}}\bar{\mathfrak{P}}_{S,T,x,y}^N|^2\,\mathrm{d}S\right)^p \overset{\bullet}{=} \Psi_2. \tag{6.19}$$

The second line requires dropping the indicator from the bound before. Combining (6.12), (6.14), (6.17), and (6.19) provides the following upper bound which is familiar from the proof of time-regularity in Proposition 3.2 in [10]:

$$\left\| \widetilde{\mathfrak{V}}^{-1} | \nabla^{\mathbf{T}}_{-\mathfrak{t}_{\text{reg}}} \bar{\mathfrak{P}}^N_{T,x} (\bar{\mathfrak{V}}^N \mathrm{d}\xi^N) | \right\|^{2p}_{\omega;2p} \lesssim_p \sum_{\mathsf{i}=1,2} \| \widetilde{\mathfrak{V}}^{-1} | \Phi_\mathsf{i} | \|^{2p}_{\omega;2p}$$

$$\lesssim_p \sum_{\mathsf{i}=1,2} \sup_{\mathfrak{l} \in \mathbb{Z}_{\geqslant 1}} \mathfrak{l}^{-2p} \| \Phi_\mathsf{i} \mathbf{1}_{(\mathfrak{V})^{\frac{1}{2}} \leqslant \mathfrak{l}+1} \|^{2p}_{\omega;2p} \lesssim_p \sum_{\mathsf{i}=1,2} \Psi_\mathsf{i}. \qquad (6.20)$$

We now follow the proof of (3.14) in Proposition 3.2 in [10] to bound $\Psi_\mathsf{i}$-terms from (6.17) and (6.19) and use the on-diagonal heat kernel bounds in Corollary A.2 in [10] that hold for the $\bar{\mathfrak{P}}^N$-heat kernel for order 1 times; see Remark 4.2. For any $\delta \in \mathbb{R}_{>0}$, we obtain the following final estimate in this proof:

$$\left\| \widetilde{\mathfrak{V}}^{-1} | \nabla^{\mathbf{T}}_{-\mathfrak{t}_{\text{reg}}} \bar{\mathfrak{P}}^N_{T,x} (\bar{\mathfrak{V}}^N \mathrm{d}\xi^N) | \right\|_{\omega;2p} \lesssim_{p,\delta} \mathfrak{t}^{\frac{1}{4}-\delta}_{\text{reg}} \lesssim N^{2\delta} \mathfrak{t}^{\frac{1}{4}}_{\text{reg}}. \qquad (6.21)$$

The final bound follows from the assumption $\mathfrak{t}_{\text{reg}} \geqslant N^{-2}$. This completes the proof for $\mathfrak{t}_{\text{reg}} \geqslant N^{-2}$. In the case $\mathfrak{t}_{\text{reg}} \leqslant N^{-2}$, we use the same strategy although Lemma B.1 provides, instead of the $\Psi_\mathsf{i}$-terms in (6.20), the upper bound of $N^{-1/2}$ times $p$-dependent factors so we are done in the case $T \geqslant \mathfrak{t}_{\text{reg}}$. To remove this final assumption for general times we can write the $\mathfrak{t}_{\text{reg}}$-time-gradient as a time-gradient on a time-scale $\mathfrak{t} \in \mathbb{R}_{\geqslant 0}$ which is at most $\mathfrak{t}_{\text{reg}}$ because we extend everything to negative times by their value at time 0. The previous argument then gives $\mathfrak{t}$-versions of the estimates above which are certainly controlled by the $\mathfrak{t}_{\text{reg}}$-estimates above since $\mathfrak{t} \leqslant \mathfrak{t}_{\text{reg}}$. This was the same observation that we made in the proof of Lemma 6.3. This completes the proof. □

*6.3. Proof of Lemma 6.4.* The differences between $\| \| \|_{\mathfrak{t}_{\text{st}}; \mathbb{T}_N}$ and $[ ]_{\mathfrak{t}_{\text{st}}; \mathbb{T}_N}$ come via behavior between times in the discretization $\mathfrak{I}_{\mathfrak{t}_{\text{st}}}$ defining $[ ]_{\mathfrak{t}_{\text{st}}; \mathbb{T}_N}$. To apply this to $\bar{\mathfrak{Z}}^N$, we note the following SDE-type formulation for $\bar{\mathfrak{Z}}^N$-dynamics and explain it after:

$$\mathrm{d}\bar{\mathfrak{Z}}^N_{\mathfrak{t},x} = \mathscr{L}^{!!} \bar{\mathfrak{Z}}^N_{\mathfrak{t},x} \mathrm{d}t + \bar{\mathfrak{Z}}^N_{\mathfrak{t},x} \mathrm{d}\xi^N_{\mathfrak{t},x} + \bar{\Phi}^{N,2}_{\mathfrak{t},x} \mathrm{d}t + \bar{\Phi}^{N,3}_{\mathfrak{t},x} \mathrm{d}t. \qquad (6.22)$$

By SDE-type we mean $\mathrm{d}\xi^N$ is a discrete approximation to Brownian motions. We recall the notation in this equation in Sect. 4. We emphasize (6.22) can be checked by taking the time-differential of the defining linear equation for $\bar{\mathfrak{Z}}^N$ in Definition 4.3.

We observe the SDE-type Eq. (6.22) admits a continuous part corresponding to the first, third, and fourth terms on the RHS of (6.22) as well as the continuous "compensation" part of the compensated Poisson process $\mathrm{d}\xi^N$. This latter continuous part in $\mathrm{d}\xi^N$ is a random drift with speed bounded by $N^2$ times universal factors. If (6.22) only had these continuous parts, then (6.22) would be a random *linear* ODE. Short-time estimates would then follow by standard procedure for linear ODE.

As for jumps in (6.22), by definition of $\mathrm{d}\xi^N$ from Proposition 2.7/(2.4) from [10], any jump in the SDE-type equation (6.22) is order $N^{-1/2}$. We will use this and control on the number of jumps in short times via tail bounds for the Poisson distribution.

**Lemma 6.6.** *Outside an event of probability at most $\kappa_{\varepsilon,C} N^{-C}$ for $C \in \mathbb{R}_{>0}$ arbitrarily large and $\varepsilon \in \mathbb{R}_{>0}$ arbitrarily small but both universal, we have the following uniformly over $0 \leqslant T \leqslant 2$ and $x \in \mathbb{T}_N$ simultaneously on this same high-probability event:*

$$\sup_{0 \leqslant \mathfrak{s} \leqslant N^{-99}} |\nabla_{\pm\mathfrak{s}}^{\mathbf{T}} \bar{\mathfrak{Z}}_{T,x}^N| \lesssim_\varepsilon N^{-\frac{1}{2}+\varepsilon} \left( 1 + \|\bar{\mathfrak{Z}}^N\|_{T;\mathbb{T}_N}^2 \right) = N^{-\frac{1}{2}+\varepsilon} + N^{-\frac{1}{2}+\varepsilon} \|\bar{\mathfrak{Z}}^N\|_{T;\mathbb{T}_N}^2.$$

(6.23)

*Proof of Lemma 6.4.* We will first prove (6.3). Provided any $\mathfrak{t} \in [0, \mathfrak{t}_{st}]$, we first employ the following decomposition in which $\mathfrak{t}_j \in \mathbb{R}_{\geqslant 0}$ denotes the time *in the discretization* $\mathfrak{I}_{\mathfrak{t}_{st}} = \{\mathfrak{t}_{st} \cdot jN^{-100}\}_{j=1}^{N^{100}}$ that is closest to $\mathfrak{t}$, and $x \in \mathbb{T}_N$ is any arbitrary point:

$$|\nabla_{-\mathfrak{t}_{reg}} \bar{\mathfrak{Z}}_{\mathfrak{t},x}^N| = |\bar{\mathfrak{Z}}_{\mathfrak{t}-\mathfrak{t}_{reg},x}^N - \bar{\mathfrak{Z}}_{\mathfrak{t},x}^N|$$

$$\leqslant |\bar{\mathfrak{Z}}_{\mathfrak{t}-\mathfrak{t}_{reg},x}^N - \bar{\mathfrak{Z}}_{\mathfrak{t}_j-\mathfrak{t}_{reg},x}^N| + |\bar{\mathfrak{Z}}_{\mathfrak{t}_j-\mathfrak{t}_{reg},x}^N - \bar{\mathfrak{Z}}_{\mathfrak{t}_j,x}^N| + |\bar{\mathfrak{Z}}_{\mathfrak{t}_j,x}^N - \bar{\mathfrak{Z}}_{\mathfrak{t},x}^N|$$ (6.24)

$$\leqslant |\bar{\mathfrak{Z}}_{\mathfrak{t}-\mathfrak{t}_{reg},x}^N - \bar{\mathfrak{Z}}_{\mathfrak{t}_j-\mathfrak{t}_{reg},x}^N| + |\bar{\mathfrak{Z}}_{\mathfrak{t}_j,x}^N - \bar{\mathfrak{Z}}_{\mathfrak{t},x}^N| + [\nabla_{-\mathfrak{t}_{reg}}^{\mathbf{T}} \bar{\mathfrak{Z}}_{T,x}^N]_{\mathfrak{t}_{st};\mathbb{T}_N}.$$ (6.25)

Observe $|\mathfrak{t} - \mathfrak{t}_j| \leqslant N^{-100}$, as the points in $\mathfrak{I}_{\mathfrak{t}_{st}}$ are separated by at most $N^{-100}$. Thus, estimating the first two terms in (6.25) with the desired high probability via Lemma 6.6 with $T = \mathfrak{t} - \mathfrak{t}_{reg}$, $\mathfrak{t}$ and $\mathfrak{s} = |\mathfrak{t} - \mathfrak{t}_j|$ gives (6.3) as $\mathfrak{t} - \mathfrak{t}_{reg}$, $\mathfrak{t} \leqslant \mathfrak{t}_{st}$ so we may bound the RHS of (6.23) for these choices of $T$ with $1 + \|\bar{\mathfrak{Z}}^N\|_{\mathfrak{t};\mathbb{T}_N}^2$, $1 + \|\bar{\mathfrak{Z}}^N\|_{\mathfrak{t}-\mathfrak{t}_{reg};\mathbb{T}_N}^2 \leqslant 1 + \|\bar{\mathfrak{Z}}^N\|_{\mathfrak{t}_{st};\mathbb{T}_N}^2$ and control all norms by $\|\bar{\mathfrak{Z}}^N\|_{\mathfrak{t}_{st};\mathbb{T}_N}$.

We now prove (6.4). If $\mathfrak{t}_{st} \leqslant \tilde{\mathfrak{t}}_{st}$, there is nothing to do, so assume $\tilde{\mathfrak{t}}_{st} \leqslant \mathfrak{t}_{st}$. Take any $\mathfrak{t} \in [\tilde{\mathfrak{t}}_{st}, \mathfrak{t}_{st}]$ and any $x \in \mathbb{T}_N$. Like above,

$$|\nabla_{-\mathfrak{t}_{reg}} \bar{\mathfrak{Z}}_{\mathfrak{t},x}^N| = |\bar{\mathfrak{Z}}_{\mathfrak{t}-\mathfrak{t}_{reg},x}^N - \bar{\mathfrak{Z}}_{\mathfrak{t},x}^N|$$

$$\leqslant |\bar{\mathfrak{Z}}_{\mathfrak{t}-\mathfrak{t}_{reg},x}^N - \bar{\mathfrak{Z}}_{\tilde{\mathfrak{t}}_{st}-\mathfrak{t}_{reg},x}^N| + |\bar{\mathfrak{Z}}_{\tilde{\mathfrak{t}}_{st}-\mathfrak{t}_{reg},x}^N - \bar{\mathfrak{Z}}_{\tilde{\mathfrak{t}}_{st},x}^N| + |\bar{\mathfrak{Z}}_{\tilde{\mathfrak{t}}_{st},x}^N - \bar{\mathfrak{Z}}_{\mathfrak{t},x}^N|$$ (6.26)

$$\leqslant |\bar{\mathfrak{Z}}_{\mathfrak{t}-\mathfrak{t}_{reg},x}^N - \bar{\mathfrak{Z}}_{\tilde{\mathfrak{t}}_{st}-\mathfrak{t}_{reg},x}^N| + |\bar{\mathfrak{Z}}_{\tilde{\mathfrak{t}}_{st},x}^N - \bar{\mathfrak{Z}}_{\mathfrak{t},x}^N| + \|\nabla_{-\mathfrak{t}_{reg}}^{\mathbf{T}} \bar{\mathfrak{Z}}_{T,x}^N\|_{\tilde{\mathfrak{t}}_{st};\mathbb{T}_N}.$$ (6.27)

Observe that $|\mathfrak{t} - \tilde{\mathfrak{t}}_{st}| = |\mathfrak{t} - \mathfrak{t}_{reg} - (\tilde{\mathfrak{t}}_{st} - \mathfrak{t}_{reg})| \leqslant N^{-100}$ because $|\mathfrak{t}_{st} - \tilde{\mathfrak{t}}_{st}| \leqslant N^{-100}$ by construction. Thus, as with deducing (6.3) by (6.25), we get (6.4) with the required high-probability from (6.27) via Lemma 6.6 with $T = \mathfrak{t} - \mathfrak{t}_{reg}$, $\mathfrak{t}$ and $\mathfrak{s} = |\tilde{\mathfrak{t}}_{st} - \mathfrak{t}|$. ☐

*Proof of Lemma 6.6.* We prove (6.23) for forwards time-gradient, so $+\mathfrak{s}$. The proof for $-\mathfrak{s}$ amounts to the following argument but backwards in time. Define $\mathfrak{I}_j \overset{\bullet}{=} [\mathfrak{t}_j, \mathfrak{t}_{j+1}]$ with $\mathfrak{t}_j = jN^{-99}$ and $j \in [\![0, 3N^{99}]\!]$. We restrict to the following event denoted by $\mathscr{E}$.

- The number of ringings among all Poisson clocks on $\mathbb{T}_N$ in any $\mathfrak{I}_j$-interval is at most $N^\delta$ with $\delta \in \mathbb{R}_{>0}$ small but universal.

We first observe that if $\mathscr{E}^C$ is the complement event to $\mathscr{E}$, we have $\mathbf{P}[\mathscr{E}^C] \lesssim_{\delta,C} N^{-C}$. Indeed, the sum of all Poisson clocks over all points in $\mathbb{T}_N$ is another Poisson clock of rate at most $N^{10}$ times uniformly bounded factors upon considering there are at most $N^{5/4+\varepsilon_{cpt}}$-many clocks each of rate at most $N^2$, both again times uniformly bounded factors. Thus, the probability that there are more than $N^\delta$-many ringings in any interval $\mathfrak{I}_j$ of time-scale $N^{-99}$ is exponentially small in $N \in \mathbb{Z}_{>0}$. The estimate for $\mathbf{P}[\mathscr{E}^C]$ then follows by taking a union bound of these *exponentially small* probabilities over all *polynomially many* $j \in [\![0, 3N^{99}]\!]$.

We proceed to show that on the event $\mathscr{E}$ above, the proposed estimate (6.23) holds. We make the following observations.

- Observe with probability 1 on the event $\mathscr{E}$, for any $0 \leqslant T \leqslant 2$ the total number of Poisson clocks ringing in the time-window $[T, T + N^{-99}]$ is at most $N^{2\delta}$, as this time-window is covered by a uniformly bounded number of intervals $\mathfrak{I}_\mathfrak{j}$ for $\mathfrak{j} \in [\![0, 3N^{99}]\!]$.
- Between jumps (6.22) gives $\partial_T \bar{\mathfrak{Z}}^N_{T,x} = \mathscr{Q}\bar{\mathfrak{Z}}^N_{T,x}$, where $\mathscr{Q}$ is a linear operator satisfying the following with $\kappa \in \mathbb{R}_{\geqslant 0}$ universal:

$$\sup_{x \in \mathbb{T}_N} |\mathscr{Q}\bar{\mathfrak{Z}}^N_{T,x}| \leqslant \kappa N^2 \sup_{x \in \mathbb{T}_N} |\bar{\mathfrak{Z}}^N_{T,x}|. \tag{6.28}$$

This can be checked directly. Thus, provided any $T \in [0, \mathfrak{t}_{st}]$, for all $T + \mathfrak{s} \in \mathbb{R}_{\geqslant 0}$ before the first jump time after $T \in \mathbb{R}_{\geqslant 0}$, assuming $\mathfrak{s} \leqslant N^{-99}$ we have the following by expanding the relevant exponential matrix/operator below in a power series:

$$|\nabla^{\mathbf{T}}_{\mathfrak{s}} \bar{\mathfrak{Z}}^N_{T,x}| = |\left(\mathbf{Exp}(\mathfrak{s}\mathscr{Q})\bar{\mathfrak{Z}}^N\right)_{T,x} - \bar{\mathfrak{Z}}^N_{T,x}| \leqslant \sum_{\mathfrak{l}=1}^\infty \mathfrak{s}^{\mathfrak{l}}(\mathfrak{l}!)^{-1}|\mathscr{Q}^{\mathfrak{l}}\bar{\mathfrak{Z}}^N_{T,x}|$$

$$\leqslant \sum_{\mathfrak{l}=1}^\infty N^{-99\mathfrak{l}}(\mathfrak{l}!)^{-1}\kappa^{\mathfrak{l}}N^{2\mathfrak{l}}\|\bar{\mathfrak{Z}}^N\|_{T;\mathbb{T}_N} \lesssim N^{-97}\|\bar{\mathfrak{Z}}^N\|_{T;\mathbb{T}_N}. \tag{6.29}$$

- The jumps of $\bar{\mathfrak{Z}}^N$ at any space-time point have sizes that are bounded above by $N^{-1/2}$ times the value of $\bar{\mathfrak{Z}}^N$ at that space-time point times uniformly bounded factors. This follows by definition of the martingale differential $d\xi^N$ within (2.4) of [10].

Iterating the last two bullet points at most $N^{2\delta}$-many times, given the first bullet point in that list, yields the proof if we replace $1 + \|\bar{\mathfrak{Z}}^N_{T,x}\|^2_{T;\mathbb{T}_N}$ with $\|\bar{\mathfrak{Z}}^N_{T,x}\|_{T;\mathbb{T}_N}$ in the proposed bound. But $\|\bar{\mathfrak{Z}}^N_{T,x}\|_{T;\mathbb{T}_N} \lesssim 1 + \|\bar{\mathfrak{Z}}^N_{T,x}\|^2_{T;\mathbb{T}_N}$ by the Cauchy-Schwarz inequality, so we deduce the proposed estimate as written as well. $\qquad \square$

# 7. Key Estimates for Pseudo-Gradients and $\Phi^{N,2}$

We will control $\bar{\mathfrak{P}}^N(\bar{\Phi}^{N,2})$ in the defining equation for $\bar{\mathfrak{Z}}^N$ via the following version of Pseudo-Proposition 2.11.

**Proposition 7.1.** *Consider a random time $\mathfrak{t}_{st} \in \mathbb{R}_{\geqslant 0}$ with $\mathfrak{t}_{st} \leqslant 1$ with probability 1. There exist universal constants $\beta_{univ,1} \in \mathbb{R}_{>0}$ and $\beta_{univ,2} \in \mathbb{R}_{>0}$ such that outside an event of probability at most $N^{-\beta_{univ,1}}$ times uniformly bounded factors, we have*

$$\|\bar{\mathfrak{P}}^N(\bar{\Phi}^{N,2})\|_{\mathfrak{t}_{st};\mathbb{T}_N} \lesssim N^{-\beta_{univ,2}} + N^{-\beta_{univ,2}}\|\bar{\mathfrak{Z}}^N\|^2_{\mathfrak{t}_{st};\mathbb{T}_N}. \tag{7.1}$$

*We recall $\bar{\Phi}^{N,2}$ from the defining Eq. (4.3) for $\bar{\mathfrak{Z}}^N$. It contains the relevant pseudo-gradient data for $\bar{\mathfrak{Z}}^N$-dynamics.*

We invite the reader to take Proposition 7.1 for granted and proceed to the next section to see how we use it to get Theorem 1.7, at least in a first reading. The proof of Proposition 7.1 is quite technical. It consists of the following three ingredients, which we overviewed in the would-be-proof of Pseudo-Proposition 2.11.

- We replace the pseudo-gradient-spatial average in $\bar{\Phi}^{N,2}$ with a cutoff via Proposition 5.2 in view of Pseudo-Proposition 2.12.

- We then replace the spatial-average of pseudo-gradients, now with a cutoff given the previous bullet point, with its dynamic average on the time-scale $\mathfrak{t}_{\mathrm{av},\mathfrak{m}_+} \in \mathbb{R}_{>0}$ from the multiscale estimate Corollary 5.12. We do this by replacement with dynamic averages on progressively larger time-scales. Here, time-regularity for the heat kernel $\bar{\mathfrak{P}}^N$ and for $\bar{\mathfrak{Z}}^N$, the latter established in Proposition 6.2, will be crucial. The multiscale estimate in Corollary 5.14 will be crucial too. We do something similar for the term in $\bar{\Phi}^{N,2}$ of order $N^{\beta_X}$ but via Corollary 5.16 instead of Corollary 5.14. This step resembles Pseudo-Proposition 2.13.
- The third step is to estimate the pseudo-gradient content of $\bar{\Phi}^{N,2}$ with their dynamic averages from the previous bullet point. We employ another multiscale strategy via replacements by progressively sharper cutoff on progressively bigger time-scale as outlined in the would-be-proof of Pseudo-Proposition 2.14. Corollary 5.12 and Corollary 5.13 are important for this step.

We now introduce some useful notation which is sometimes referred to as a squared version of the "Japanese bracket" in PDE.

**Definition 7.2.** For $\mathfrak{t}_{\mathrm{st}} \in \mathbb{R}_{\geqslant 0}$, define $\langle \rangle_{\mathfrak{t}_{\mathrm{st}};\mathbb{T}_N} \overset{\bullet}{=} 1 + \|\|_{\mathfrak{t}_{\mathrm{st}};\mathbb{T}_N}^2$ acting on $\mathfrak{f} : \mathbb{R}_{\geqslant 0} \times \mathbb{T}_N \to \mathbb{R}$. Via the Cauchy-Schwarz inequality, we have the following estimate, which we will frequently use in this section, that compares this bracket to the $\|\|_{\mathfrak{t}_{\mathrm{st}};\mathbb{T}_N}$-norm:

$$\|\mathfrak{f}\|_{\mathfrak{t}_{\mathrm{st}};\mathbb{T}_N} \leqslant \langle \mathfrak{f} \rangle_{\mathfrak{t}_{\mathrm{st}};\mathbb{T}_N}. \tag{7.2}$$

A significant portion of our upcoming analysis for order-$N^{1/2}$ terms in $\bar{\Phi}^{N,2}$ holds equally well for order-$N^{\beta_X}$ terms in $\bar{\Phi}^{N,2}$ with a possible exception of elementary but ultimately negligible adjustments in power-counting and application of Corollary 5.13/Corollary 5.16 instead of Corollary 5.12/Corollary 5.14. So, we make the following "formal replacement" occasionally:

$$N^{\frac{1}{2}} \to N^{\beta_X} \quad \text{and} \quad \mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}) \to \widetilde{\sum}_{\mathfrak{l}=1,\ldots,N^{\beta_X}} \widetilde{\mathfrak{g}}^{\mathfrak{l}}. \tag{7.3}$$

*7.1. Proof of Proposition 7.1.* First we provide the key inputs each of which correspond to a bullet point in the list following Proposition 7.1. We then combine them to prove Proposition 7.1. The rest of the section is dedicated to proofs of each input.

We first introduce some notation for the aforementioned key inputs which capture the errors in the first two bullet points in the outline above. Again, these errors come from introducing cutoff into a spatial-average and replacements-by-time-average.

- Recall the spatial-average of pseudo-gradients $\mathscr{A}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g})$ in Proposition 2.7 and its cutoff $\mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g})$ in Definition 5.1. We define

$$\Psi^{N,1} \overset{\bullet}{=} N^{\frac{1}{2}} \mathscr{A}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}) - N^{\frac{1}{2}} \mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}). \tag{7.4}$$

- Taking $\mathfrak{t}_{\mathrm{av},\mathfrak{l}_+} \in \mathbb{R}_{>0}$ as the "maximal time" defined in Corollary 5.14, we define the difference-via-time-average-replacement below in which we recall the space-time averages in Definition 5.3 and the cutoff-spatial-average in Definition 5.1:

$$\Psi^{N,2} \overset{\bullet}{=} N^{\frac{1}{2}} \mathscr{A}_{\mathfrak{t}_{\mathrm{av},\mathfrak{l}_+}}^{\mathbf{T},+} \mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}) - N^{\frac{1}{2}} \mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}). \tag{7.5}$$

Taking $\mathfrak{t}^{\sim}_{\mathrm{av},\mathfrak{l}^{\sim}_+} \in \mathbb{R}_{>0}$ as the "maximal time" defined in Corollary 5.16, we define the difference-via-time-average-replacement below in which we recall the time-average in Definition 5.8 of the functional-with-pseudo-gradient-factor:

$$\widetilde{\Psi}^{N,2} \triangleq N^{\beta_X} \widetilde{\sum_{\mathfrak{l}=1,\ldots,N^{\beta_X}}} \left( \mathscr{A}^{\mathbf{T},+}_{\mathfrak{t}^{\sim}_{\mathrm{av},\mathfrak{l}^{\sim}_+}} (\widetilde{\mathfrak{g}}^{\mathfrak{l}}) - \widetilde{\mathfrak{g}}^{\mathfrak{l}} \right). \tag{7.6}$$

The first technical ingredient corresponds to the first bullet point from the list following the statement of Proposition 7.1. The proof of Lemma 7.3 below is a direct consequence of the Markov inequality and Proposition 5.2, in short.

**Lemma 7.3.** *Consider a random time* $\mathfrak{t}_{\mathrm{st}} \in \mathbb{R}_{\geqslant 0}$ *with* $\mathfrak{t}_{\mathrm{st}} \leqslant 1$ *with probability 1. There exist universal constants* $\beta_{\mathrm{univ},1} \in \mathbb{R}_{>0}$ *and* $\beta_{\mathrm{univ},2} \in \mathbb{R}_{>0}$ *such that outside an event of probability at most* $N^{-\beta_{\mathrm{univ},1}}$ *times uniformly bounded factors, we have*

$$\|\bar{\mathfrak{P}}^N_{T,x}(\Psi^{N,1}\bar{\mathfrak{Z}}^N)\|_{\mathfrak{t}_{\mathrm{st}};\mathbb{T}_N} \lesssim N^{-\beta_{\mathrm{univ},2}} \langle \bar{\mathfrak{Z}}^N \rangle_{\mathfrak{t}_{\mathrm{st}};\mathbb{T}_N}. \tag{7.7}$$

The second technical ingredient we require is the replacement of $\mathscr{C}^{\mathbf{X},-}$ in $\Psi^{N,1}$ within the previous Lemma 7.3 with its time-average on mesoscopic time-scale that corresponds to the second bullet point in the list following the statement of Proposition 7.1. Roughly speaking we may replace local functionals with their time-averages inside the heat operator $\bar{\mathfrak{P}}^N$ up to error terms controlled by the time-regularity of the heat kernel and of $\bar{\mathfrak{Z}}^N$. We apply Proposition 6.2 and time-regularity of the heat kernel in Lemma A.1 to control these errors. We give details for this heuristic after our proof of Proposition 7.1.

**Lemma 7.4.** *Consider a random time* $\mathfrak{t}_{\mathrm{st}} \in \mathbb{R}_{\geqslant 0}$ *with* $\mathfrak{t}_{\mathrm{st}} \leqslant 1$ *with probability 1. There exist universal constants* $\beta_{\mathrm{univ},1} \in \mathbb{R}_{>0}$ *and* $\beta_{\mathrm{univ},2} \in \mathbb{R}_{>0}$ *such that outside an event of probability at most* $N^{-\beta_{\mathrm{univ},1}}$ *times uniformly bounded factors, we have*

$$\|\bar{\mathfrak{P}}^N_{T,x}(\Psi^{N,2}\bar{\mathfrak{Z}}^N)\|_{\mathfrak{t}_{\mathrm{st}};\mathbb{T}_N} \lesssim N^{-\beta_{\mathrm{univ},2}} \langle \bar{\mathfrak{Z}}^N \rangle_{\mathfrak{t}_{\mathrm{st}};\mathbb{T}_N}. \tag{7.8}$$

**Lemma 7.5.** *Consider a random time* $\mathfrak{t}_{\mathrm{st}} \in \mathbb{R}_{\geqslant 0}$ *with* $\mathfrak{t}_{\mathrm{st}} \leqslant 1$ *with probability 1. There exist universal constants* $\beta_{\mathrm{univ},1} \in \mathbb{R}_{>0}$ *and* $\beta_{\mathrm{univ},2} \in \mathbb{R}_{>0}$ *such that outside an event of probability at most* $N^{-\beta_{\mathrm{univ},1}}$ *times uniformly bounded factors, we have*

$$\|\bar{\mathfrak{P}}^N_{T,x}(\widetilde{\Psi}^{N,2}\bar{\mathfrak{Z}}^N)\|_{\mathfrak{t}_{\mathrm{st}};\mathbb{T}_N} \lesssim N^{-\beta_{\mathrm{univ},2}} \langle \bar{\mathfrak{Z}}^N \rangle_{\mathfrak{t}_{\mathrm{st}};\mathbb{T}_N}. \tag{7.9}$$

The third ingredient is the following pair of estimates for time-averages on maximal time-scales $\mathfrak{t}_{\mathrm{av},\mathfrak{m}_+}, \mathfrak{t}^{\sim}_{\mathrm{av},\mathfrak{m}^{\sim}_+}$ from Corollary 5.12 and Corollary 5.13. This corresponds to the third bullet point from the list following Proposition 7.1. As alluded to therein, the following is a consequence of a scheme that avoids controlling time-averages on the aforementioned maximal time-scales $\mathfrak{t}_{\mathrm{av},\mathfrak{m}_+}, \mathfrak{t}^{\sim}_{\mathrm{av},\mathfrak{m}^{\sim}_+}$ directly and instead controls time-averages on smaller time-scales, glues these shorter-time estimates into longer-time estimates, and iterates until we arrive at the maximal time-scales from Corollary 5.12 and Corollary 5.13.

**Lemma 7.6.** *Consider a random time* $\mathfrak{t}_{\mathrm{st}} \in \mathbb{R}_{\geqslant 0}$ *with* $\mathfrak{t}_{\mathrm{st}} \leqslant 1$ *with probability 1. There exist universal constants* $\beta_{\mathrm{univ},1} \in \mathbb{R}_{>0}$ *and* $\beta_{\mathrm{univ},2} \in \mathbb{R}_{>0}$ *such that outside an event of probability at most* $N^{-\beta_{\mathrm{univ},1}}$ *times uniformly bounded factors, we have*

$$\|\bar{\mathfrak{P}}^N_{T,x}(N^{\frac{1}{2}}\mathscr{A}^{\mathbf{T},+}_{\mathfrak{t}_{\mathrm{av},\mathfrak{m}_+}} \mathscr{C}^{\mathbf{X},-}_{N^{\beta_X}}(\mathfrak{g}) \cdot \bar{\mathfrak{Z}}^N)\|_{\mathfrak{t}_{\mathrm{st}};\mathbb{T}_N} \lesssim N^{-\beta_{\mathrm{univ},2}} \langle \bar{\mathfrak{Z}}^N \rangle_{\mathfrak{t}_{\mathrm{st}};\mathbb{T}_N}. \tag{7.10}$$

*Above, the time-scale* $\mathfrak{t}_{\mathrm{av},\mathfrak{m}_+} \in \mathbb{R}_{>0}$ *is the "maximal" time-scale from* Corollary *5.12, which we emphasize is equal to the "maximal" time-scale* $\mathfrak{t}_{\mathrm{av},\mathfrak{l}_+} \in \mathbb{R}_{>0}$ *in* Corollary *5.14. Moreover, if instead we take the time-scales* $\widetilde{\mathfrak{t}}_{\mathrm{av},\mathfrak{m}_+^\sim}$, $\widetilde{\mathfrak{t}}_{\mathrm{av},\mathfrak{l}_+^\sim}$ *in* Corollary *5.13 and* Corollary *5.16, which are also equal to each other, we have the following with at least the same probability and same constant* $\beta_{\mathrm{univ},2} \in \mathbb{R}_{>0}$:

$$\|\bar{\mathfrak{P}}_{T,x}^N \left( N^{\beta_X} \left( \widetilde{\sum}_{\mathfrak{l}=1,\dots,N^{\beta_X}} \mathscr{A}_{\widetilde{\mathfrak{t}}_{\mathrm{av},\mathfrak{m}_+^\sim}}^{\mathbf{T},+} (\widetilde{\mathfrak{g}}^{\mathfrak{l}}) \right) \cdot \bar{\mathfrak{z}}^N \right) \|_{\mathfrak{t}_{\mathrm{st}};\mathbb{T}_N} \lesssim N^{-\beta_{\mathrm{univ},2}} \langle \bar{\mathfrak{z}}^N \rangle_{\mathfrak{t}_{\mathrm{st}};\mathbb{T}_N}. \tag{7.11}$$

*Proof of Proposition 7.1.* Throughout we will make statements which hold with high probability. Here high probability means that the complement event has probability at most $N^{-\beta_{\mathrm{univ},1}}$ times uniformly bounded constants, with $\beta_{\mathrm{univ},1} \in \mathbb{R}_{>0}$ universal. Because we make a uniformly bounded number of such statements, the conclusion holds with high-probability too.

By definition of $\bar{\Phi}^{N,2}$ in the stochastic equation for $\bar{\mathfrak{z}}^N$ in Definition 4.3 and the triangle inequality for $\|\|_{\mathfrak{t}_{\mathrm{st}};\mathbb{T}_N}$, we first have

$$\|\bar{\mathfrak{P}}^N(\bar{\Phi}^{N,2})\|_{\mathfrak{t}_{\mathrm{st}};\mathbb{T}_N} \leqslant \|\bar{\mathfrak{P}}_{T,x}^N(N^{\frac{1}{2}} \mathscr{A}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}) \cdot \bar{\mathfrak{z}}^N)\|_{\mathfrak{t}_{\mathrm{st}};\mathbb{T}_N}$$
$$+ \|\bar{\mathfrak{P}}_{T,x}^N \left( N^{\beta_X} \widetilde{\sum}_{\mathfrak{l}=1,\dots,N^{\beta_X}} \widetilde{\mathfrak{g}}^{\mathfrak{l}} \cdot \bar{\mathfrak{z}}^N \right) \|_{\mathfrak{t}_{\mathrm{st}};\mathbb{T}_N} + \Xi, \tag{7.12}$$

where the last quantity corresponds to gradient terms inside $\bar{\Phi}^{N,2}$; precisely, the $\Xi$ term is

$$\Xi \doteq \widetilde{\sum}_{\mathfrak{l}=1,\dots,N^{\beta_X}} N^{-\frac{1}{2}} \|\bar{\mathfrak{P}}_{T,x}^N(\bar{\nabla}_{-7\mathfrak{l}\mathfrak{m}}^{!}(\mathfrak{b}^{\mathfrak{l}}\bar{\mathfrak{z}}^N))\|_{\mathfrak{t}_{\mathrm{st}};\mathbb{T}_N}. \tag{7.13}$$

Recall $\mathfrak{m} \in \mathbb{Z}_{>0}$ is the maximal jump-length in the particle random walks in the model, and $\bar{\nabla}^{!} \doteq N\bar{\nabla}$ is the continuum-scaling of the discrete gradient acting on functions $\mathbb{T}_N \to \mathbb{R}$ on the torus. We denote the first and second terms on the RHS of (7.12) by $\Xi_1$ and $\Xi_2$, respectively. We adopt the notation in Lemmas 7.3, 7.4, and 7.6 with $\mathfrak{t}_{\mathrm{av},\mathfrak{m}_+}$ from Corollary 5.12/5.14 to first get

$$\Xi_1 \leqslant \|\bar{\mathfrak{P}}_{T,x}^N(\Psi^{N,1}\bar{\mathfrak{z}}^N)\|_{\mathfrak{t}_{\mathrm{st}};\mathbb{T}_N} + \|\bar{\mathfrak{P}}_{T,x}^N(\Psi^{N,2}\bar{\mathfrak{z}}^N)\|_{\mathfrak{t}_{\mathrm{st}};\mathbb{T}_N}$$
$$+ \|\bar{\mathfrak{P}}_{T,x}^N(N^{\frac{1}{2}} \mathscr{A}_{\mathfrak{t}_{\mathrm{av},\mathfrak{m}_+}}^{\mathbf{T},+} \mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}) \cdot \bar{\mathfrak{z}}^N)\|_{\mathfrak{t}_{\mathrm{st}};\mathbb{T}_N} \tag{7.14}$$

Combining the last bound with Lemmas 7.3, 7.4, and 7.6, with the required high probability we have $\Xi_1 \lesssim N^{-\beta_{\mathrm{univ},2}} \langle \bar{\mathfrak{z}}^N \rangle_{\mathfrak{t}_{\mathrm{st}};\mathbb{T}_N}$. We now study $\Xi_2$. Adopting the same notation and also that of Lemma 7.5 but with $\widetilde{\mathfrak{t}}_{\mathrm{av},\mathfrak{m}_+^\sim}$ from Corollary 5.13/5.16, we get

$$\Xi_2 \leqslant \|\bar{\mathfrak{P}}_{T,x}^N(\widetilde{\Psi}^{N,2}\bar{\mathfrak{z}}^N)\|_{\mathfrak{t}_{\mathrm{st}};\mathbb{T}_N} + \|\bar{\mathfrak{P}}_{T,x}^N \left( N^{\beta_X} \left( \widetilde{\sum}_{\mathfrak{l}=1,\dots,N^{\beta_X}} \mathscr{A}_{\widetilde{\mathfrak{t}}_{\mathrm{av},\mathfrak{m}_+^\sim}}^{\mathbf{T},+}(\widetilde{\mathfrak{g}}^{\mathfrak{l}}) \right) \cdot \bar{\mathfrak{z}}^N \right) \|_{\mathfrak{t}_{\mathrm{st}};\mathbb{T}_N}. \tag{7.15}$$

We now apply Lemma 7.5 and Lemma 7.6 to deduce $\Xi_2 \lesssim N^{-\beta_{\mathrm{univ},2}} \langle \bar{\mathfrak{z}}^N \rangle_{\mathfrak{t}_{\mathrm{st}};\mathbb{T}_N}$ also with the required high-probability. It remains to estimate $\Xi$. To this end, we estimate each term inside the averaged summation defining $\Xi$ given right after the first triangle inequality estimate (7.12). Estimating all of these terms amounts to analytic estimates in

Lemma A.1. In particular, we do not use randomness and we only require the uniform bound $|\mathfrak{b}^{\mathfrak{l}}| \lesssim 1$. To be precise, we employ the regularity estimate (A.7) to get

$$
\begin{aligned}
\Xi \doteq \widetilde{\sum}_{\mathfrak{l}=1}^{N^{\beta_X}} N^{-\frac{1}{2}} \|\bar{\mathfrak{P}}_{T,x}^N(\bar{\nabla}_{-7\mathfrak{l}\mathfrak{m}}^!(\mathfrak{b}^{\mathfrak{l}}\bar{\mathfrak{z}}^N))\|_{\mathfrak{t}_{st};\mathbb{T}_N} \\
\lesssim_{\mathfrak{m}} \widetilde{\sum}_{\mathfrak{l}=1}^{N^{\beta_X}} N^{-\frac{1}{2}} |\mathfrak{l}| \|\bar{\mathfrak{z}}^N\|_{\mathfrak{t}_{st};\mathbb{T}_N} \lesssim N^{-\frac{1}{2}+\beta_X} \|\bar{\mathfrak{z}}^N\|_{\mathfrak{t}_{st};\mathbb{T}_N} \\
\leqslant N^{-\frac{1}{6}+\varepsilon_{X,1}} \langle\bar{\mathfrak{z}}^N\rangle_{\mathfrak{t}_{st};\mathbb{T}_N}.
\end{aligned}
\tag{7.16}
$$

Indeed, the far RHS of (7.16) follows by the middle upon recalling $\beta_X = \frac{1}{3} + \varepsilon_{X,1}$ and applying the bracket inequality (7.2). We now combine (7.16) with (7.12) and the previous $\Xi_1 + \Xi_2 \lesssim N^{-\beta_{\mathrm{univ},2}} \langle\bar{\mathfrak{z}}^N\rangle_{\mathfrak{t}_{st};\mathbb{T}_N}$ for the first two terms on the RHS of (7.12). □

We now prove ingredients used in the proof of Proposition 7.1. For all of Lemmas 7.3, 7.4, 7.5, and 7.6, we give ingredients for their proofs, combine them, and then defer their technical proofs to the end of the section to avoid obscuring key ideas.

### 7.2. Proof of Lemma 7.3.

The proof of replacement-by-cutoff we give here provides a good template for later arguments in this section. As $\mathfrak{t}_{st} \leqslant 1$ with probability 1, for any $T \leqslant \mathfrak{t}_{st} \leqslant 1$ we have $\|\|_{\mathfrak{t}_{st};\mathbb{T}_N} \leqslant \|\|_{1;\mathbb{T}_N}$ as norms, and the deterministic bound

$$
\begin{aligned}
\|\bar{\mathfrak{P}}_{T,x}^N(\Psi^{N,1}\bar{\mathfrak{z}}^N)\|_{\mathfrak{t}_{st};\mathbb{T}_N} &\leqslant \|\bar{\mathfrak{z}}^N\|_{\mathfrak{t}_{st};\mathbb{T}_N} \|\bar{\mathfrak{P}}_{T,x}^N(N^{\frac{1}{2}}|\mathscr{A}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}) - \mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g})|)\|_{1;\mathbb{T}_N} \\
&\leqslant \langle\bar{\mathfrak{z}}^N\rangle_{\mathfrak{t}_{st};\mathbb{T}_N} \|\bar{\mathfrak{P}}_{T,x}^N(N^{\frac{1}{2}}|\mathscr{A}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}) - \mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g})|)\|_{1;\mathbb{T}_N}.
\end{aligned}
$$

The second bound follows from applying (7.2) to the $\bar{\mathfrak{z}}^N$-norm in the first bound. By Proposition 5.2 and the Markov inequality, the $\|\|_{1;\mathbb{T}_N}$-norm in this last estimate is at most $N^{-\beta_{\mathrm{univ},2}}$ times uniformly bounded factors with the required high probability for $\beta_{\mathrm{univ},2} \in \mathbb{R}_{>0}$ universal. Precisely, by the Markov inequality and expectation bound in Proposition 5.2, we deduce

$$
\begin{aligned}
\mathbf{P}&\left(\|\bar{\mathfrak{P}}_{T,x}^N(N^{\frac{1}{2}}|\mathscr{A}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}) - \mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g})|)\|_{1;\mathbb{T}_N} \geqslant N^{-\beta_{\mathrm{univ},2}}\right) \\
&\leqslant N^{\beta_{\mathrm{univ},2}} \mathbf{E}\|\bar{\mathfrak{P}}_{T,x}^N(N^{\frac{1}{2}}|\mathscr{A}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}) - \mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g})|)\|_{1;\mathbb{T}_N} \\
&\lesssim N^{-\beta_{\mathrm{univ}}+\beta_{\mathrm{univ},2}}.
\end{aligned}
$$

We emphasize that $\beta_{\mathrm{univ}} \in \mathbb{R}_{>0}$ is universal and $\beta_{\mathrm{univ},2} \in \mathbb{R}_{>0}$ is a universal constant of our choosing. Choosing $\beta_{\mathrm{univ},2} = \frac{1}{2}\beta_{\mathrm{univ}}$ guarantees the event in the probability on the LHS fails with the required high-probability. Outside such an event, the proposed estimate in Lemma 7.3 holds courtesy of the first display in this proof. This completes the proof of Lemma 7.3. □

### 7.3. Proof of Lemma 7.4.

The first ingredient towards the proof for Lemma 7.4 is the following inductive estimate that we will iterate to prove Lemma 7.4. This identifies a cost of replacing a time-average in the heat operator by the same time-average but on a slightly larger time-scale. This cost is ultimately controlled by time-regularity of the heat kernel $\bar{\mathfrak{P}}^N$ and of $\bar{\mathfrak{z}}^N$, both on the larger of the two time-scales that we time-average with respect to, times the time-average on the shorter time-scale.

**Lemma 7.7.** *Consider any possibly random time* $t_{st} \in \mathbb{R}_{\geqslant 0}$ *satisfying* $t_{st} \leqslant 1$ *with probability 1. Consider any times* $t_1, t_2 \in \mathbb{R}_{\geqslant 0}$ *satisfying* $N^{-100} \lesssim t_1 \leqslant t_2 \lesssim N^{-1}$ *and* $t_2 t_1^{-1} \in \mathbb{Z}_{\geqslant 0}$. *There exist universal constants* $\beta_{univ,1}, \beta_{univ,2} \in \mathbb{R}_{>0}$ *such that outside an event of probability at most* $N^{-\beta_{univ,1}}$ *times uniformly bounded factors, we have the following estimate for any* $\varepsilon \in \mathbb{R}_{>0}$:

$$\|\bar{\mathfrak{P}}_{T,x}^N(\Psi^{N,2,t_1,t_2}\bar{\mathfrak{Z}}^N)\|_{t_{st};\mathbb{T}_N} \lesssim_\varepsilon N^{-\beta_{univ,2}}\langle\bar{\mathfrak{Z}}^N\rangle_{t_{st};\mathbb{T}_N}$$
$$+ \langle\bar{\mathfrak{Z}}^N\rangle_{t_{st};\mathbb{T}_N}N^\varepsilon t_2^{\frac{1}{4}}\|\bar{\mathfrak{P}}_{T,x}^N(N^{\frac{1}{2}}|\mathscr{A}_{t_1}^{\mathbf{T},+}\mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g})|)\|_{1;\mathbb{T}_N}. \quad (7.17)$$

*Above we recalled the space-time average of pseudo-gradients in* Definition *5.3, and we also introduced the following quantity:*

$$\Psi^{N,2,t_1,t_2} \overset{\bullet}{=} N^{\frac{1}{2}}\mathscr{A}_{t_2}^{\mathbf{T},+}\mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}) - N^{\frac{1}{2}}\mathscr{A}_{t_1}^{\mathbf{T},+}\mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}). \quad (7.18)$$

We will require another version of Lemma 7.7 but when the initial time-scale is $t_1 = 0$, so the following result estimates the error in introducing a time-average in the first place that will kickstart the inductive procedure referenced at the beginning of this subsection that is based on Lemma 7.7. The proof of this upcoming estimate first replaces a time-0 time-average by a time-average on a very sub-microscopic time-scale. Because in expectation we expect nothing to happen on a very sub-microscopic time-scale, we directly get a little "regularity" of the local functional we are time-averaging for this initial replacement-by-time-average. In particular, we do not use regularity of the heat kernel or $\bar{\mathfrak{Z}}^N$ for the initial step. After, we use Lemma 7.7.

**Lemma 7.8.** *Assume the setting and notation in* Lemma 7.7 *but for* $t_1 = 0$ *and* $t_2 = t_{av,0}$ *in* Corollary *5.14. There are universal constants* $\beta_{univ,1}, \beta_{univ,2} \in \mathbb{R}_{>0}$ *such that outside an event of probability at most* $N^{-\beta_{univ,1}}$ *times uniformly bounded factors, we have*

$$\|\bar{\mathfrak{P}}_{T,x}^N(\Psi^{N,2,0,t_{av,0}}\bar{\mathfrak{Z}}^N)\|_{t_{st};\mathbb{T}_N} \lesssim N^{-\beta_{univ,2}}\langle\bar{\mathfrak{Z}}^N\rangle_{t_{st};\mathbb{T}_N}. \quad (7.19)$$

*Proof of Lemma 7.4.* We make the same note about "high-probability" as in the beginning of the proof of Proposition 7.1.

Recall the times $t_{av,\ell} \in \mathbb{R}_{>0}$ in Corollary 5.14, including $t_{av,0} = \mathscr{O}(N^{-2+\varepsilon_{X,2}})$. We will iterate the estimate (7.17) in Lemma 7.7 to replace the time-average on time-scale $t_{av,\ell}$ by the time-average on time-scale $t_{av,\ell+1}$ until we reach the maximal time-scale $t_{av,\mathfrak{l}_+} \in \mathbb{R}_{>0}$. We start this off by replacing the time-0 average, namely the identity operator, by a time-$t_{av,0}$ average. Estimating the errors in this first step of introducing a time-$t_{av,0}$-average and the subsequent "multiscale" steps then completes the proof.

We now make the previous paragraph precise. Let us first observe the following telescoping sum that follows by definition. It replaces the time-0 average with a time-$t_{av,\mathfrak{l}_+}$ average via intermediate steps. Recall $\Psi^{N,2}$ defined before/for Lemma 7.4:

$$\bar{\mathfrak{P}}_{T,x}^N(\Psi^{N,2}\bar{\mathfrak{Z}}^N) = \bar{\mathfrak{P}}_{T,x}^N(\Psi^{N,2,0,t_{av,\mathfrak{l}_+}}\bar{\mathfrak{Z}}^N)$$
$$= \bar{\mathfrak{P}}_{T,x}^N(\Psi^{N,2,0,t_{av,0}}\bar{\mathfrak{Z}}^N) + \sum_{\ell=0}^{\mathfrak{l}_+-1}\bar{\mathfrak{P}}_{T,x}^N(\Psi^{N,2,t_{av,\ell},t_{av,\ell+1}}\bar{\mathfrak{Z}}^N). \quad (7.20)$$

We first address the first term on the RHS of (7.20). By Lemma 7.8 with high-probability we get the estimate:

$$\|\bar{\mathfrak{P}}_{T,x}^{N}(\Psi^{N,2,0,\mathsf{t}_{\mathrm{av},0}}\bar{\mathfrak{Z}}^{N})\|_{\mathsf{t}_{\mathrm{st}};\mathbb{T}_{N}} \lesssim N^{-\beta_{\mathrm{univ},2}}\langle\bar{\mathfrak{Z}}^{N}\rangle_{\mathsf{t}_{\mathrm{st}};\mathbb{T}_{N}}. \tag{7.21}$$

For the second term on the RHS of (7.20), we use the triangle inequality for $\|\,\|_{\mathsf{t}_{\mathrm{st}};\mathbb{T}_{N}}$ and the uniform boundedness for $\mathfrak{l}_{+} \in \mathbb{Z}_{\geq 0}$ from Corollary 5.14 to get the following which controls a sum by the number of terms in the sum times a supremum:

$$\|\textstyle\sum_{\ell=0}^{\mathfrak{l}_{+}-1}\bar{\mathfrak{P}}_{T,x}^{N}(\Psi^{N,2,\mathsf{t}_{\mathrm{av},\ell},\mathsf{t}_{\mathrm{av},\ell+1}}\bar{\mathfrak{Z}}^{N})\|_{\mathsf{t}_{\mathrm{st}};\mathbb{T}_{N}} \lesssim \sup_{\ell=0,\ldots,\mathfrak{l}_{+}-1} \|\bar{\mathfrak{P}}_{T,x}^{N}(\Psi^{N,2,\mathsf{t}_{\mathrm{av},\ell},\mathsf{t}_{\mathrm{av},\ell+1}}\bar{\mathfrak{Z}}^{N})\|_{\mathsf{t}_{\mathrm{st}};\mathbb{T}_{N}}. \tag{7.22}$$

By (7.17) in Lemma 7.7, we estimate each term in the supremum on the RHS of (7.22) with high probability *uniformly* in $\ell$:

$$\|\bar{\mathfrak{P}}_{T,x}^{N}(\Psi^{N,2,\mathsf{t}_{\mathrm{av},\ell},\mathsf{t}_{\mathrm{av},\ell+1}}\bar{\mathfrak{Z}}^{N})\|_{\mathsf{t}_{\mathrm{st}};\mathbb{T}_{N}} \lesssim N^{-\beta_{\mathrm{univ},2}}\langle\bar{\mathfrak{Z}}^{N}\rangle_{\mathsf{t}_{\mathrm{st}};\mathbb{T}_{N}}$$
$$+ \langle\bar{\mathfrak{Z}}^{N}\rangle_{\mathsf{t}_{\mathrm{st}};\mathbb{T}_{N}} N^{\varepsilon}\mathsf{t}_{\mathrm{av},\ell+1}^{1/4}\|\bar{\mathfrak{P}}_{T,x}^{N}(N^{\frac{1}{2}}|\mathscr{A}_{\mathsf{t}_{\mathrm{av},\ell}}^{\mathbf{T},+}\mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g})|)\|_{1;\mathbb{T}_{N}}. \tag{7.23}$$

We may now apply the expectation estimate in Corollary 5.14 along with the Markov inequality to deduce the following bound with high-probability for the second term on the RHS of (7.23) in which $\beta_{\mathrm{univ},2} \in \mathbb{R}_{>0}$ is possibly updated from (7.21)/(7.23):

$$\sup_{\ell=0,\ldots,\mathfrak{l}_{+}-1} N^{\varepsilon}\mathsf{t}_{\mathrm{av},\ell+1}^{1/4}\|\bar{\mathfrak{P}}_{T,x}^{N}(N^{\frac{1}{2}}|\mathscr{A}_{\mathsf{t}_{\mathrm{av},\ell}}^{\mathbf{T},+}\mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g})|)\|_{1;\mathbb{T}_{N}} \lesssim N^{-\beta_{\mathrm{univ},2}}. \tag{7.24}$$

We now combine (7.22), (7.23), and (7.24) to get the following high-probability bound:

$$\|\textstyle\sum_{\ell=0}^{\mathfrak{l}_{+}-1}\bar{\mathfrak{P}}_{T,x}^{N}(\Psi^{N,2,\mathsf{t}_{\mathrm{av},\ell},\mathsf{t}_{\mathrm{av},\ell+1}}\bar{\mathfrak{Z}}^{N})\|_{\mathsf{t}_{\mathrm{st}};\mathbb{T}_{N}} \lesssim N^{-\beta_{\mathrm{univ},2}}\langle\bar{\mathfrak{Z}}^{N}\rangle_{\mathsf{t}_{\mathrm{st}};\mathbb{T}_{N}}. \tag{7.25}$$

Combining (7.20), (7.21), and (7.25) with the triangle inequality for $\|\,\|_{\mathsf{t}_{\mathrm{st}};\mathbb{T}_{N}}$ completes the proof.                                                                                            □

*7.4. Proof of Lemma 7.5.* The proof of Lemma 7.5 is near identical to the proof of Lemma 7.4. We just make the replacement (7.3) and adjust power-counting in $N \in \mathbb{Z}_{>0}$ in a fashion that is ultimately negligible. This starts with an analog of Lemma 7.7.

**Lemma 7.9.** *Retain the notation of* Lemma 7.7. *There exist universal constants* $\beta_{\mathrm{univ},1}$, $\beta_{\mathrm{univ},2} \in \mathbb{R}_{>0}$ *such that outside an event of probability at most* $N^{-\beta_{\mathrm{univ},1}}$ *times uniformly bounded factors, we have the following estimate for any* $\varepsilon \in \mathbb{R}_{>0}$:

$$\|\bar{\mathfrak{P}}_{T,x}^{N}(\widetilde{\Psi}^{N,2,\mathsf{t}_1,\mathsf{t}_2}\bar{\mathfrak{Z}}^{N})\|_{\mathsf{t}_{\mathrm{st}};\mathbb{T}_{N}} \lesssim_{\varepsilon} N^{-\beta_{\mathrm{univ},2}}\langle\bar{\mathfrak{Z}}^{N}\rangle_{\mathsf{t}_{\mathrm{st}};\mathbb{T}_{N}}$$
$$+ \langle\bar{\mathfrak{Z}}^{N}\rangle_{\mathsf{t}_{\mathrm{st}};\mathbb{T}_{N}} N^{\varepsilon}\mathsf{t}_{2}^{\frac{1}{4}}\left(\widetilde{\textstyle\sum}_{\mathfrak{l}=1}^{N^{\beta_X}}\|\bar{\mathfrak{P}}_{T,x}^{N}(N^{\beta_X}|\mathscr{A}_{\mathsf{t}_1}^{\mathbf{T},+}(\widetilde{\mathfrak{g}}^{\mathfrak{l}})|)\|_{1;\mathbb{T}_{N}}\right). \tag{7.26}$$

*We recall the time-average introduced in* Definition 5.8 *of the functionals* $\mathfrak{g}^{\mathfrak{l}}$ *with pseudo-gradient factor. Above we also introduced*

$$\widetilde{\Psi}^{N,2,\mathsf{t}_1,\mathsf{t}_2} \stackrel{\bullet}{=} N^{\beta_X}\widetilde{\textstyle\sum}_{\mathfrak{l}=1,\ldots,N^{\beta_X}}\left(\mathscr{A}_{\mathsf{t}_2}^{\mathbf{T},+}(\widetilde{\mathfrak{g}}^{\mathfrak{l}}) - \mathscr{A}_{\mathsf{t}_1}^{\mathbf{T},+}(\widetilde{\mathfrak{g}}^{\mathfrak{l}})\right). \tag{7.27}$$

We emphasize that the only difference between this and Lemma 7.7 is the replacement (7.3). In particular, as we suggested at the beginning of this subsection its proof is similar to that of Lemma 7.7. The details of what we time-average in Lemma 7.7 and Lemma 7.9 are unimportant, just that we are replacing a time-average by another on a larger time-scale.

The second preliminary ingredient we will need towards the proof of Lemma 7.5 is another parallel to a step in the proof for Lemma 7.4. This second ingredient is an analog of Lemma 7.8. It serves to introduce a preliminary time-average that kicks off the inductive procedure driven by the previous estimate in Lemma 7.9. This is similar to the purpose of Lemma 7.8.

**Lemma 7.10.** *Assume the setting and notation in* Lemma *7.7 but for* $t_1 = 0$ *and* $t_2 = t_{\widetilde{av},0}$ *in* Corollary *5.16. There are universal constants* $\beta_{univ,1}, \beta_{univ,2} \in \mathbb{R}_{>0}$ *such that outside an event of probability at most* $N^{-\beta_{univ,1}}$ *times uniformly bounded factors, we have*

$$\|\bar{\mathfrak{P}}_{T,x}^{N}(\widetilde{\Psi}^{N,2,0,t_{\widetilde{av},0}}\bar{\mathfrak{Z}}^{N})\|_{t_{st};\mathbb{T}_N} \lesssim N^{-\beta_{univ,2}}\langle\bar{\mathfrak{Z}}^{N}\rangle_{t_{st};\mathbb{T}_N}. \tag{7.28}$$

*Proof of Lemma 7.5.* It suffices to follow the proof of Lemma 7.4 verbatim, make the replacement (7.3), use the bound (7.26) in place of (7.17), use time-scales in Corollary 5.16 in place of those in Corollary 5.14, and use the estimate in Corollary 5.16 in place of that in Corollary 5.14. This also means using Lemma 7.9/7.10 instead of Lemma 7.7/7.8. □

## 7.5. Proof of Lemma 7.6.

Recall that we have replaced the spatial average of the pseudo-gradient in $\bar{\Phi}^{N,2}$ with its CLT-type cutoff in Lemma 7.3, and that we have replaced this spatial-average-with-cutoff with a dynamic average on a mesoscopic time scale in Lemma 7.4. In Lemma 7.5, we replaced the functionals $\widetilde{\mathfrak{g}}^{\mathfrak{l}}$ with pseudo-gradient factors with mesoscopic time-averages as well. For Lemma 7.6, we introduce the following inductive scheme transferring a priori estimates, that are of main interest in Propositions 5.6 and 5.11, and Corollaries 5.12 and 5.13, between time-averages on progressively larger time-scales. This inductive mechanism, again, is a precise implementation of the would-be-proof of Pseudo-Proposition 2.14.

**Lemma 7.11.** *Recall the sequences of time-scales* $\{t_{av,m}\}_{m=1}^{m_+}$ *and of exponents* $\{\beta_m\}_{m=0}^{m_+}$ *within* Corollary *5.12 and the space-time averages with cutoffs in* Definition *5.5. Provided any* $m \in \mathbb{Z}_{\geqslant 1}$, *we define the following double averaging which groups scale-$t_{av,m}$ time-averages with a priori upper bounds into blocks of time-scale $t_{av,m+1}$, and these blocks, composed of scale-$t_{av,m}$ time-averages-with-upper-bound-cutoff, are grouped into the "maximal" time-scale $t_{av,m_+} \in \mathbb{R}_{>0}$. Precisely, we define the following double average in which scale-$t_{av,m}$ time-averages with the upper-bound-cutoff of order $N^{-\beta_m}$ are defined afterwards/in the second line below:*

$$\Gamma_{T,x}^{N,m,m+1} \overset{\bullet}{=} \sum_{\mathfrak{l}=0}^{t_{av,m_+}t_{av,m+1}^{-1}-1} \Gamma_{T+\mathfrak{l}t_{av,m+1},x}^{N,t_{av,m}} \overset{\bullet}{=} \sum_{\mathfrak{l}=0}^{t_{av,m_+}t_{av,m+1}^{-1}-1}\sum_{j=0}^{t_{av,m+1}t_{av,m}^{-1}-1} \Gamma_{T+\mathfrak{l}t_{av,m+1},x}^{N,t_{av,m},j} \tag{7.29}$$

$$\Gamma_{T,x}^{N,t_{av,m},j} \overset{\bullet}{=} \mathscr{A}_{t_{av,m}}^{\mathbf{T},+}\mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{T+jt_{av,m},x}) \cdot \mathbf{1}\left(\sup_{0\leqslant t\leqslant t_{av,m}} t \cdot t_{av,m}^{-1}|\mathscr{A}_{t}^{\mathbf{T},+}\mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{T+jt_{av,m},x})| \leqslant N^{-\beta_m}\right). \tag{7.30}$$

*Take any random time* $\mathfrak{t}_{st} \in \mathbb{R}_{\geqslant 0}$ *satisfying* $\mathfrak{t}_{st} \leqslant 1$ *with probability 1. There exists a pair of universal constants* $\beta_{univ,1}, \beta_{univ,2} \in \mathbb{R}_{>0}$ *so that outside an event of probability at most* $N^{-\beta_{univ,1}}$ *times uniformly bounded constants, simultaneously in* $m \in [\![1, \mathfrak{m}_+ - 2]\!]$ *we have the following which replaces an average of scale-*$\mathfrak{t}_{av,m}$ *time-averages with upper bound cutoffs with an average of scale-*$\mathfrak{t}_{av,m+1}$ *time-averages with slightly improved upper bound cutoffs:*

$$\|\bar{\mathfrak{P}}_{T,x}^N(N^{\frac{1}{2}}\Gamma^{N,m,m+1}\bar{\mathfrak{z}}^N)\|_{\mathfrak{t}_{st};\mathbb{T}_N} \lesssim \|\bar{\mathfrak{P}}_{T,x}^N(N^{\frac{1}{2}}\Gamma^{N,m+1,m+2}\bar{\mathfrak{z}}^N)\|_{\mathfrak{t}_{st};\mathbb{T}_N}$$
$$+ N^{-\beta_{univ,2}}\|\bar{\mathfrak{z}}^N\|_{\mathfrak{t}_{st};\mathbb{T}_N}. \tag{7.31}$$

*The same is true after replacement* (7.3) *if we use time-scales and exponents in* Corollary *5.13 instead of those in* Corollary *5.12.*

Lemma 7.11 is difficult to justify with words; it roughly amounts to comparing $\Gamma^{N,m,m+1}$ and $\Gamma^{N,m+1,m+2}$ with bootstrapping and improving a priori bounds on time-scale $\mathfrak{t}_{av,m}$ into a priori bounds on time-scale $\mathfrak{t}_{av,m+1}$. One important observation for the proof of Lemma 7.11 is that time-averages on larger time-scales are averages of time-averages on shorter time-scales. This lets us transfer between time-scales in the proof for Lemma 7.11. We emphasize that error terms we get when we compare $\Gamma^{N,m,m+1}$ and $\Gamma^{N,m+1,m+2}$ and when we bootstrap a priori bounds to larger time-scales can be controlled by Corollaries 5.12 and 5.13.

We proceed with another preparatory lemma which starts the inductive procedure that Lemma 7.11 propagates.

**Lemma 7.12.** *We recap setting of* Lemma *7.11. Recall the sequences* $\{\mathfrak{t}_{av,m}\}_{m=1}^{\mathfrak{m}_+}$ *and* $\{\beta_m\}_{m=0}^{\mathfrak{m}_+}$ *in Corollary 5.12, and define*

$$\Gamma_{T,x}^{N,1,2} \doteq \sum_{\mathfrak{l}=0}^{\mathfrak{t}_{av,\mathfrak{m}_+}\mathfrak{t}_{av,2}^{-1}-1} \Gamma_{T+\mathfrak{l}\mathfrak{t}_{av,2},x}^{N,\mathfrak{t}_{av,1}} \doteq \sum_{\mathfrak{l}=0}^{\mathfrak{t}_{av,\mathfrak{m}_+}\mathfrak{t}_{av,2}^{-1}-1}\sum_{\mathfrak{j}=0}^{\mathfrak{t}_{av,2}\mathfrak{t}_{av,1}^{-1}-1} \Gamma_{T+\mathfrak{l}\mathfrak{t}_{av,2},x}^{N,\mathfrak{t}_{av,1},\mathfrak{j}} \tag{7.32}$$

$$\Gamma_{T,x}^{N,\mathfrak{t}_{av,1},\mathfrak{j}} \doteq \mathscr{A}_{\mathfrak{t}_{av,1}}^{\mathbf{T},+}\mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{T+\mathfrak{j}\mathfrak{t}_{av,1},x}) \cdot \mathbf{1}\left(\sup_{0 \leqslant \mathfrak{t} \leqslant \mathfrak{t}_{av,1}} \mathfrak{t} \cdot \mathfrak{t}_{av,1}^{-1}|\mathscr{A}_{\mathfrak{t}}^{\mathbf{T},+}\mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{T+\mathfrak{j}\mathfrak{t}_{av,1},x})| \leqslant N^{-\beta_1}\right). \tag{7.33}$$

*Consider any random time* $\mathfrak{t}_{st} \in \mathbb{R}_{\geqslant 0}$ *satisfying* $\mathfrak{t}_{st} \leqslant 1$ *with probability 1. There exist universal constants* $\beta_{univ,1}, \beta_{univ,2} \in \mathbb{R}_{>0}$ *such that outside an event of probability at most* $N^{-\beta_{univ,1}}$ *times uniformly bounded constants, we have*

$$\|\bar{\mathfrak{P}}_{T,x}^N(N^{\frac{1}{2}}\mathscr{A}_{\mathfrak{t}_{av,\mathfrak{m}_+}}^{\mathbf{T},+}\mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}) \cdot \bar{\mathfrak{z}}^N)\|_{\mathfrak{t}_{st};\mathbb{T}_N} \lesssim \|\bar{\mathfrak{P}}_{T,x}^N(N^{\frac{1}{2}}\Gamma^{N,1,2}\bar{\mathfrak{z}}^N)\|_{\mathfrak{t}_{st};\mathbb{T}_N}$$
$$+ N^{-\beta_{univ,2}}\|\bar{\mathfrak{z}}^N\|_{\mathfrak{t}_{st};\mathbb{T}_N}. \tag{7.34}$$

*The same is true after replacement* (7.3) *if we use time-scales and exponents in* Corollary *5.13 instead of those in* Corollary *5.12.*

*Proof of Lemma 7.6.* We first make the same disclaimer concerning high-probability statements and events as we made at the beginning of the proof of Proposition 7.1. We also focus first on showing the first estimate (7.10) from the statement of Lemma 7.6. We use the first bound for order $N^{1/2}$ quantities in Lemma 7.12 and iterate the first estimate for order $N^{1/2}$ terms in Lemma 7.11. As $\mathfrak{m}_+ \lesssim 1$ by Corollary 5.12, we iterate Lemma 7.11 a uniformly bounded number of times, so with high probability

$$\|\bar{\mathfrak{P}}_{T,x}^N(N^{\frac{1}{2}}\mathscr{A}_{\mathfrak{t}_{av,\mathfrak{m}_+}}^{\mathbf{T},+}\mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}) \cdot \bar{\mathfrak{z}}^N)\|_{\mathfrak{t}_{st};\mathbb{T}_N} \lesssim \|\bar{\mathfrak{P}}_{T,x}^N(N^{\frac{1}{2}}\Gamma^{N,\mathfrak{m}_+-1,\mathfrak{m}_+}\bar{\mathfrak{z}}^N)\|_{\mathfrak{t}_{st};\mathbb{T}_N}$$
$$+ N^{-\beta_{univ,2}}\|\bar{\mathfrak{z}}^N\|_{\mathfrak{t}_{st};\mathbb{T}_N}. \tag{7.35}$$

By definition in Lemma 7.11 $\Gamma^{N,\mathfrak{m}_+-1,\mathfrak{m}_+}$ is an average of terms that are at most $N^{-\beta_{\mathfrak{m}_+-1}}$. As $\beta_{\mathfrak{m}_+-1} = \beta_{\mathfrak{m}_+} - \varepsilon_{X,3} \geqslant \frac{1}{2} + 2\varepsilon_{X,3}$, by the space-time convolution estimate (A.3), we control the first term on the RHS of (7.35) directly/deterministically and get

$$\|\bar{\mathfrak{P}}_{T,x}^N(N^{\frac{1}{2}}\Gamma^{N,\mathfrak{m}_+-1,\mathfrak{m}_+}\bar{\mathfrak{Z}}^N)\|_{\mathrm{t_{st};T_N}} \lesssim N^{\frac{1}{2}}\|\Gamma^{N,\mathfrak{m}_+-1,\mathfrak{m}_+}\bar{\mathfrak{Z}}^N\|_{\mathrm{t_{st};T_N}}$$
$$\lesssim N^{\frac{1}{2}-\beta_{\mathfrak{m}_+-1}}\|\bar{\mathfrak{Z}}^N\|_{\mathrm{t_{st};T_N}} \leqslant N^{-2\varepsilon_{X,3}}\langle\bar{\mathfrak{Z}}^N\rangle_{\mathrm{t_{st};T_N}}. \tag{7.36}$$

Combining (7.35) and (7.36) yields the first estimate (7.10). The bound (7.11) follows via the same argument except that for our application of Lemmas 7.11 and 7.12, and also for (7.35) and (7.36) above, we make the replacement (7.3), and we adopt the time-scales $\{\widetilde{\mathrm{t}}_{\mathrm{av},m}\}_{m=1}^{\widetilde{\mathfrak{m}}_+}$, exponents $\{\widetilde{\beta_m}\}_{m=0}^{\widetilde{\mathfrak{m}}_+}$, and estimates in Corollary 5.13 instead of in Corollary 5.12. $\qquad\square$

*7.6. Proofs of Technical Estimates.* We now give proofs of all technical ingredients in the order that they are presented.

*Proof of Lemma 7.7.* Recall the notation $\Psi^{N,2,\mathrm{t_1},\mathrm{t_2}}$ in Lemma 7.7. The architecture of our proof will be computing explicitly the heat-operator term inside the $\|\|\|_{\mathrm{t_{st};T_N}}$-norm on the LHS of the proposed estimate and then estimating each term in our resulting calculations. This starts with the following elementary consideration in which we de-compose a scale-$\mathrm{t_2}$ time-average in terms of time-shifted scale-$\mathrm{t_1}$ time-averages and then we remove these time-shifts to get the $\Phi_{\mathrm{j}}$-errors below. We clarify this after:

$$\bar{\mathfrak{P}}_{T,x}^N(N^{\frac{1}{2}}\mathscr{A}_{\mathrm{t_2}}^{\mathbf{T},+}\mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{S,y})\cdot\bar{\mathfrak{Z}}_{S,y}^N) = \bar{\mathfrak{P}}_{T,x}^N\left(N^{\frac{1}{2}}\widetilde{\sum_{\mathrm{j}=0}^{\mathrm{t_2}\mathrm{t_1}^{-1}-1}}\mathscr{A}_{\mathrm{t_1}}^{\mathbf{T},+}\mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{S+\mathrm{j}\mathrm{t_1},y})\cdot\bar{\mathfrak{Z}}_{S,y}^N\right) \tag{7.37}$$

$$= \bar{\mathfrak{P}}_{T,x}^N(N^{\frac{1}{2}}\mathscr{A}_{\mathrm{t_1}}^{\mathbf{T},+}\mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{S,y})\cdot\bar{\mathfrak{Z}}_{S,y}^N) + \widetilde{\sum_{\mathrm{j}=0}^{\mathrm{t_2}\mathrm{t_1}^{-1}-1}}\Phi_{\mathrm{j}}. \tag{7.38}$$

The first line follows from the observation that a time-average with respect to the time-scale $\mathrm{t_2}$ is an average of suitably shifted time-averages with respect to the time-scale $\mathrm{t_1}$. We recall from the statement of Lemma 7.7 that $\mathrm{t_2}$ is a positive integer multiple of $\mathrm{t_1}$, so that the sum and its limits of indices are all well-defined. For example, the previous decomposition yielding (7.38) is a relative of the fact that an average of 10 terms may be written as the average of 5 other terms, each of which is the average of 2 "adjacent" terms. We have also introduced the following $\Phi_{\mathrm{j}}$-errors obtained after removing the $\mathrm{j}\mathrm{t_1}$-time shift to get (7.38):

$$\Phi_{\mathrm{j}} \overset{\bullet}{=} \bar{\mathfrak{P}}_{T,x}^N\left(N^{\frac{1}{2}}\mathscr{A}_{\mathrm{t_1}}^{\mathbf{T},+}\mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{S+\mathrm{j}\mathrm{t_1},y})\cdot\bar{\mathfrak{Z}}_{S,y}^N\right)$$
$$- \bar{\mathfrak{P}}_{T,x}^N\left(N^{\frac{1}{2}}\mathscr{A}_{\mathrm{t_1}}^{\mathbf{T},+}\mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{S,y})\cdot\bar{\mathfrak{Z}}_{S,y}^N\right)$$
$$= \bar{\mathfrak{P}}_{T,x}^N\left(N^{\frac{1}{2}}\nabla_{\mathrm{j}\mathrm{t_1}}^{\mathbf{T}}\left(\mathscr{A}_{\mathrm{t_1}}^{\mathbf{T},+}\mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{S,y})\right)\bar{\mathfrak{Z}}_{S,y}^N\right).$$

From (7.38), we subtract the first term on the RHS from both sides. By definition of $\Psi^{N,2,t_1,t_2}\bar{\mathfrak{Z}}^N$, we get the following formula:

$$
\begin{aligned}
&\bar{\mathfrak{P}}_{T,x}^N (\Psi^{N,2,t_1,t_2}\bar{\mathfrak{Z}}^N) \\
&\quad = \bar{\mathfrak{P}}_{T,x}^N (N^{\frac{1}{2}} \mathscr{A}_{t_2}^{\mathbf{T},+} \mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{S,y}) \cdot \bar{\mathfrak{Z}}_{S,y}^N) - \bar{\mathfrak{P}}_{T,x}^N (N^{\frac{1}{2}} \mathscr{A}_{t_1}^{\mathbf{T},+} \mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{S,y}) \cdot \bar{\mathfrak{Z}}_{S,y}^N) \\
&\quad = \widetilde{\sum}_{j=0}^{t_2 t_1^{-1}-1} \Phi_j .
\end{aligned}
\tag{7.39}
$$

We proceed to rewrite $\Phi_j$ to transform the time-gradient of the space-time average defining $\Phi_j$-terms into time-gradients of the heat operator and the microscopic Cole-Hopf transform. We will achieve this by a time-discrete-version of integration-by-parts for the time-integral in the heat operator $\bar{\mathfrak{P}}^N$. To make this precise, we first use the following time-discrete-type Leibniz rule:

$$
\begin{aligned}
&\nabla_{jt_1}^{\mathbf{T}} \left( \mathscr{A}_{t_1}^{\mathbf{T},+} \mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{S,y}) \right) \bar{\mathfrak{Z}}_{S,y}^N \\
&\quad = \nabla_{jt_1}^{\mathbf{T}} \left( \mathscr{A}_{t_1}^{\mathbf{T},+} \mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{S,y}) \bar{\mathfrak{Z}}_{S-jt_1,y}^N \right) + \mathscr{A}_{t_1}^{\mathbf{T},+} \mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{S,y}) \nabla_{-jt_1}^{\mathbf{T}} \bar{\mathfrak{Z}}_{S,y}^N .
\end{aligned}
\tag{7.40}
$$

We will now plug the previous display into the heat operator in the definition of $\Phi_j$/the far RHS of the display preceding (7.39) to get the following representation of $\Phi_j$ in which the time-gradient below that acts on the space-time pseudo-gradient average times $\bar{\mathfrak{Z}}^N$ will be transferred to the heat kernel in $\bar{\mathfrak{P}}^N$ via an integration-by-parts-type consideration as we will soon see:

$$
\begin{aligned}
\Phi_j &= \bar{\mathfrak{P}}_{T,x}^N \left( N^{\frac{1}{2}} \nabla_{jt_1}^{\mathbf{T}} \left( \mathscr{A}_{t_1}^{\mathbf{T},+} \mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{S,y}) \bar{\mathfrak{Z}}_{S-jt_1,y}^N \right) \right) \\
&\quad + \bar{\mathfrak{P}}_{T,x}^N \left( N^{\frac{1}{2}} \mathscr{A}_{t_1}^{\mathbf{T},+} \mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{S,y}) \nabla_{-jt_1}^{\mathbf{T}} \bar{\mathfrak{Z}}_{S,y}^N \right) \overset{\bullet}{=} \Phi_{j,1} + \Phi_{j,2} .
\end{aligned}
\tag{7.41}
$$

The identities (7.39) and (7.41) combined with the triangle inequality yield the following estimate:

$$
\begin{aligned}
\| \bar{\mathfrak{P}}_{T,x}^N (\Psi^{N,2,t_1,t_2}\bar{\mathfrak{Z}}^N) \|_{t_{st};\mathbb{T}_N} &\leqslant \widetilde{\sum}_{j=0}^{t_2 t_1^{-1}-1} \| \Phi_j \|_{t_{st};\mathbb{T}_N} \\
&\leqslant \widetilde{\sum}_{j=0}^{t_2 t_1^{-1}-1} \sum_{\mathfrak{k}=1}^{2} \| \Phi_{j,\mathfrak{k}} \|_{t_{st};\mathbb{T}_N} \\
&= 2 \widetilde{\sum}_{j=0}^{t_2 t_1^{-1}-1} \widetilde{\sum}_{\mathfrak{k}=1}^{2} \| \Phi_{j,\mathfrak{k}} \|_{t_{st};\mathbb{T}_N} .
\end{aligned}
\tag{7.42}
$$

To control the average of the norms on the far RHS of (7.42), it suffices to estimate each summand in (7.42) uniformly in both indices on the same high-probability event. We do this by heat operator/time-regularity estimates. First note $jt_1 \lesssim t_2 \lesssim N^{-1}$.

- We bound $\Phi_{j,1}$ in (7.41) by deterministic analytic means. In particular we move the time-gradient onto the heat operator and take advantage of smoothness for the heat operator in time with a sufficiently integrable short-time singularity. This is done in Lemma A.2 which yields the following given any $\delta \in \mathbb{R}_{>0}$ arbitrarily small but universal for which we recall $jt_1 \lesssim t_2 \lesssim N^{-1}$ and boundedness of $\mathscr{A}^{\mathbf{T},+}\mathscr{C}^{\mathbf{X},-}(\mathfrak{g})$. Note $\Phi_{j,1}$ in (7.41) sees $\bar{\mathfrak{Z}}^N$ until time $t_{st}$, which is why its $\|\|_{t_{st};\mathbb{T}_N}$-norm appears below. We also clarify such application of Lemma A.2 is done with $\bar{\varphi}$ equal to the

space-time average times the time-shifted $\bar{\mathfrak{Z}}^N$-process in the time-gradient in $\Phi_{j,1}$ and with the time-scale $t_{reg} = jt_1$ as well as $t_{st}$ therein equal to $t_{st}$ here:

$$\|\Phi_{j,1}\|_{t_{st};\mathbb{T}_N} \lesssim_\delta N^{\frac{1}{2}+2\delta}jt_1\|\bar{\mathfrak{Z}}^N\|_{t_{st};\mathbb{T}_N} \lesssim N^{\frac{1}{2}+2\delta}t_2\|\bar{\mathfrak{Z}}^N\|_{t_{st};\mathbb{T}_N}$$
$$\lesssim N^{-\frac{1}{2}+2\delta}\|\bar{\mathfrak{Z}}^N\|_{t_{st};\mathbb{T}_N} \leqslant N^{-\frac{1}{2}+2\delta}\langle\bar{\mathfrak{Z}}^N\rangle_{t_{st};\mathbb{T}_N}. \tag{7.43}$$

- It remains to estimate $\Phi_{j,2}$. Observe that all of our steps/bounds have been deterministic thus far. We will now use a random high-probability estimate for $\Phi_{j,2}$. It will be convenient to first use the notation $\kappa_{j,t_1,\varepsilon} = N^{-2+\varepsilon} + N^\varepsilon jt_1$. We think of $\varepsilon \in \mathbb{R}_{>0}$ as arbitrarily small but universal in such notation. We also emphasize $\Phi_{j,2}$ is the heat operator acting on the space-time average of the pseudo-gradient times the time-gradient of $\bar{\mathfrak{Z}}^N$ before time $t_{st} \in \mathbb{R}_{\geqslant 0}$.

Via Proposition 6.2, with probability at least $1 - \kappa_{\varepsilon,C}N^{-C}$ for any $\varepsilon, C \in \mathbb{R}_{>0}$, where in this application of Proposition 6.2 we choose $t_{st} \in \mathbb{R}_{\geqslant 0}$ therein equal to $t_{st} \in \mathbb{R}_{\geqslant 0}$ here and we choose $t_{reg} = jt_1$, we may control the time-gradients of $\bar{\mathfrak{Z}}^N$ in $\Phi_{j,2}$ uniformly in the integral on the same high-probability event. We then establish the following estimate upon observing again the $\bar{\mathfrak{Z}}^N$-process is only evaluated before time $t_{st} \in \mathbb{R}_{\geqslant 0}$ in the integral/heat operator $\Phi_{j,2}$. We give more explanation after:

$$|\Phi_{j,2}| \lesssim \|\nabla^{\mathbf{T}}_{-jt_1}\bar{\mathfrak{Z}}^N\|_{t_{st};\mathbb{T}_N}\bar{\mathfrak{P}}^N_{T,x}(N^{\frac{1}{2}}|\mathscr{A}^{\mathbf{T},+}_{t_1}\mathscr{C}^{\mathbf{X},-}_{N^{\beta_X}}(\mathfrak{g}_{S,y})|)$$
$$\lesssim \kappa^{1/4}_{j,t_1,\varepsilon}\langle\bar{\mathfrak{Z}}^N\rangle_{t_{st};\mathbb{T}_N}\|\bar{\mathfrak{P}}^N_{T,x}(N^{\frac{1}{2}}|\mathscr{A}^{\mathbf{T},+}_{t_1}\mathscr{C}^{\mathbf{X},-}_{N^{\beta_X}}(\mathfrak{g})|)\|_{1;\mathbb{T}_N}. \tag{7.44}$$

Let us emphasize (7.44) holds uniformly over $[0, t_{st}] \times \mathbb{T}_N$ with high-probability. We clarify that the second estimate, again, follows by estimating the time-gradient in $\Phi_{j,2}$ via the Holder regularity bound of $\kappa^{1/4}_{j,t_1,\varepsilon}$ times the bracket-norm of $\bar{\mathfrak{Z}}^N$, which holds uniformly in integration variables in the heat operator $\bar{\mathfrak{P}}^N$ in $\Phi_{j,2}$ with high probability by Proposition 6.2.

We recall (7.44) holds uniformly over $[0, t_{st}] \times \mathbb{T}_N$. Thus, upon recalling $\kappa_{j,t_1,\varepsilon}$ defined at the beginning of this bullet point, we get the following high-probability estimate for which we provide a little more explanation afterwards as well:

$$\|\Phi_{j,2}\|_{t_{st};\mathbb{T}_N}$$
$$\lesssim N^{-1/2+\varepsilon/4}\langle\bar{\mathfrak{Z}}^N\rangle_{t_{st};\mathbb{T}_N}\|\bar{\mathfrak{P}}^N_{T,x}(N^{1/2}|\mathscr{A}^{\mathbf{T},+}_{t_1}\mathscr{C}^{\mathbf{X},-}_{N^{\beta_X}}(\mathfrak{g})|)\|_{1;\mathbb{T}_N}$$
$$+ N^{\varepsilon/4}j^{1/4}t_1^{1/4}\langle\bar{\mathfrak{Z}}^N\rangle_{t_{st};\mathbb{T}_N}\|\bar{\mathfrak{P}}^N_{T,x}(N^{1/2}|\mathscr{A}^{\mathbf{T},+}_{t_1}\mathscr{C}^{\mathbf{X},-}_{N^{\beta_X}}(\mathfrak{g})|)\|_{1;\mathbb{T}_N}$$
$$\lesssim N^{-1/6+\varepsilon/4}\langle\bar{\mathfrak{Z}}^N\rangle_{t_{st};\mathbb{T}_N} + N^{\varepsilon/4}j^{1/4}t_1^{1/4}\langle\bar{\mathfrak{Z}}^N\rangle_{t_{st};\mathbb{T}_N}\|\bar{\mathfrak{P}}^N_{T,x}(N^{1/2}|\mathscr{A}^{\mathbf{T},+}_{t_1}\mathscr{C}^{\mathbf{X},-}_{N^{\beta_X}}(\mathfrak{g})|)\|_{1;\mathbb{T}_N}. \tag{7.45}$$

The line (7.45) follows from the line above by the convolution bound (A.3) applied to the first term in the line before and the bound $|\mathscr{A}^{\mathbf{T},+}\mathscr{C}^{\mathbf{X},-}(\mathfrak{g})| \lesssim N^{-\beta_X/2-\varepsilon_{X,1}/2} \leqslant N^{-1/6}$ that follows by time-averaging the a priori $\mathscr{C}^{\mathbf{X},-}$-bound; see Definition 5.1.

We observe the previous two bullet points provide high-probability estimates for $\|\Phi_{j,\ell}\|_{t_{st};\mathbb{T}_N}$ *for each* of the indices $j, \ell \in \mathbb{Z}_{\geqslant 0}$. By "high-probability" in this claim, we mean an event whose complement has probability $\kappa_C N^{-C}$ for any $C \in \mathbb{R}_{\geqslant 0}$. With a union bound over the complement of all events over all indices $j, \ell$, we deduce the event on which all of our $\|\Phi_{j,\ell}\|_{t_{st};\mathbb{T}_N}$-estimates hold *simultaneously* has complement with probability at most $\kappa_C N^{-C}t_2t_1^{-1} \lesssim \kappa_C N^{-C+100}$ with universal implied constant. This final

probability bound follows from the assumed constraints $N^{-100} \lesssim t_1, t_2 \lesssim N^{-1}$ in the statement of Lemma 7.7. Combining these simultaneous high-probability $\|\Phi_{j,\mathfrak{k}}\|_{t_{st};\mathbb{T}_N}$-estimates with (7.42) yields the desired estimate. □

*Proof of Lemma 7.8.* Let us first define the intermediate sub-microscopic time-scale of $t_{1,2} \stackrel{\bullet}{=} N^{-100}$ between $t_1, t_2 \in \mathbb{R}_{\geqslant 0}$. The triangle inequality gives the following in which we first replace a scale-0 time-average by the intermediate time-scale $t_{1,2}$:

$$\|\bar{\mathfrak{P}}_{T,x}^N(\Psi^{N,2,0,t_{av,0}}\bar{\mathfrak{Z}}^N)\|_{t_{st};\mathbb{T}_N} \leqslant \|\bar{\mathfrak{P}}_{T,x}^N(\Psi^{N,2,0,t_{1,2}}\bar{\mathfrak{Z}}^N)\|_{t_{st};\mathbb{T}_N}$$
$$+ \|\bar{\mathfrak{P}}_{T,x}^N(\Psi^{N,2,t_{1,2},t_{av,0}}\bar{\mathfrak{Z}}^N)\|_{t_{st};\mathbb{T}_N}. \tag{7.46}$$

Let us now explain the utility of this intermediate time-scale $t_{1,2} = N^{-100}$. The first term on the RHS of (7.46) can be estimated via first moment after pulling the $\bar{\mathfrak{Z}}^N$-factor outside the heat operator. This first moment bound would yield a "time-regularity" bound for local functionals of the particle system. It says at times $t_{1,2} = N^{-100}$, we do not expect to see any of the Poisson clocks ring, and thus local functionals are "constant" or "smooth" on the time-scale $t_{1,2} = N^{-100}$. The second term on the RHS of (7.46) is then estimated by (7.17) in Lemma 7.7, as we have introduced a preliminary time-average $t_{1,2} = N^{-100}$ for this second term.

We make the above precise. Pulling $\|\bar{\mathfrak{Z}}^N\|_{t_{st};\mathbb{T}_N} \leqslant \langle\bar{\mathfrak{Z}}^N\rangle_{t_{st};\mathbb{T}_N}$ from the heat operator in the first term on the RHS of (7.46) gives

$$\|\bar{\mathfrak{P}}_{T,x}^N(\Psi^{N,2,0,t_{1,2}}\bar{\mathfrak{Z}}^N)\|_{t_{st};\mathbb{T}_N} \lesssim \langle\bar{\mathfrak{Z}}^N\rangle_{t_{st};\mathbb{T}_N}\|\left(\int_0^T \sum_{y\in\mathbb{T}_N}|\Psi_{S,y}^{N,2,0,t_{1,2}}|\,dS\right)\|_{t_{st};\mathbb{T}_N}$$
$$\lesssim \langle\bar{\mathfrak{Z}}^N\rangle_{t_{st};\mathbb{T}_N}\int_0^1 \sum_{y\in\mathbb{T}_N}|\Psi_{S,y}^{N,2,0,t_{1,2}}|\,dS. \tag{7.47}$$

The estimate before the far RHS of (7.47)/in the middle of the above display follows by a maximum principle $\mathbf{1}_{T\geqslant S}\bar{\mathfrak{P}}_{S,T,x,y}^N \leqslant 1$. The far RHS of (7.47) follows by extending the integration-domain in the integral of a non-negative term from $[0, T] \to [0, 1]$.

To get a high-probability estimate for the RHS, by the Markov inequality it suffices to bound expectation of the RHS of (7.47) *without* the bracket factor in front. To this end, we pull said expectation past the integral and summation so it hits the $\Psi$-factor. To bound expectation of this $\Psi$-factor, recall by definition this $\Psi$-term is the difference between the spatial-average-with-cutoff $\mathscr{C}^{\mathbf{X},-}(\mathfrak{g})$ and its time-average on scale-$t_{1,2} = N^{-100}$. As $\mathscr{C}^{\mathbf{X},-}(\mathfrak{g})$ is uniformly bounded with support with size of order $N^{\beta_X}$, this $\Psi$-term, in absolute value, is controlled by the number of times any of the Poisson clocks in the support of $\mathscr{C}^{\mathbf{X},-}(\mathfrak{g})$, all speed at most $N^2$ times uniformly bounded factors, ring in the time-scale $t_{1,2} = N^{-100}$. As a sum of independent Poissons is Poisson via additive rates, this is asking for the number of times a Poisson clock of speed $N^{2+\beta_X-100}$ rings in order 1 time. The probability that such Poisson clock rings at all is controlled by its speed if its speed is uniformly bounded. Thus, as $\beta_X = \frac{1}{3} + \varepsilon_{X,1} \leqslant 1$,

$$\mathbf{E}|\Psi_{S,y}^{N,2,0,t_{1,2}}| \lesssim N^{2+\beta_X-100} \lesssim N^{-97}. \tag{7.48}$$

As a consequence (7.47), the proof of the Markov inequality gives the required estimate for the first term on the RHS of (7.46) outside an event of probability bounded by the following in which we employ (7.48) along with the Fubini theorem and control on an

unaveraged sum/integral by the length of the sum/integration set times a supremum; recall $|\mathbb{T}_N| \lesssim N^2$ from Sect. 4:

$$N^{\beta_{\text{univ},2}} \mathbf{E} \int_0^1 \sum\nolimits_{y \in \mathbb{T}_N} |\Psi_{S,y}^{N,2,0,\mathfrak{t}_{1,2}}| \, dS = N^{\beta_{\text{univ},2}} \int_0^1 \sum\nolimits_{y \in \mathbb{T}_N} \mathbf{E} |\Psi_{S,y}^{N,2,0,\mathfrak{t}_{1,2}}| \, dS$$
$$\lesssim N^{\beta_{\text{univ},2}} |\mathbb{T}_N| N^{-97} \lesssim N^{-97+2+\beta_{\text{univ},2}}. \tag{7.49}$$

The last bound follows from (7.48). Choosing $\beta_{\text{univ},2} \in \mathbb{R}_{>0}$ sufficiently small but still universal gives that with high-probability, the first term from the RHS of (7.46) is controlled by the RHS of the proposed estimate in Lemma 7.8. By (7.46), the proposed estimate in Lemma 7.8 follows if we control the second term on the RHS of (7.46) with high probability.

To control the second term from the RHS of (7.46) with required high probability, we use Lemma 7.7 with $\mathfrak{t}_1 = \mathfrak{t}_{1,2} = N^{-100}$ and $\mathfrak{t}_2 = \mathfrak{t}_{\text{av},0} \lesssim N^{-2+\varepsilon_{X,2}}$ for $\varepsilon_{X,2} \in \mathbb{R}_{>0}$ arbitrarily small but universal. We also recall $|\mathscr{A}^{\mathbf{T},+}\mathscr{C}^{\mathbf{X},-}(\mathfrak{g})| \lesssim N^{-1/6}$. Ultimately, with the required high probability we have the following estimate for the second term on the RHS of (7.46):

$$\|\bar{\mathfrak{P}}_{T,x}^N(\Psi^{N,2,\mathfrak{t}_{1,2},\mathfrak{t}_{\text{av},0}}\bar{\mathfrak{Z}}^N)\|_{\mathfrak{t}_{\text{st}};\mathbb{T}_N}$$
$$\lesssim_\varepsilon N^{-\beta_{\text{univ},2}} \langle \bar{\mathfrak{Z}}^N \rangle_{\mathfrak{t}_{\text{st}};\mathbb{T}_N} + \langle \bar{\mathfrak{Z}}^N \rangle_{\mathfrak{t}_{\text{st}};\mathbb{T}_N} N^\varepsilon \mathfrak{t}_{\text{av},0}^{1/4} \|\bar{\mathfrak{P}}_{T,x}^N(N^{\frac{1}{2}}|\mathscr{A}_{\mathfrak{t}_{1,2}}^{\mathbf{T},+}\mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g})|)\|_{1;\mathbb{T}_N} \tag{7.50}$$

$$\lesssim N^{-\beta_{\text{univ},2}} \langle \bar{\mathfrak{Z}}^N \rangle_{\mathfrak{t}_{\text{st}};\mathbb{T}_N} + \langle \bar{\mathfrak{Z}}^N \rangle_{\mathfrak{t}_{\text{st}};\mathbb{T}_N} N^{\frac{1}{2}+\varepsilon} N^{-\frac{1}{6}} \mathfrak{t}_{\text{av},0}^{1/4}. \tag{7.51}$$

Elementary power-counting with $\mathfrak{t}_{\text{av},0} \lesssim N^{-2+\varepsilon_{X,2}}$ bounds the second term on the RHS of (7.46). This completes the proof. □

*Proof of Lemma 7.9.* We observe that Lemma 7.9 differs from the earlier result Lemma 7.7 just through the replacement (7.3). The proof of Lemma 7.9 amounts to verbatim following proof of Lemma 7.7 upon making the replacement (7.3). In particular, the only important thing is that we replace the time-scale of time-averaging from $\mathfrak{t}_1 \rightsquigarrow \mathfrak{t}_2$. The details of what we time-average are irrelevant. Actually, the replacement of (7.3) in the proof of Lemma 7.7 changes power-counting in the proof of Lemma 7.7 but in ultimately irrelevant ways if instead of the bound $|\mathscr{C}^{\mathbf{X},-}(\mathfrak{g})| \lesssim N^{-\frac{1}{6}}$, we instead use $|\widetilde{\mathfrak{g}}^{\mathfrak{t}}| \lesssim 1$ in this power-counting. □

*Proof of Lemma 7.10.* We follow the proof of Lemma 7.8 with the minor changes mentioned in the proof of Lemma 7.9. □

*Proof of Lemma 7.11.* The architecture behind this argument is to relate $\Gamma^{N,m,m+1}$ and $\Gamma^{N,m+1,m+2}$ from the statement of Lemma 7.11 via exact identities. We then estimate the errors obtained when relating $\Gamma^{N,m,m+1}$ and $\Gamma^{N,m+1,m+2}$ in these exact identities. Before we start, we invite the reader to refer to Lemma 7.11 to recall definitions of $\Gamma$-terms during this proof. First, notation.

- Take $m, \mathfrak{l}, \mathfrak{k}, \mathfrak{j} \in \mathbb{Z}_{\geq 0}$. For $T \in \mathbb{R}_{\geq 0}$ let $T_{\mathfrak{l},\mathfrak{k},\mathfrak{j}}^m \overset{\bullet}{=} T + \mathfrak{l}\mathfrak{t}_{\text{av},m+2} + \mathfrak{k}\mathfrak{t}_{\text{av},m+1} + \mathfrak{j}\mathfrak{t}_{\text{av},m}$ be time-shifts at three $m$-dependent time-scales. In particular each subscript in this notation indicates shifting at one of the three time-scales $\mathfrak{t}_{\text{av},m+2}, \mathfrak{t}_{\text{av},m+1}, \mathfrak{t}_{\text{av},m}$ in this order.

- By definition the term $\Gamma^{N,m,m+1}$ from the LHS of the proposed estimate is an average of $t_{av,m_+}t_{av,m+1}^{-1}$-many time-shifted terms $\Gamma^{N,t_{av,m}}$ each of which is the average of scale-$t_{av,m}$ time-averages equipped with a priori upper bound cutoff of $N^{-\beta_m}$. To relate this to $\Gamma^{N,m+1,m+2}$, we will group each of the $t_{av,m_+}t_{av,m+1}^{-1}$-many time-shifted terms $\Gamma^{N,t_{av,m}}$ into $t_{av,m_+}t_{av,m+2}^{-1}$-many groups. We will then relate each of these $t_{av,m_+}t_{av,m+2}^{-1}$-many groups to $\Gamma^{N,t_{av,m+1}}$-terms. Because the term $\Gamma^{N,m,m+1}$ is the average of these $t_{av,m_+}t_{av,m+2}^{-1}$-many $\Gamma^{N,t_{av,m+1}}$-terms, making such a relation and estimating error terms will suffice. Precisely, the first step is to compute $\Gamma^{N,m,m+1}$ by decompositions with respect to $t_{av,m+2}$-, $t_{av,m+1}$-, and $t_{av,m}$-time scales. We collect all $t_{av,m_+}t_{av,m+1}^{-1}$-many length-$t_{av,m+1}$ $\Gamma^{N,t_{av,m}}$-terms inside $\Gamma^{N,m,m+1}$ into length-$t_{av,m+2}$ sums, each with $t_{av,m+2}t_{av,m+1}^{-1}$-many summands. Equivalently, we rewrite the sum in $\Gamma^{N,m,m+1}$ with time-scale increments given by multiples of $t_{av,m+1}$ by grouping time-scales by multiples of $t_{av,m+2}$. For example a sum of 10 terms is a sum of 5 other terms which are each sums of 2 terms with neighboring indices:

$$
\begin{aligned}
\Gamma_{T,x}^{N,m,m+1} &= \widetilde{\sum}_{\mathfrak{l}=0}^{t_{av,m_+}t_{av,m+2}^{-1}-1}\widetilde{\sum}_{\mathfrak{k}=0}^{t_{av,m+2}t_{av,m+1}^{-1}-1}\Gamma_{T+[\mathfrak{l}t_{av,m+2}+\mathfrak{k}t_{av,m+1},x}^{N,t_{av,m}}\\
&= \widetilde{\sum}_{\mathfrak{l}=0}^{t_{av,m_+}t_{av,m+2}^{-1}-1}\widetilde{\sum}_{\mathfrak{k}=0}^{t_{av,m+2}t_{av,m+1}^{-1}-1}\Gamma_{T_{\mathfrak{l},\mathfrak{k},0}^m,x}^{N,t_{av,m}}.
\end{aligned}\tag{7.52}
$$

Recall in Corollary 5.12 that $t_{av,m}$ are integer multiples of each other. Our goal is now to relate the $\mathfrak{k}$-sum in (7.52) to $\Gamma^{N,t_{av,m+1}}$.

- Building off the previous bullet point, we recall $\Gamma^{N,t_{av,m+1}}$ is an average of time-shifted time-averages of $\mathscr{C}^{\mathbf{X},-}(\mathfrak{g})$ each on the time-scale $t_{av,m+1}$ with upper bound cutoffs of $N^{-\beta_{m+1}}$. The first step we take towards getting such $\Gamma^{N,t_{av,m+1}}$-terms out of the inner summations from the RHS of (7.52) is to first equip the scale-$t_{av,m+1}$ inner summations in (7.52) with an upper bound cutoff of $N^{-\beta_m}$ for a scale-$t_{av,m+1}$ time-average which we later upgrade to $N^{-\beta_{m+1}}$. Precisely, with explanation given after,

$$
\begin{aligned}
\Gamma_{T,x}^{N,m,m+1} &= \widetilde{\sum}_{\mathfrak{l}=0}^{t_{av,m_+}t_{av,m+2}^{-1}-1}\widetilde{\sum}_{\mathfrak{k}=0}^{t_{av,m+2}t_{av,m+1}^{-1}-1}\Gamma_{\mathfrak{l},\mathfrak{k},1}\\
&\quad + \widetilde{\sum}_{\mathfrak{l}=0}^{t_{av,m_+}t_{av,m+2}^{-1}-1}\widetilde{\sum}_{\mathfrak{k}=0}^{t_{av,m+2}t_{av,m+1}^{-1}-1}\Gamma_{\mathfrak{l},\mathfrak{k},2} \overset{\bullet}{=} \Phi_1 + \Phi_2.
\end{aligned}\tag{7.53}
$$

Here $\Gamma_{\mathfrak{l},\mathfrak{k},i} \overset{\bullet}{=} \Gamma_{T_{\mathfrak{l},\mathfrak{k},0}^m,x}^{N,t_{av,m}}\mathbf{1}[\mathscr{E}_i]$ with $\mathscr{E}_2$ the complement of $\mathscr{E}_1$ below and thus $\Gamma_{T_{\mathfrak{l},\mathfrak{k},0}^m,x}^{N,t_{av,m}} = \Gamma_{T_{\mathfrak{l},\mathfrak{k},0}^m,x}^{N,t_{av,m}}\mathbf{1}[\mathscr{E}_1]+\Gamma_{T_{\mathfrak{l},\mathfrak{k},0}^m,x}^{N,t_{av,m}}\mathbf{1}[\mathscr{E}_2] = \Gamma_{\mathfrak{l},\mathfrak{k},1}+\Gamma_{\mathfrak{l},\mathfrak{k},2}$. In particular, plugging this last identity into the far RHS of (7.52) produces the identity (7.53). The $\mathscr{E}_1$ event is defined as

$$
\mathbf{1}[\mathscr{E}_1] \overset{\bullet}{=} \mathbf{1}\left(\sup_{0\leqslant \mathfrak{t}\leqslant t_{av,m+1}}\mathfrak{t}\cdot t_{av,m+1}^{-1}|\mathscr{A}_{\mathfrak{t}}^{\mathbf{T},+}\mathscr{C}_{N^{\beta}X}^{\mathbf{X},-}(\mathfrak{g}_{T_{\mathfrak{l},\mathfrak{k},0}^m,x})| \leqslant N^{-\beta_m}\right).\tag{7.54}
$$

The $\mathscr{E}_1$-event gives a cutoff for time-averages on scale $t_{av,m+1}$. The $\Phi_2$-term from the far RHS of (7.53) will be treated as an error term at the end of this proof. We unfold the $\Phi_1$-term by unfolding each $\Gamma_{\mathfrak{l},\mathfrak{k},1}$-summand. This summand is a time-shifted $\Gamma^{N,t_{av,m}}$-term with the a priori upper bound cutoff in $\mathscr{E}_1$. Each of the $\Gamma^{N,t_{av,m}}$-terms is an average of time-shifted time-averages of $\mathscr{C}^{\mathbf{X},-}(\mathfrak{g})$ each with an a priori upper

bound cutoff of $N^{-\beta_m}$. If we take away these upper bound cutoffs for the time-shifted time-averages inside $\Gamma^{N,\mathsf{t}_{\mathrm{av},m}}$, then the resulting $\Gamma^{N,\mathsf{t}_{\mathrm{av},m}}$-term would be the average of time-shifted time-averages of $\mathscr{C}^{\mathbf{X},-}(\mathfrak{g})$ on adjacent time-intervals of length $\mathsf{t}_{\mathrm{av},m}$ which would glue together into a time-average of $\mathscr{C}^{\mathbf{X},-}(\mathfrak{g})$ on the time-scale $\mathsf{t}_{\mathrm{av},m+1}$. The indicator function $\mathscr{E}_1$ would then grant the resulting scale-$\mathsf{t}_{\mathrm{av},m+1}$ average an a priori upper bound cutoff of $N^{-\beta_m}$. Thus, the next step we will take is to adjust $\Gamma_{\mathfrak{l},\mathfrak{k},1}$ by keeping the $\mathbf{1}[\mathscr{E}_1]$-factor but removing the a priori upper bound cutoffs inside the time-averages $\Gamma^{N,\mathsf{t}_{\mathrm{av},m},\mathsf{j}}$ defining the $\Gamma^{N,\mathsf{t}_{\mathrm{av},m}}$-factor in $\Gamma_{\mathfrak{l},\mathfrak{k},1}$. Precisely, recalling $\Gamma_{\mathfrak{l},\mathfrak{k},1}$ defined immediately after (7.53) and looking at definitions of $\Gamma^{N,\mathsf{t}_{\mathrm{av},m}}$ and $\Gamma^{N,\mathsf{t}_{\mathrm{av},m},\mathsf{j}}$ in Lemma 7.11, we get, with explanation to come, a scale-$\mathsf{t}_{\mathrm{av},m}/\mathsf{t}_{\mathrm{av},m+1}$ expansion below in which we treat each of the scale-$\mathsf{t}_{\mathrm{av},m}$ $\Gamma^{N,\mathsf{t}_{\mathrm{av},m},\mathsf{j}}$-terms in $\Gamma^{N,\mathsf{t}_{\mathrm{av},m}}$ by removing their indicator functions to get an honest scale-$\mathsf{t}_{\mathrm{av},m}$ time-average. This is the first term on the far RHS below. We get an error of the same scale-$\mathsf{t}_{\mathrm{av},m}$ average although with lower bound cutoff complement to the upper bound event in $\Gamma^{N,\mathsf{t}_{\mathrm{av},m},\mathsf{j}}$. This is the second term on the far RHS below:

$$\Gamma_{\mathfrak{l},\mathfrak{k},1} = \mathbf{1}[\mathscr{E}_1]\Gamma^{N,\mathsf{t}_{\mathrm{av},m}}_{T^m_{\mathfrak{l},\mathfrak{k},0},x} = \widetilde{\sum}_{\mathsf{j}=0}^{\mathsf{t}_{\mathrm{av},m+1}\mathsf{t}_{\mathrm{av},m}^{-1}}\mathbf{1}[\mathscr{E}_1]\Gamma^{N,\mathsf{t}_{\mathrm{av},m},\mathsf{j}}_{T^m_{\mathfrak{l},\mathfrak{k},0},x}$$

$$= \widetilde{\sum}_{\mathsf{j}=0}^{\mathsf{t}_{\mathrm{av},m+1}\mathsf{t}_{\mathrm{av},m}^{-1}-1}\Gamma_{\mathfrak{l},\mathfrak{k},\mathsf{j},1,1} + \widetilde{\sum}_{\mathsf{j}=0}^{\mathsf{t}_{\mathrm{av},m+1}\mathsf{t}_{\mathrm{av},m}^{-1}-1}\Gamma_{\mathfrak{l},\mathfrak{k},\mathsf{j},1,2}. \qquad (7.55)$$

The terms in the two-term decomposition within the RHS of (7.55) are defined below with an event defined afterwards. The first term is a scale-$\mathsf{t}_{\mathrm{av},m}$ average over a window contained inside the length-$\mathsf{t}_{\mathrm{av},m+1}$ block starting at the time $T^m_{\mathfrak{l},\mathfrak{k},0}$ associated to $\Gamma_{\mathfrak{l},\mathfrak{k},1}$, equipped with the event $\mathscr{E}_1$ which is defined at scale $\mathsf{t}_{\mathrm{av},m+1}$/depends only on $\mathfrak{l},\mathfrak{k}$ variables. The second term below is the same but now with an additional indicator function $\mathbf{1}[\mathscr{F}]$ obtained by removing the indicator function in $\Gamma^{N,\mathsf{t}_{\mathrm{av},m},\mathsf{j}}$-terms:

$$\Gamma_{\mathfrak{l},\mathfrak{k},\mathsf{j},1,1} \overset{\bullet}{=} \mathscr{A}^{\mathbf{T},+}_{\mathsf{t}_{\mathrm{av},m}}\mathscr{C}^{\mathbf{X},-}_{N^{\beta_X}}(\mathfrak{g}_{T^m_{\mathfrak{l},\mathfrak{k},\mathsf{j}},x})\cdot\mathbf{1}[\mathscr{E}_1] \quad \text{and} \quad \Gamma_{\mathfrak{l},\mathfrak{k},\mathsf{j},1,2}$$

$$\overset{\bullet}{=} -\mathscr{A}^{\mathbf{T},+}_{\mathsf{t}_{\mathrm{av},m}}\mathscr{C}^{\mathbf{X},-}_{N^{\beta_X}}(\mathfrak{g}_{T^m_{\mathfrak{l},\mathfrak{k},\mathsf{j}},x})\cdot\mathbf{1}[\mathscr{E}_1]\mathbf{1}[\mathscr{F}].$$

We introduced the following complement of the event in the indicator function defining $\Gamma^{N,\mathsf{t}_{\mathrm{av},m},\mathsf{j}}$. Indeed, it appears in (7.55) by removing said indicator function/observing the indicator function in $\Gamma^{N,\mathsf{t}_{\mathrm{av},m},\mathsf{j}}$ is equal to $1 - \mathbf{1}[\mathscr{F}]$:

$$\mathbf{1}[\mathscr{F}] \overset{\bullet}{=} \mathbf{1}\left(\sup_{0\leqslant\mathsf{t}\leqslant\mathsf{t}_{\mathrm{av},m}}\mathsf{t}\cdot\mathsf{t}_{\mathrm{av},m}^{-1}|\mathscr{A}^{\mathbf{T},+}_{\mathsf{t}}\mathscr{C}^{\mathbf{X},-}_{N^{\beta_X}}(\mathfrak{g}_{T^m_{\mathfrak{l},\mathfrak{k},\mathsf{j}},x})| > N^{-\beta_m}\right). \qquad (7.56)$$

Thus, averaging the previous expansion (7.55) over $\mathfrak{l},\mathfrak{k}$-indices gives the decomposition

$$\Phi_1 = \widetilde{\sum}_{\mathfrak{l}=0}^{\mathsf{t}_{\mathrm{av},m+}\mathsf{t}_{\mathrm{av},m+2}^{-1}-1}\widetilde{\sum}_{\mathfrak{k}=0}^{\mathsf{t}_{\mathrm{av},m+2}\mathsf{t}_{\mathrm{av},m+1}^{-1}-1}\widetilde{\sum}_{\mathsf{j}=0}^{\mathsf{t}_{\mathrm{av},m+1}\mathsf{t}_{\mathrm{av},m}^{-1}-1}\Gamma_{\mathfrak{l},\mathfrak{k},\mathsf{j},1,1}$$

$$+ \widetilde{\sum}_{\mathfrak{l}=0}^{\mathsf{t}_{\mathrm{av},m+}\mathsf{t}_{\mathrm{av},m+2}^{-1}-1}\widetilde{\sum}_{\mathfrak{k}=0}^{\mathsf{t}_{\mathrm{av},m+2}\mathsf{t}_{\mathrm{av},m+1}^{-1}-1}\widetilde{\sum}_{\mathsf{j}=0}^{\mathsf{t}_{\mathrm{av},m+1}\mathsf{t}_{\mathrm{av},m}^{-1}-1}\Gamma_{\mathfrak{l},\mathfrak{k},\mathsf{j},1,2}. \qquad (7.57)$$

We denote the first sum on the RHS of (7.57) by $\Phi_3$ and the second by $\Phi_4$. We treat $\Phi_4$ as an error like $\Phi_2$. We now study $\Phi_3$. We first take the sum over the index $\mathsf{j}$. We observe the event $\mathscr{E}_1$ defined after (7.53) is independent of this sum variable since it is defined by a constraint at scale $\mathsf{t}_{\mathrm{av},m+1}$/with respect to $\mathfrak{l},\mathfrak{k}$-indices. By definition of

$\Gamma_{\mathfrak{l},\mathfrak{k},j,1,1}$ given after (7.55), we sum over the j-variable/the finest $t_{\mathrm{av},m}$ time-scale in $\Phi_3$ and glue such scale-$t_{\mathrm{av},m}$ terms into scale-$t_{\mathrm{av},m+1}/\mathfrak{l}$, $\mathfrak{k}$-terms:

$$\sum_{j=0}^{\widetilde{t_{\mathrm{av},m+1}t_{\mathrm{av},m}^{-1}-1}} \Gamma_{\mathfrak{l},\mathfrak{k},j,1,1} = \mathbf{1}[\mathscr{E}_1] \cdot \sum_{j=0}^{\widetilde{t_{\mathrm{av},m+1}t_{\mathrm{av},m}^{-1}-1}} \mathscr{A}_{t_{\mathrm{av},m}}^{\mathbf{T},+} \mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{T_{\mathfrak{l},\mathfrak{k},j}^m},x)$$
$$= \mathbf{1}[\mathscr{E}_1] \cdot \mathscr{A}_{t_{\mathrm{av},m+1}}^{\mathbf{T},+} \mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{T_{\mathfrak{l},\mathfrak{k},0}^m},x). \tag{7.58}$$

Above, (7.58) follows from observing the $t_{\mathrm{av},m+1}$-scale time average is an average of the preceding suitably time-shifted time averages at the smaller $t_{\mathrm{av},m}$-scale. We emphasize (7.55) and (7.58) are a precise implementation of the statement previous to (7.55) concerning adjustment of $\Gamma_{\mathfrak{l},\mathfrak{k},1}$ to turn it into a time-average of $\mathscr{C}^{\mathbf{X},-}(\mathfrak{g})$ on time-scale $t_{\mathrm{av},m+1}$ with an a priori upper bound cutoff of $N^{-\beta_m}$. By definition of $\Phi_3$ as the first triple sum from the RHS of (7.57), from (7.58) we then get the following representation of $\Phi_3$ as an average of time-shifted time-averages of $\mathscr{C}^{\mathbf{X},-}(\mathfrak{g})$ on scales $t_{\mathrm{av},m+1}$ with upper bound cutoff $N^{-\beta_m}$:

$$\Phi_3 = \sum_{\mathfrak{l}=0}^{\widetilde{t_{\mathrm{av},m+}t_{\mathrm{av},m+2}^{-1}-1}} \sum_{\mathfrak{k}=0}^{\widetilde{t_{\mathrm{av},m+2}t_{\mathrm{av},m+1}^{-1}-1}} \mathbf{1}[\mathscr{E}_1] \cdot \mathscr{A}_{t_{\mathrm{av},m+1}}^{\mathbf{T},+} \mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{T_{\mathfrak{l},\mathfrak{k},0}^m},x). \tag{7.59}$$

- We build off the previous bullet point starting with (7.59). Let us observe that $\Phi_3$ is *almost* equal to $\Gamma^{N,m+1,m+2}$. This would give us an explicit formula to relate $\Gamma^{N,m,m+1}$ and $\Gamma^{N,m+1,m+2}$ because $\Phi_3$ is related to $\Phi_1$ with an error term $\Phi_4$ via (7.57), and $\Phi_1$ is related to $\Gamma^{N,m,m+1}$ through an error term $\Phi_2$ via (7.53). The "almost"-ness of the equality between $\Phi_3$ and $\Gamma^{N,m+1,m+2}$ is because the term in (7.58) or equivalently the summand in (7.59) is *almost* equal to the time-average of $\mathscr{C}^{\mathbf{X},-}$ on time-scale $t_{\mathrm{av},m+1}$ with upper bound cutoff of $N^{-\beta_{m+1}}$, and such a time-average with $N^{-\beta_{m+1}}$-cutoff is equal to $\Gamma^{N,t_{\mathrm{av},m+1},\mathfrak{k}}$ by definition from the statement of Lemma 7.11, and similarly by definition $\Gamma^{N,m+1,m+2}$ averages these $\Gamma^{N,t_{\mathrm{av},m+1},\mathfrak{k}}$-terms. However, the cutoff in $\mathscr{E}_1$ is the upper bound of $N^{-\beta_m}$, not $N^{-\beta_{m+1}}$. Thus, the next step we take is to upgrade the cutoff $N^{-\beta_m}$ in $\mathscr{E}_1$ to $N^{-\beta_{m+1}}$. Precisely,

$$\mathbf{1}[\mathscr{E}_1] \cdot \mathscr{A}_{t_{\mathrm{av},m+1}}^{\mathbf{T},+} \mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{T_{\mathfrak{l},\mathfrak{k},0}^m},x) = (\mathbf{1}[\mathscr{G}_1] + \mathbf{1}[\mathscr{E}_1]\mathbf{1}[\mathscr{G}_2]) \cdot \mathscr{A}_{t_{\mathrm{av},m+1}}^{\mathbf{T},+} \mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{T_{\mathfrak{l},\mathfrak{k},0}^m},x) \tag{7.60}$$

where $\mathscr{G}_1$ is the following upgrade of $\mathscr{E}_1$ that is certainly contained in $\mathscr{E}_1$, so $\mathbf{1}[\mathscr{G}_1]\mathbf{1}[\mathscr{E}_1] = \mathbf{1}[\mathscr{G}_1]$, and $\mathscr{G}_2$ is its complement:

$$\mathbf{1}[\mathscr{G}_1] \overset{\bullet}{=} \mathbf{1}\left(\sup_{0\leqslant t\leqslant t_{\mathrm{av},m+1}} t \cdot t_{\mathrm{av},m+1}^{-1} |\mathscr{A}_t^{\mathbf{T},+} \mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{T_{\mathfrak{l},\mathfrak{k},0}^m},x)| \leqslant N^{-\beta_{m+1}}\right). \tag{7.61}$$

By (7.59), (7.60), and definition of $\Gamma^{N,t_{\mathrm{av},m+1}}$ and $\Gamma^{N,t_{\mathrm{av},m+1},\mathfrak{k}}$ in the statement of Lemma 7.11, we average over $\mathfrak{l}$, $\mathfrak{k}$-indices to get

$$\Phi_3 = \sum_{\mathfrak{l}=0}^{\widetilde{t_{\mathrm{av},m+}t_{\mathrm{av},m+2}^{-1}-1}} \sum_{\mathfrak{k}=0}^{\widetilde{t_{\mathrm{av},m+2}t_{\mathrm{av},m+1}^{-1}-1}} \mathbf{1}[\mathscr{E}_1] \cdot \mathscr{A}_{t_{\mathrm{av},m+1}}^{\mathbf{T},+} \mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{T_{\mathfrak{l},\mathfrak{k},0}^m},x) \tag{7.62}$$

$$= \sum_{\mathfrak{l}=0}^{\widetilde{t_{\mathrm{av},m+}t_{\mathrm{av},m+2}^{-1}-1}} \sum_{\mathfrak{k}=0}^{\widetilde{t_{\mathrm{av},m+2}t_{\mathrm{av},m+1}^{-1}-1}} \mathbf{1}[\mathscr{G}_1] \cdot \mathscr{A}_{t_{\mathrm{av},m+1}}^{\mathbf{T},+} \mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{T_{\mathfrak{l},\mathfrak{k},0}^m},x) + \Phi_5 \tag{7.63}$$

$$= \sum_{\mathfrak{l}=0}^{\widetilde{t_{\mathrm{av},m+}t_{\mathrm{av},m+2}^{-1}-1}} \sum_{\mathfrak{k}=0}^{\widetilde{t_{\mathrm{av},m+2}t_{\mathrm{av},m+1}^{-1}-1}} \Gamma_{T+\mathfrak{l}t_{\mathrm{av},m+2},x}^{N,t_{\mathrm{av},m+1},\mathfrak{k}} + \Phi_5 \tag{7.64}$$

$$= \sum_{\mathfrak{l}=0}^{\widetilde{t_{\mathrm{av},m+}t_{\mathrm{av},m+2}^{-1}-1}} \Gamma_{T+\mathfrak{l}t_{\mathrm{av},m+2},x}^{N,t_{\mathrm{av},m+1}} + \Phi_5 = \Gamma_{T,x}^{N,m+1,m+2} + \Phi_5. \tag{7.65}$$

The $\Phi_5$-term introduced in the derivation of (7.65) above is defined below:

$$\Phi_5 \doteq \widetilde{\sum}_{\mathfrak{l}=0}^{\mathfrak{t}_{\mathrm{av},\mathfrak{m}_+}\mathfrak{t}_{\mathrm{av},m+2}^{-1}-1} \widetilde{\sum}_{\mathfrak{k}=0}^{\mathfrak{t}_{\mathrm{av},m+2}\mathfrak{t}_{\mathrm{av},m+1}^{-1}-1} \mathbf{1}[\mathscr{E}_1]\mathbf{1}[\mathscr{G}_2] \cdot \mathscr{A}_{\mathfrak{t}_{\mathrm{av},m+1}}^{\mathbf{T},+} \mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{T_{\mathfrak{l},\mathfrak{k},0}^m},x). \tag{7.66}$$

We repeat that the first two identities in the calculation to get (7.65) follow by (7.59), (7.60), respectively and the remaining lines follow by definitions in the statement of the current Lemma 7.11. We explain this as follows in case of interest/clarity. In the following explanation all of the references to the definitions of $\Gamma$-terms can be found in the statement of Lemma 7.11. In the calculation to get (7.65), the third identity follows from observing the summand in the second line in that calculation is a time-shifted time-average of $\mathscr{C}^{\mathbf{X},-}(\mathfrak{g})$ on time-scale $\mathfrak{t}_{\mathrm{av},m+1}$ with an a priori upper bound cutoff of $N^{-\beta_{m+1}}$ in $\mathscr{G}_1$, and this is precisely the $\Gamma^{N,\mathfrak{t}_{\mathrm{av},m+1},\mathfrak{k}}$-type term in the statement of the current Lemma 7.11. The fourth line follows by recalling $\Gamma^{N,\mathfrak{t}_{\mathrm{av},m+1}}$ collects-via-averaging $\mathfrak{t}_{\mathrm{av},m+2}\mathfrak{t}_{\mathrm{av},m+1}^{-1}$-many time-shifted versions of $\Gamma^{N,\mathfrak{t}_{\mathrm{av},m+1},\mathfrak{k}}$-terms by definition as well. The fifth and final identity in (7.65) follows by similarly recalling that $\Gamma^{N,m+1,m+2}$ collects $\mathfrak{t}_{\mathrm{av},\mathfrak{m}_+}\mathfrak{t}_{\mathrm{av},m+2}^{-1}$-many time-shifted versions of $\Gamma^{N,\mathfrak{t}_{\mathrm{av},m+1}}$.

By (7.53), (7.57), and (7.65), we have a transfer-of-scales with errors $\Phi_2$, $\Phi_4$, $\Phi_5$ we are left to bound after multiplying by $N^{1/2}$:

$$\Gamma_{T,x}^{N,m,m+1} = \Gamma_{T,x}^{N,m+1,m+2} + \Phi_2 + \Phi_4 + \Phi_5. \tag{7.67}$$

Each of the error terms $\Phi_2$, $\Phi_4$, $\Phi_5$, by their respective definitions, come from manipulating time-averages of $\mathscr{C}^{\mathbf{X},-}(\mathfrak{g})$ on time-scales $\mathfrak{t}_{\mathrm{av},m}$ or $\mathfrak{t}_{\mathrm{av},m+1}$ by putting on indicator functions which all give either an upper bound cutoff or a lower bound cutoff for said time-averages. Thus, as we make precise shortly, these error terms will all eventually be treated via Corollary 5.12.

We now move to analysis. We first define $\Psi_{\mathfrak{u}} \doteq \bar{\mathfrak{P}}^N(\Phi_{\mathfrak{u}}\bar{\mathfrak{Z}}^N)$ and observe, via the proof of the Markov inequality, the bound

$$\mathbf{P}\left(\|\bar{\mathfrak{Z}}^N\|_{\mathfrak{t}_{\mathrm{st}};\mathbb{T}_N}^{-1}\sum_{\mathfrak{u}=2,4,5}\|\Psi_{\mathfrak{u}}\|_{\mathfrak{t}_{\mathrm{st}};\mathbb{T}_N} \gtrsim N^{-\frac{1}{2}-\beta_{\mathrm{univ},2}}\right)$$
$$\lesssim N^{\frac{1}{2}+\beta_{\mathrm{univ},2}}\sum_{\mathfrak{u}=2,4,5}\mathbf{E}[\|\bar{\mathfrak{Z}}^N\|_{\mathfrak{t}_{\mathrm{st}};\mathbb{T}_N}^{-1}\|\Psi_{\mathfrak{u}}\|_{\mathfrak{t}_{\mathrm{st}};\mathbb{T}_N}]. \tag{7.68}$$

Given (7.67), we treat each of the three expectations on the RHS of (7.68). Let us first reemphasize that all three $\Psi_{\mathfrak{u}}$-terms will be turned into space-time averages of pseudo-gradients with both upper and lower bound cutoffs. We then use Corollary 5.12.

- We first treat $\mathfrak{u} = 2$. Recalling $\Psi_2 = \bar{\mathfrak{P}}^N(\Phi_2 \cdot \bar{\mathfrak{Z}}^N)$ and $\Phi_2$ in (7.53), via the triangle inequality we get the deterministic bound

$$|\Psi_2| \leqslant \widetilde{\sum}_{\mathfrak{l}=0}^{\mathfrak{t}_{\mathrm{av},\mathfrak{m}_+}\mathfrak{t}_{\mathrm{av},m+2}^{-1}-1} \widetilde{\sum}_{\mathfrak{k}=0}^{\mathfrak{t}_{\mathrm{av},m+2}\mathfrak{t}_{\mathrm{av},m+1}^{-1}-1} \bar{\mathfrak{P}}_{T,x}^N(|\Gamma_{\mathfrak{l},\mathfrak{k},2}|) \cdot \|\bar{\mathfrak{Z}}^N\|_{\mathfrak{t}_{\mathrm{st}};\mathbb{T}_N}. \tag{7.69}$$

Thus, from (7.69) and pushing all $\|\|\|_{t_{st};\mathbb{T}_N}$-norms to $\|\|\|_{1;\mathbb{T}_N}$, we get the following in which we bound averages by suprema:

$$
\mathbf{E}[\|\bar{\mathfrak{Z}}^N\|_{t_{st};\mathbb{T}_N}^{-1}\|\Psi_2\|_{t_{st};\mathbb{T}_N}]
$$

$$
\leqslant \widetilde{\sum}_{\mathfrak{l}=0}^{t_{av,m+}t_{av,m+2}^{-1}-1}\widetilde{\sum}_{\mathfrak{k}=0}^{t_{av,m+2}t_{av,m+1}^{-1}-1}\mathbf{E}\|\bar{\mathfrak{P}}_{T,x}^N(|\Gamma_{\mathfrak{l},\mathfrak{k},2}|)\|_{t_{st};\mathbb{T}_N} \tag{7.70}
$$

$$
\leqslant \sup_{\mathfrak{l}=0,\ldots,t_{av,m+}t_{av,m+2}^{-1}-1}\sup_{\mathfrak{k}=0,\ldots,t_{av,m+2}t_{av,m+1}^{-1}-1}\mathbf{E}\|\bar{\mathfrak{P}}_{T,x}^N(|\Gamma_{\mathfrak{l},\mathfrak{k},2}|)\|_{1;\mathbb{T}_N}. \tag{7.71}
$$

We control these expectations uniformly over $\mathfrak{l}$, $\mathfrak{k}$. We first make the following observation regarding the event $\mathscr{E}_2$ which we defined earlier as the complement of the $\mathscr{E}_1$-event introduced after (7.53). This effectively takes an a priori lower bound for a scale-$t_{av,m+1}$ time-average from $\mathscr{E}_2$ and provides a priori lower bounds for *at least one* of its scale-$t_{av,m}$ "pieces":

$$
\mathbf{1}[\mathscr{E}_2] \leqslant \sum_{\mathfrak{n}=0}^{t_{av,m+1}t_{av,m}^{-1}-1}\mathbf{1}\left(\sup_{0\leqslant t\leqslant t_{av,m}}t\cdot t_{av,m}^{-1}|\mathscr{A}_t^{\mathbf{T},+}\mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{T_{\mathfrak{l},\mathfrak{k},\mathfrak{n}}^m,x})|\geqslant N^{-\beta_m}\right)
$$

$$
\doteq \sum_{\mathfrak{n}=0}^{t_{av,m+1},t_{av,m}^{-1}-1}\mathbf{1}[\mathscr{E}_{2,\mathfrak{n},T}^{\mathfrak{l},\mathfrak{k}}]. \tag{7.72}
$$

Indeed, observe that the dynamic-average $\mathscr{A}^{\mathbf{T},+}\mathscr{C}^{\mathbf{X},-}$ with respect to time-scales $t \leqslant t_{av,m+1}$ is an average of $t_{av,m+1}t_{av,m}^{-1}$-many suitably time-shifted time-averages $\mathscr{A}^{\mathbf{T},+}\mathscr{C}^{\mathbf{X},-}$ that are each defined with respect to time-scales $t \leqslant t_{av,m}$. Thus, if the former is bigger than $N^{-\beta_m}$, at least one of the pieces that it is averaging over must be bigger than $N^{-\beta_m}$ as well. The summation over $\mathfrak{n} = 0, \ldots, t_{av,m+1}t_{av,m}^{-1} - 1$ is a union bound that accounts for which of the aforementioned time-scale $t_{av,m} \in \mathbb{R}_{>0}$ pieces does the job, of which there certainly may be multiple. By the definition of $\Gamma_{\mathfrak{l},\mathfrak{k},2}$ in (7.53), from (7.72) we trade in the $\mathbf{1}[\mathscr{E}_2]$-factor in $\Gamma_{\mathfrak{l},\mathfrak{k},2}$ for the sum of the $\mathbf{1}[\mathscr{E}_{2,\mathfrak{n},T}^{\mathfrak{l},\mathfrak{k}}]$-indicator functions which we then rewrite as an average of these latter indicator functions upon inserting the normalization factor back. Using this and the triangle inequality, we get the following upper bound which controls $\Gamma_{\mathfrak{l},\mathfrak{k},2}$ by controlling the $\mathbf{1}[\mathscr{E}_2]$-factor therein by the sum of the events $\mathbf{1}[\mathscr{E}_{2,\mathfrak{n},T}^{\mathfrak{l},\mathfrak{k}}]$:

$$
\mathbf{E}\|\bar{\mathfrak{P}}_{T,x}^N(|\Gamma_{\mathfrak{l},\mathfrak{k},2}|)\|_{1;\mathbb{T}_N} \leqslant t_{av,m+1}t_{av,m}^{-1}\widetilde{\sum}_{\mathfrak{n}=0}^{t_{av,m+1}t_{av,m}^{-1}-1}\mathbf{E}\|\bar{\mathfrak{P}}_{T,x}^N(|\Gamma_{S_{\mathfrak{l},\mathfrak{k},0}^m,y}^{N,t_{av,m}}|\cdot\mathbf{1}[\mathscr{E}_{2,\mathfrak{n},S}^{\mathfrak{l},\mathfrak{k}}])\|_{1;\mathbb{T}_N}. \tag{7.73}
$$

Observe that $t_{av,m+1}t_{av,m}^{-1} \lesssim N^\varepsilon$ with $\varepsilon \in \mathbb{R}_{>0}$ arbitrarily small but universal as any of these time-scales with adjacent indices are constructed as arbitrarily small but universal powers of $N \in \mathbb{Z}_{>0}$ times each other. This may be verified by an elementary calculation with the definition of these times in Corollary 5.12. Recalling $\Gamma^{N,t_{av,m}}$ as an average over $\Gamma^{N,t_{av,m},\mathfrak{j}}$, (7.73) gives

$$
\mathbf{E}\|\bar{\mathfrak{P}}_{T,x}^N(|\Gamma_{\mathfrak{l},\mathfrak{k},2}|)\|_{1;\mathbb{T}_N}
$$

$$
\lesssim N^\varepsilon\sup_{\mathfrak{n}=0,\ldots,t_{av,m+1}t_{av,m}^{-1}-1}\sup_{\mathfrak{j}=0,\ldots,t_{av,m+1}t_{av,m}^{-1}-1}\mathbf{E}\|\bar{\mathfrak{P}}_{T,x}^N(|\Gamma_{S_{\mathfrak{l},\mathfrak{k},0}^m,y}^{N,t_{av,m},\mathfrak{j}}|\cdot\mathbf{1}[\mathscr{E}_{2,\mathfrak{n},S}^{\mathfrak{l},\mathfrak{k}}])\|_{1;\mathbb{T}_N}. \tag{7.74}
$$

In this final expectation, we first observe that $\Gamma^{N,t_{av,m},\mathfrak{j}}$-term associated to indices $\mathfrak{l}$, $\mathfrak{k}$, $\mathfrak{j}$ is the time-average $\mathscr{A}^{\mathbf{T},+}\mathscr{C}^{\mathbf{X},-}$ evaluated at time $S + \mathfrak{l}t_{av,m+2} + \mathfrak{k}t_{av,m+1} + \mathfrak{j}t_{av,m}$

by definition of $\Gamma^{N,\mathsf{t}_{\mathrm{av},m},\mathsf{j}}$ in the statement of Lemma 7.11. It also carries the upper bound cutoff of $N^{-\beta_m}$ which therefore induces an upper bound cutoff of $N^{-\beta_{m-1}}$ for trivial reasons. Observe now that courtesy of the $\mathscr{E}_{2,\mathfrak{n},S}^{\mathfrak{l},\mathfrak{k}}$-event hitting this $\Gamma^{N,\mathsf{t}_{\mathrm{av},m},\mathsf{j}}$-term it also carries a lower bound cutoff $N^{-\beta_m}$ for the same time-average but now evaluated at the time $S + \mathfrak{l}\mathsf{t}_{\mathrm{av},m+2} + \mathfrak{k}\mathsf{t}_{\mathrm{av},m+1} + \mathfrak{n}\mathsf{t}_{\mathrm{av},m}$. In particular, observe these times are separated by $|\mathsf{j} - \mathfrak{n}|\mathsf{t}_{\mathrm{av},m} \lesssim \mathsf{t}_{\mathrm{av},m+1} \lesssim N^\varepsilon \mathsf{t}_{\mathrm{av},m}$, because $\mathsf{j}, \mathfrak{n} \in \mathbb{Z}_{\geqslant 0}$ satisfy $\mathsf{j}, \mathfrak{n} \leqslant \mathsf{t}_{\mathrm{av},m+1}\mathsf{t}_{\mathrm{av},m}^{-1} - 1$. We may thus use Corollary 5.12 for $\mathsf{t}_{\mathfrak{s}} = \mathfrak{l}\mathsf{t}_{\mathrm{av},m+2} + \mathfrak{k}\mathsf{t}_{\mathrm{av},m+1} + (\mathsf{j}\wedge\mathfrak{n})\mathsf{t}_{\mathrm{av},m}$ with $\mathsf{t}_{N,\varepsilon,m} = (\mathsf{j}\vee\mathfrak{n})\mathsf{t}_{\mathrm{av},m} - (\mathsf{j}\wedge\mathfrak{n})\mathsf{t}_{\mathrm{av},m}$ to get the following for $\beta_{\mathrm{univ}} \in \mathbb{R}_{>0}$ a universal constant. Our application of Corollary 5.12 is used with $\mathfrak{i} = 2$ if the lower-bound cutoff in $\mathbf{1}[\mathscr{E}_{2,\mathfrak{n},S}^{\mathfrak{l},\mathfrak{k}}]$ is shifted ahead of the time-average $\Gamma^{N,\mathsf{t}_{\mathrm{av},m},\mathsf{j}}$ so that $\mathfrak{n} \geqslant \mathsf{j}$ and otherwise $\mathfrak{i} = 1$; the index $\mathfrak{i} \in \{1, 2\}$ here is that in the superscript of the time-average-with-cutoff $\mathscr{C}^{\mathbf{T},+,\mathsf{t}_{N,\varepsilon,m},\mathfrak{i}}$ in Corollary 5.12:

$$\mathbf{E}\|\bar{\mathfrak{P}}_{T,x}^N(|\Gamma_{S_{\mathfrak{l},\mathfrak{k},0}^m,y}^{N,\mathsf{t}_{\mathrm{av},m},\mathsf{j}}| \cdot \mathbf{1}[\mathscr{E}_{2,\mathfrak{n},S}^{\mathfrak{l},\mathfrak{k}}])\|_{1;\mathbb{T}_N} \;\lesssim\; N^{-1/2-\beta_{\mathrm{univ}}}. \tag{7.75}$$

Combining (7.71), (7.74), and (7.75) gives the following if we choose $2\varepsilon \leqslant \beta_{\mathrm{univ}}$ sufficiently small but universal:

$$\mathbf{E}[\|\bar{\mathfrak{Z}}^N\|_{\mathsf{t}_{\mathrm{st}};\mathbb{T}_N}^{-1}\|\Psi_2\|_{\mathsf{t}_{\mathrm{st}};\mathbb{T}_N}] \;\lesssim\; N^{-\frac{1}{2}-\frac{1}{2}\beta_{\mathrm{univ}}}. \tag{7.76}$$

- We move to $\mathfrak{u} = 4$. Following the observation (7.71) made for $\mathfrak{u} = 2$ and recalling $\Psi_4 = \bar{\mathfrak{P}}^N(\Phi_4 \cdot \bar{\mathfrak{Z}}^N)$ with $\Phi_4$ from (7.57), we get the following estimate in which we remove the $\bar{\mathfrak{Z}}^N$-process from $\Psi_4$ via pulling out its $\|\|_{\mathsf{t}_{\mathrm{st}};\mathbb{T}_N}$-norm and dividing it out to get just the $\Gamma_{\mathfrak{l},\mathfrak{k},\mathsf{j},1,2}$-terms in $\Phi_4$ from (7.57); we additionally push the remaining $\|\|_{\mathsf{t}_{\mathrm{st}};\mathbb{T}_N}$-norm to the $\|\|_{1;\mathbb{T}_N}$-norm:

$$\mathbf{E}[\|\bar{\mathfrak{Z}}^N\|_{\mathsf{t}_{\mathrm{st}};\mathbb{T}_N}^{-1}\|\Psi_4\|_{\mathsf{t}_{\mathrm{st}};\mathbb{T}_N}]$$
$$\leqslant \sup_{\mathfrak{l}=0,\ldots,\mathsf{t}_{\mathrm{av},m}+\mathsf{t}_{\mathrm{av},m+2}^{-1}-1} \;\; \sup_{\mathfrak{k}=0,\ldots,\mathsf{t}_{\mathrm{av},m+2}\mathsf{t}_{\mathrm{av},m+1}^{-1}-1} \;\; \sup_{\mathsf{j}=0,\ldots,\mathsf{t}_{\mathrm{av},m+1}\mathsf{t}_{\mathrm{av},m}^{-1}-1} \mathbf{E}\|\bar{\mathfrak{P}}_{T,x}^N(|\Gamma_{\mathfrak{l},\mathfrak{k},\mathsf{j},1,2}|)\|_{1;\mathbb{T}_N}. \tag{7.77}$$

To bound expectations on the RHS of (7.77), recall the definition of $\Gamma_{\mathfrak{l},\mathfrak{k},\mathsf{j},1,2}$. This leads us to studying the event $\mathscr{E}_1$ introduced after (7.53). We make the following observation. We will explain it in detail shortly but it roughly transfers a priori estimates in $\mathscr{E}_1$ after (7.53) for scale-$\mathsf{t}_{\mathrm{av},m+1}$ time-integrals to scale-$\mathsf{t}_{\mathrm{av},m}$ time-integrals by controlling the latter "smaller-scale" integrals by integrals on intervals of the form $T_{\mathfrak{l},\mathfrak{k},0}^m + [0, \mathsf{t}]$ with $\mathsf{t} \leqslant \mathsf{t}_{\mathrm{av},m+1}$ for which we have integral estimates in $\mathscr{E}_1$. Such a "transfer" is performed by a close relative of the fact that we may control an integral on $[1, 2]$ by integrals on $[0, 2]$ and $[0, 1]$. Roughly, we control integrals on scale-$\mathsf{t}_{\mathrm{av},m}$ intervals by referring to the scale-$\mathsf{t}_{\mathrm{av},m+1}$ block/$\mathfrak{k}$-index the scale-$\mathsf{t}_{\mathrm{av},m}$ intervals live in:

$$\mathbf{1}[\mathscr{E}_1] \;\leqslant\; \prod_{\mathfrak{n}=0}^{\mathsf{t}_{\mathrm{av},m+1}\mathsf{t}_{\mathrm{av},m}^{-1}-1} \mathbf{1}\left(\sup_{0\leqslant\mathsf{t}\leqslant\mathsf{t}_{\mathrm{av},m}} \mathsf{t}\cdot\mathsf{t}_{\mathrm{av},m+1}^{-1}|\mathscr{A}_{\mathsf{t}}^{\mathbf{T},+}\mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{T_{\mathfrak{l},\mathfrak{k},\mathfrak{n}}^m},x)| \lesssim N^{-\beta_m}\right) \tag{7.78}$$

$$\leqslant\; \prod_{\mathfrak{n}=0}^{\mathsf{t}_{\mathrm{av},m+1}\mathsf{t}_{\mathrm{av},m}^{-1}-1} \mathbf{1}\left(\sup_{0\leqslant\mathsf{t}\leqslant\mathsf{t}_{\mathrm{av},m}} \mathsf{t}\cdot\mathsf{t}_{\mathrm{av},m}^{-1}|\mathscr{A}_{\mathsf{t}}^{\mathbf{T},+}\mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{T_{\mathfrak{l},\mathfrak{k},\mathfrak{n}}^m},x)| \lesssim N^{-\beta_{m-1}}\right). \tag{7.79}$$

To get (7.78), note the $\mathfrak{t}\mathscr{A}^{\mathbf{T},+}\mathscr{C}^{\mathbf{X},-}$-term in the sup in the indicator function on the RHS of (7.78) is an *unnormalized* integral of $\mathscr{C}^{\mathbf{X},-}$ over $T_{\mathfrak{l},\mathfrak{k},\mathfrak{n}}^m+[0,\mathfrak{t}]$. We control it by integrals on $[T_{\mathfrak{l},\mathfrak{k},0}^m,T_{\mathfrak{l},\mathfrak{k},\mathfrak{n}}^m]$ and $[T_{\mathfrak{l},\mathfrak{k},0}^m,T_{\mathfrak{l},\mathfrak{k},\mathfrak{n}}^m+\mathfrak{t}]$, whose set difference is $T_{\mathfrak{l},\mathfrak{k},\mathfrak{n}}^m+[0,\mathfrak{t}]$, using linearity of the integral with respect to domain of integration. Uniformly over $0\leqslant\mathfrak{t}\leqslant\mathfrak{t}_{\mathrm{av},m}$ these last two integrals are bounded on the event $\mathscr{E}_1$ by definition of $\mathscr{E}_1$ if $0\leqslant\mathfrak{n}\leqslant\mathfrak{t}_{\mathrm{av},m+1}\mathfrak{t}_{\mathrm{av},m}^{-1}-1$ because these two domains of integration $[T_{\mathfrak{l},\mathfrak{k},0}^m,T_{\mathfrak{l},\mathfrak{k},\mathfrak{n}}^m]$ and $[T_{\mathfrak{l},\mathfrak{k},0}^m,T_{\mathfrak{l},\mathfrak{k},\mathfrak{n}}^m+\mathfrak{t}]$ are both of the form $T_{\mathfrak{l},\mathfrak{k},0}^m+[0,\mathfrak{s}]$ with $\mathfrak{s}\leqslant\mathfrak{t}_{\mathrm{av},m+1}$ on which we have integral bounds for $\mathscr{C}^{\mathbf{X},-}$ on $\mathscr{E}_1$. Thus the $N^{-\beta_m}$-cutoff in $\mathscr{E}_1$ gets transferred to the terms inside the indicator functions on the RHS of (7.78) *on the event* $\mathscr{E}_1$. On the other hand, (7.79) follows from (7.78) and $\mathfrak{t}_{\mathrm{av},m}^{-1}\lesssim\mathfrak{t}_{\mathrm{av},m+1}^{-1}N^{-\beta_{m-1}+\beta_m}$ which can be checked using the definitions in Corollary 5.12. In particular, we may trade the $\mathfrak{t}_{\mathrm{av},m+1}^{-1}$-factor within the indicator function on the RHS of (7.78) with the more-singular $\mathfrak{t}_{\mathrm{av},m}^{-1}$-factor if we loosen the upper bound $N^{-\beta_m}$ on the RHS of (7.78) to $N^{-\beta_{m-1}}$ by the previous inequality for $\mathfrak{t}_{\mathrm{av},m+1}^{-1}$ and $\mathfrak{t}_{\mathrm{av},m}^{-1}$.

Looking at the definition of $\Gamma_{\mathfrak{l},\mathfrak{k},j,1,2}$ given after (7.55), observe that it is the time-average $\mathscr{A}^{\mathbf{T},+}\mathscr{C}^{\mathbf{X},-}$ with the a priori lower-bound cutoff of $N^{-\beta_m}$. Moreover, upon taking the $\mathfrak{n}=j$ factor within (7.79), we may impose the upper bound cutoff of $N^{-\beta_{m-1}}$ for the same time average, therefore controlling $\Gamma_{\mathfrak{l},\mathfrak{k},j,1,2}$ by a $\mathscr{C}^{\mathbf{T},+}\mathscr{C}^{\mathbf{X},-}$-term. Thus, Corollary 5.12 implies the following with $\beta_{\mathrm{univ}}\in\mathbb{R}_{>0}$ universal upon choosing $\mathfrak{t}_{\mathfrak{s}}=\mathfrak{l}\mathfrak{t}_{\mathrm{av},m+2}+\mathfrak{k}\mathfrak{t}_{\mathrm{av},m+1}+j\mathfrak{t}_{\mathrm{av},m}$ and $\mathfrak{t}_{N,\varepsilon,m}=0$ and for either $\mathfrak{i}\in\{1,2\}$ since $\mathfrak{t}_{N,\varepsilon,m}=0$:

$$\mathbf{E}\|\bar{\mathfrak{P}}_{T,x}^N(|\Gamma_{\mathfrak{l},\mathfrak{k},j,1,2}|)\|_{1;\mathbb{T}_N}\lesssim N^{-1/2-\beta_{\mathrm{univ}}}.\tag{7.80}$$

Combining this with (7.77) yields the expectation estimate

$$\mathbf{E}[\|\bar{\mathfrak{Z}}^N\|_{\mathfrak{t}_{\mathrm{st}};\mathbb{T}_N}^{-1}\|\Psi_4\|_{\mathfrak{t}_{\mathrm{st}};\mathbb{T}_N}]\lesssim N^{-\frac{1}{2}-\beta_{\mathrm{univ}}}.\tag{7.81}$$

- We are left with $\mathfrak{u}=5$ on the RHS of (7.68). As with (7.71) and (7.77), by definition of $\Phi_5$ given after (7.65), we get the next bound again upon removing the $\bar{\mathfrak{Z}}^N$-process from $\Psi_5=\bar{\mathfrak{P}}^N(\Phi_5\cdot\bar{\mathfrak{Z}}^N)$ by pulling out its $\|\|_{\mathfrak{t}_{\mathrm{st}};\mathbb{T}_N}$-norm and then dividing it out:

$$\mathbf{E}[\|\bar{\mathfrak{Z}}^N\|_{\mathfrak{t}_{\mathrm{st}};\mathbb{T}_N}^{-1}\|\Psi_5\|_{\mathfrak{t}_{\mathrm{st}};\mathbb{T}_N}]$$
$$\leqslant\sup_{\substack{\mathfrak{l}=0,\ldots,\mathfrak{t}_{\mathrm{av},m+}\mathfrak{t}_{\mathrm{av},m+2}^{-1}-1\\\mathfrak{k}=0,\ldots,\mathfrak{t}_{\mathrm{av},m+2}\mathfrak{t}_{\mathrm{av},m+1}^{-1}-1}}\mathbf{E}\|\bar{\mathfrak{P}}_{T,x}^N(\mathbf{1}[\mathscr{E}_1]\mathbf{1}[\mathscr{G}_2]\cdot|\mathscr{A}_{\mathfrak{t}_{\mathrm{av},m+1}}^{\mathbf{T},+}\mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{S_{\mathfrak{l},\mathfrak{k},0}^m,y})|)\|_{1;\mathbb{T}_N}.$$
$$\tag{7.82}$$

The quantity that we are integrating in the heat-operator on the RHS of (7.82), per every pair of indices $\mathfrak{l},\mathfrak{k}$, is a time-average term $\mathscr{A}^{\mathbf{T},+}\mathscr{C}^{\mathbf{X},-}$ that, by definition of $\mathscr{E}_1$, carries an upper bound cutoff of $N^{-\beta_m}$ and, by definition of $\mathscr{G}_2$, carries a lower bound cutoff of $N^{-\beta_{m+1}}$. Thus, the quantity we are integrating in the heat operator on the RHS of (7.82) is a $\mathscr{C}^{\mathbf{T},+}\mathscr{C}^{\mathbf{X},-}$ that is treated in Corollary 5.12. In particular, by Corollary 5.12 with $\mathfrak{t}_{\mathfrak{s}}=\mathfrak{l}\mathfrak{t}_{\mathrm{av},m+2}+\mathfrak{k}\mathfrak{t}_{\mathrm{av},m+1}$ and $\mathfrak{t}_{N,\varepsilon,m}=0$ and either $\mathfrak{i}\in\{1,2\}$, we get

$$\mathbf{E}\|\bar{\mathfrak{P}}_{T,x}^N(\mathbf{1}[\mathscr{E}_1]\mathbf{1}[\mathscr{G}_2]\cdot|\mathscr{A}_{\mathfrak{t}_{\mathrm{av},m+1}}^{\mathbf{T},+}\mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{S_{\mathfrak{l},\mathfrak{k},0}^m,y})|)\|_{1;\mathbb{T}_N}\lesssim N^{-1/2-\beta_{\mathrm{univ}}}.\tag{7.83}$$

From this last estimate combined with (7.82), we get the expectation estimate

$$\mathbf{E}[\|\bar{\mathfrak{Z}}^N\|_{\mathfrak{t}_{\mathrm{st}};\mathbb{T}_N}^{-1}\|\Psi_5\|_{\mathfrak{t}_{\mathrm{st}};\mathbb{T}_N}]\lesssim N^{-\frac{1}{2}-\beta_{\mathrm{univ}}}.\tag{7.84}$$

By (7.68) and (7.76), (7.81), and (7.84), with the required high-probability, for some $\beta_{\mathrm{univ},2} \in \mathbb{R}_{>0}$ universal, we get

$$\|\bar{\mathfrak{z}}^N\|_{\mathfrak{t}_{\mathrm{st}};\mathbb{T}_N}^{-1} \sum\nolimits_{\mathfrak{u}=2,4,5} \|\bar{\mathfrak{P}}_{T,x}^N(\Phi_\mathfrak{u}\bar{\mathfrak{z}}^N)\|_{\mathfrak{t}_{\mathrm{st}};\mathbb{T}_N}$$
$$= \|\bar{\mathfrak{z}}^N\|_{\mathfrak{t}_{\mathrm{st}};\mathbb{T}_N}^{-1} \sum\nolimits_{\mathfrak{u}=2,4,5} \|\Psi_\mathfrak{u}\|_{\mathfrak{t}_{\mathrm{st}};\mathbb{T}_N} \lesssim N^{-\frac{1}{2}-\beta_{\mathrm{univ},2}}. \tag{7.85}$$

Indeed we recall $\Psi_\mathfrak{u} = \bar{\mathfrak{P}}_{T,x}^N(\Phi_\mathfrak{u}\bar{\mathfrak{z}}^N)$ by definition given right after the equation (7.67). Applying (7.85) with (7.67) finishes the proof for the proposed estimate that is explicitly written in the statement of Lemma 7.11. To get the estimate after replacement (7.3) and a replacement of time-scales/exponents by those in Corollary 5.13, it suffices to write these replacements in formally and then employ an identical argument. In particular, the details of what we are time-averaging and cutting off are not relevant, as are the details of the exponents as long as we use the estimates in Corollary 5.13 instead of those in Corollary 5.12. □

*Proof of Lemma 7.12.* In principle we have already written a proof for Lemma 7.12 in the middle of the proof for Lemma 7.11. However we provide details anyway for clarity. The architecture for the proof of Lemma 7.12 is similar to that of Lemma 7.11. We first relate $\Gamma^{N,1,2}$ within the RHS of the proposed estimate to the space-time average $\mathscr{A}^{\mathbf{T},+}\mathscr{C}^{\mathbf{X},-}$ within the LHS of the same proposed estimate through exact identities, and then we estimate the error terms in such relation/the quantities in these exact identities. The approach that we take towards relating these two quantities from the previous sentence/writing exact identities is again by upgrading in time-scale $\mathfrak{t}_{\mathrm{av},1} \to \mathfrak{t}_{\mathrm{av},2}$ and cutoffs then estimating errors which turn out to be time-averages of $\mathscr{C}^{\mathbf{X},-}(\mathfrak{g})$ with upper/lower bound cutoffs. First, as in the proof of Lemma 7.11 we introduce notation for time-shifts at two time-scales:

- Provided any $\mathfrak{l}, \mathfrak{j} \in \mathbb{Z}_{\geq 0}$ and $T \in \mathbb{R}_{\geq 0}$ define the shift $T_{\mathfrak{l},\mathfrak{j}} \overset{\bullet}{=} T + \mathfrak{l}\mathfrak{t}_{\mathrm{av},2} + \mathfrak{j}\mathfrak{t}_{\mathrm{av},1}$ with respect to time-scales $\mathfrak{t}_{\mathrm{av},2}$, $\mathfrak{t}_{\mathrm{av},1}$, respectively.

Recall the $\mathscr{A}^{\mathbf{T},+}\mathscr{C}^{\mathbf{X},-}(\mathfrak{g})$-term is a time-average with respect to the "maximal" scale $\mathfrak{t}_{\mathrm{av},\mathfrak{m}_+} \in \mathbb{R}_{>0}$. With more explanation after,

$$\mathscr{A}_{\mathfrak{t}_{\mathrm{av},\mathfrak{m}_+}}^{\mathbf{T},+}\mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{T,x}) = \widetilde{\sum}_{\mathfrak{l}=0}^{\mathfrak{t}_{\mathrm{av},\mathfrak{m}_+}\mathfrak{t}_{\mathrm{av},2}^{-1}-1} \widetilde{\sum}_{\mathfrak{j}=0}^{\mathfrak{t}_{\mathrm{av},2}\mathfrak{t}_{\mathrm{av},1}^{-1}-1} \mathscr{A}_{\mathfrak{t}_{\mathrm{av},1}}^{\mathbf{T},+}\mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{T_{\mathfrak{l},\mathfrak{j}},x}). \tag{7.86}$$

Indeed, we interpret the LHS of (7.86) to be the average of suitably time-shifted dynamic-averages with respect to the smaller time-scale $\mathfrak{t}_{\mathrm{av},2}$ in Corollary 5.12, and we further decompose each of these time averages with respect to $\mathfrak{t}_{\mathrm{av},2}$ into more suitably time-shifted time averages with respect to the smallest time-scale $\mathfrak{t}_{\mathrm{av},1}$. As $\mathscr{C}^{\mathbf{X},-}$ has the upper bound cutoff $N^{-\beta_0}$ by definition, where $\beta_0 = \frac{1}{2}\beta_X - \varepsilon_{X,2}$ with $\varepsilon_{X,2} \geq \frac{1}{2}\varepsilon_{X,1}$ arbitrarily small but universal in Corollary 5.12, for free by (7.86) we get

$$\mathscr{A}_{\mathfrak{t}_{\mathrm{av},\mathfrak{m}_+}}^{\mathbf{T},+}\mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{T,x}) = \widetilde{\sum}_{\mathfrak{l}=0}^{\mathfrak{t}_{\mathrm{av},\mathfrak{m}_+}\mathfrak{t}_{\mathrm{av},2}^{-1}-1} \widetilde{\sum}_{\mathfrak{j}=0}^{\mathfrak{t}_{\mathrm{av},2}\mathfrak{t}_{\mathrm{av},1}^{-1}-1} \mathscr{A}_{\mathfrak{t}_{\mathrm{av},1}}^{\mathbf{T},+}\mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{T_{\mathfrak{l},\mathfrak{j}},x}) \cdot \mathbf{1}[\mathscr{E}_{\mathfrak{l},\mathfrak{j}}] \tag{7.87}$$

provided that we have defined the following event. We clarify that the supremum in the following event is the supremum over *integrals* on time-scales $0 \leq \mathfrak{t} \leq \mathfrak{t}_{\mathrm{av},1}$ of the cutoff spatial average $\mathscr{C}^{\mathbf{X},-}$ then reweighted by $\mathfrak{t}_{\mathrm{av},1}^{-1}$. In particular, every $\mathfrak{t}$-indexed term in the

supremum below is bounded by $|\mathscr{C}^{\mathbf{X},-}| \lesssim N^{-\beta_0}$ as every $\mathfrak{t}$-indexed term time-averages $\mathscr{C}^{\mathbf{X},-}$ on a time-scale $\mathfrak{t} \leqslant \mathfrak{t}_{\mathrm{av},1}$:

$$\mathbf{1}[\mathscr{E}_{\mathfrak{l},\mathfrak{j}}] \stackrel{\bullet}{=} \mathbf{1}\left(\sup_{0 \leqslant \mathfrak{t} \leqslant \mathfrak{t}_{\mathrm{av},1}} \mathfrak{t} \cdot \mathfrak{t}_{\mathrm{av},1}^{-1} |\mathscr{A}_{\mathfrak{t}}^{\mathbf{T},+} \mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{T_{\mathfrak{l},\mathfrak{j}},x})| \lesssim N^{-\beta_0}\right). \tag{7.88}$$

Observe that the $\mathfrak{l},\mathfrak{j}$-term in the sum from the RHS of (7.87) is *almost* the $\Gamma^{N,\mathfrak{t}_{\mathrm{av},1},\mathfrak{j}}$-term, except the cutoff defining $\mathscr{E}_{\mathfrak{l},\mathfrak{j}}$ is of order $N^{-\beta_0}$ rather than $N^{-\beta_1}$. To this end, we define the following event $\mathscr{G}_{\mathfrak{l},\mathfrak{j},1}$ with a stricter constraint giving us the cutoff we need:

$$\mathbf{1}[\mathscr{G}_{\mathfrak{l},\mathfrak{j},1}] \stackrel{\bullet}{=} \mathbf{1}\left(\sup_{0 \leqslant \mathfrak{t} \leqslant \mathfrak{t}_{\mathrm{av},1}} \mathfrak{t} \cdot \mathfrak{t}_{\mathrm{av},1}^{-1} |\mathscr{A}_{\mathfrak{t}}^{\mathbf{T},+} \mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{T_{\mathfrak{l},\mathfrak{j}},x})| \leqslant N^{-\beta_1}\right). \tag{7.89}$$

Let $\mathscr{G}_{\mathfrak{l},\mathfrak{j},2}$ denote the complement of $\mathscr{G}_{\mathfrak{l},\mathfrak{j},1}$. Observe the containment $\mathscr{G}_{\mathfrak{l},\mathfrak{j},1} \subseteq \mathscr{E}_{\mathfrak{l},\mathfrak{j}}$ that follows because $\beta_1 \geqslant \beta_0$, so $\mathscr{G}_{\mathfrak{l},\mathfrak{j},1}$ imposes a stricter constraint than $\mathscr{E}_{\mathfrak{l},\mathfrak{j}}$. Thus $\mathbf{1}[\mathscr{E}_{\mathfrak{l},\mathfrak{j}}] = \mathbf{1}[\mathscr{G}_{\mathfrak{l},\mathfrak{j},1}] + \mathbf{1}[\mathscr{G}_{\mathfrak{l},\mathfrak{j},2}]\mathbf{1}[\mathscr{E}_{\mathfrak{l},\mathfrak{j}}]$. We apply this decomposition to (7.87) and deduce

$$\mathscr{A}_{\mathfrak{t}_{\mathrm{av},\mathfrak{m}_+}}^{\mathbf{T},+} \mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{T,x}) = \sum_{\mathfrak{l}=0}^{\widetilde{\mathfrak{t}_{\mathrm{av},\mathfrak{m}_+} \mathfrak{t}_{\mathrm{av},2}^{-1}-1}} \sum_{\mathfrak{j}=0}^{\widetilde{\mathfrak{t}_{\mathrm{av},2} \mathfrak{t}_{\mathrm{av},1}^{-1}-1}} \Phi_{1,\mathfrak{l},\mathfrak{j}}$$
$$+ \sum_{\mathfrak{l}=0}^{\widetilde{\mathfrak{t}_{\mathrm{av},\mathfrak{m}_+} \mathfrak{t}_{\mathrm{av},2}^{-1}-1}} \sum_{\mathfrak{j}=0}^{\widetilde{\mathfrak{t}_{\mathrm{av},2} \mathfrak{t}_{\mathrm{av},1}^{-1}-1}} \Phi_{2,\mathfrak{l},\mathfrak{j}} \stackrel{\bullet}{=} \Phi_1 + \Phi_2 \tag{7.90}$$

where $\Phi_{1,\mathfrak{l},\mathfrak{j}}$ replaces the free $\mathscr{E}_{\mathfrak{l},\mathfrak{j}}$-cutoff of $N^{-\beta_0}$ with the $\mathscr{G}_{\mathfrak{l},\mathfrak{j},1}$-cutoff of $N^{-\beta_1}$, and $\Phi_{2,\mathfrak{l},\mathfrak{j}}$ is the resulting error in this upgrade:

$$\Phi_{1,\mathfrak{l},\mathfrak{j}} \stackrel{\bullet}{=} \mathscr{A}_{\mathfrak{t}_{\mathrm{av},1}}^{\mathbf{T},+} \mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{T_{\mathfrak{l},\mathfrak{j}},x}) \cdot \mathbf{1}[\mathscr{G}_{\mathfrak{l},\mathfrak{j},1}] \quad \text{and} \quad \Phi_{2,\mathfrak{l},\mathfrak{j}}$$
$$\stackrel{\bullet}{=} \mathscr{A}_{\mathfrak{t}_{\mathrm{av},1}}^{\mathbf{T},+} \mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}_{T_{\mathfrak{l},\mathfrak{j}},x}) \cdot \mathbf{1}[\mathscr{E}_{\mathfrak{l},\mathfrak{j}}]\mathbf{1}[\mathscr{G}_{\mathfrak{l},\mathfrak{j},2}]. \tag{7.91}$$

We observe that $\Phi_1 = \Gamma^{N,1,2}$ just by its definition within the statement of Lemma 7.12 as it is an average of suitably time-shifted time-averages with respect to time-scale $\mathfrak{t}_{\mathrm{av},1}$ each with an upper bound cutoff of $N^{-\beta_1}$. By (7.90) and this observation we get

$$|\bar{\mathfrak{P}}_{T,x}^N(N^{\frac{1}{2}} \mathscr{A}_{\mathfrak{t}_{\mathrm{av},\mathfrak{m}_+}}^{\mathbf{T},+} \mathscr{C}_{N^{\beta_X}}^{\mathbf{X},-}(\mathfrak{g}) \cdot \bar{\mathfrak{z}}^N)| \leqslant |\bar{\mathfrak{P}}_{T,x}^N(N^{\frac{1}{2}} \Gamma^{N,1,2} \cdot \bar{\mathfrak{z}}^N)|$$
$$+ \|\bar{\mathfrak{z}}^N\|_{\mathfrak{t}_{\mathrm{st}};\mathbb{T}_N} \bar{\mathfrak{P}}_{T,x}^N(N^{\frac{1}{2}}|\Phi_2|). \tag{7.92}$$

It suffices to show that with the required high-probability, we can control the second term in (7.92):

$$\|\bar{\mathfrak{P}}_{T,x}^N(N^{\frac{1}{2}}|\Phi_2|)\|_{\mathfrak{t}_{\mathrm{st}};\mathbb{T}_N} \lesssim N^{-\beta_{\mathrm{univ},2}}. \tag{7.93}$$

We will do this using a procedure that is a simplified or specialized version of estimates for $\Psi_{\mathfrak{u}}$ with $\mathfrak{u} \in \{2,4,5\}$ in the proof for Lemma 7.11. In particular, we will control $|\Phi_2|$ in terms of quantities of the form $\mathscr{C}^{\mathbf{T},+} \mathscr{C}^{\mathbf{X},-}$ that are estimated in Corollary 5.12. Recall the $\mathfrak{l},\mathfrak{j}$-term in the sum defining $\Phi_2$ is $\Phi_{2,\mathfrak{l},\mathfrak{j}}$. To establish the previous high-probability

statement, we employ the Markov inequality as with the estimate (7.68) in the proof of
Lemma 7.11 and then unfold the definition of $\Phi_2$. As $t_{st} \leqslant 2$, we get

$$\mathbf{P}\left(\|\bar{\mathfrak{P}}^N_{T,x}(N^{\frac{1}{2}}|\Phi_2|)\|_{t_{st};\mathbb{T}_N} \gtrsim N^{-\beta_{univ,2}}\right)$$

$$\leqslant N^{\beta_{univ,2}}\mathbf{E}\|\bar{\mathfrak{P}}^N_{T,x}(N^{\frac{1}{2}}|\Phi_2|)\|_{2;\mathbb{T}_N} \tag{7.94}$$

$$\leqslant \overbrace{\sum_{\mathfrak{l}=0}^{t_{av,m_+}t_{av,2}^{-1}-1}}\overbrace{\sum_{j=0}^{t_{av,2}t_{av,1}^{-1}-1}} N^{\beta_{univ,2}}\mathbf{E}\|\bar{\mathfrak{P}}^N_{T,x}(N^{\frac{1}{2}}|\Phi_{2,\mathfrak{l},j}|)\|_{2;\mathbb{T}_N} \tag{7.95}$$

$$\leqslant \sup_{\mathfrak{l}=0,...,t_{av,m_+}t_{av,2}^{-1}-1}\sup_{j=0,...,t_{av,2}t_{av,1}^{-1}-1} N^{\beta_{univ,2}}\mathbf{E}\|\bar{\mathfrak{P}}^N_{T,x}(N^{\frac{1}{2}}|\Phi_{2,\mathfrak{l},j}|)\|_{2;\mathbb{T}_N}. \tag{7.96}$$

Courtesy of (7.96), we want to control the expectations appearing in (7.96) uniformly
in $\mathfrak{l}, j$ as this would provide the estimate (7.93). To this end, note that $\Phi_{2,\mathfrak{l},j}$, by its
definition as the $(\mathfrak{l}, j)$-summand in the summation for $\Phi_2$ that we wrote immediately
after the decomposition (7.90), is a space-time average $\mathscr{A}^{\mathbf{T},+}\mathscr{C}^{\mathbf{X},-}$ equipped with the
upper bound cutoff $N^{-\beta_0}$ courtesy of the event $\mathscr{E}_{\mathfrak{l},j}$ inside $\Phi_{2,\mathfrak{l},j}$ as well as a lower bound
cutoff of $N^{-\beta_1}$ courtesy of the $\mathscr{G}_{\mathfrak{l},j,2}$-event inside $\Phi_{2,\mathfrak{l},j}$. Thus, we may use the $m = 1$
estimate in Corollary 5.12 with $t_{\mathfrak{s}} = \mathfrak{l}t_{av,2} + jt_{av,1}$ and $t_{N,\varepsilon,m} = 0$ and either index
$i \in \{1, 2\}$ to get, for $\beta_{univ} \in \mathbb{R}_{>0}$ universal,

$$\sup_{\mathfrak{l}=0,...,t_{av,m_+}t_{av,2}^{-1}-1}\sup_{j=0,...,t_{av,2}t_{av,1}^{-1}-1}N^{\beta_{univ,2}}\mathbf{E}\|\bar{\mathfrak{P}}^N_{T,x}(N^{\frac{1}{2}}|\Phi_{2,\mathfrak{l},j}|)\|_{2;\mathbb{T}_N}$$

$$\lesssim N^{-\beta_{univ}+\beta_{univ,2}}. \tag{7.97}$$

Choosing $\beta_{univ,2} = \frac{1}{2}\beta_{univ}$ and plugging the previous estimate into (7.96) establishes
(7.93) with the required high-probability and completes the proof of the proposed bound
that is explicitly written in the statement of Lemma 7.12. To prove the estimate *after*
the replacement (7.3) and the replacement of time-scales and exponents by those in
Corollary 5.13, we follow the remarks given at the end of the proof of Lemma 7.11. In
particular, we make the aforementioned replacements *in the proof* and use the estimates
in Corollary 5.13 as opposed to those in Corollary 5.12.                                     □

## 8. Proof of Theorem 1.7

The proof of Theorem 1.7 will follow the strategy that is outlined in the bullet points
below.

- We compare the microscopic Cole-Hopf transform $\mathfrak{Z}^N$ to an auxiliary space-time
  process we denote by $\mathfrak{Y}^N$. This space-time process is defined by the same stochastic
  integral equation as $\bar{\mathfrak{Z}}^N$ in Sect. 4 but tossing out $\bar{\Phi}^{N,2}$ and replacing all $\bar{\mathfrak{Z}}^N$-terms by
  $\mathfrak{Y}^N$. In particular, besides microscopic SHE terms, the $\mathfrak{Y}^N$-equation only has data
  of weakly vanishing functionals.
- After the comparison between $\mathfrak{Z}^N$ and $\mathfrak{Y}^N$, we prove that $\mathfrak{Y}^N$ converges to the
  solution of SHE in the Skorokhod space $\mathbf{D}_1$. To this end, recall that $\mathfrak{Y}^N$ is defined
  by a microscopic version of the SHE plus some extra weakly vanishing data. In
  particular, proof of convergence to SHE for $\mathfrak{Y}^N$ will follow the analysis of weakly
  vanishing terms in [10]. Because most of the work is done in [10], we only present
  necessary adjustments for the proof of convergence to SHE for $\mathfrak{Y}^N$.

We first define the $\mathfrak{Y}^N$-process; recall $\mathbb{T}_N$ in the beginning of Sect. 4 and $\bar{\nabla}^! = N\bar{\nabla}$ where $\bar{\nabla}$ is gradient on the *torus* $\mathbb{T}_N$.

**Definition 8.1.** Define $\mathfrak{Y}^N$ as the solution to the following stochastic equation on $\mathbb{R}_{\geqslant 0} \times \mathbb{T}_N$; recall $\chi$ from Definition 4.3:

$$
\begin{aligned}
\mathfrak{Y}_{T,x}^N &= \bar{\mathfrak{P}}_{T,x}^{N,\mathbf{X}}(\bar{\mathfrak{Z}}_{0,\bullet}^N) + \bar{\mathfrak{P}}_{T,x}^N(\mathfrak{Y}^N \mathrm{d}\xi^N) + \bar{\mathfrak{P}}_{T,x}^N(\bar{\Phi}^{N,4}) \\
&= \bar{\mathfrak{P}}_{T,x}^{N,\mathbf{X}}(\chi_\bullet \mathfrak{Z}_{0,\bullet}^N) + \bar{\mathfrak{P}}_{T,x}^N(\mathfrak{Y}^N \mathrm{d}\xi^N) + \bar{\mathfrak{P}}_{T,x}^N(\bar{\Phi}^{N,4}). \quad (8.1)
\end{aligned}
$$

As in Remark 4.4 $\mathfrak{Y}^N \mathrm{d}\xi^N$ is the martingale differential corresponding to the Poisson process with jumps of $\mathrm{d}M$ in (2.4) in [10] then scaled by $\mathfrak{Y}^N$ at the same space-time point. Also $\bar{\Phi}^{N,4}$ is the following analog of $\bar{\Phi}^{N,3}$ from Definition 4.3:

$$
\bar{\Phi}_{T,x}^{N,4} \overset{\bullet}{=} \mathfrak{w}_{T,x} \mathfrak{Y}_{T,x}^N + \sum_{k=-2\mathfrak{m}}^{2\mathfrak{m}} c_k \bar{\nabla}_k^!(\mathfrak{w}_{T,x}^k \mathfrak{Y}_{T,x}^N). \quad (8.2)
$$

A function on $\mathbb{T}_N$ or $\mathbb{R}_{\geqslant 0} \times \mathbb{T}_N$, if it is not already given by restricting a function on $\mathbb{Z}$ or $\mathbb{R}_{\geqslant 0} \times \mathbb{Z}$ to the torus $\mathbb{T}_N$, is implicitly lifted to a function on $\mathbb{Z}$ or $\mathbb{R}_{\geqslant 0} \times \mathbb{Z}$ via periodic extension outside $\mathbb{T}_N$. We then lift to $\mathbb{R}$ or $\mathbb{R}_{\geqslant 0} \times \mathbb{R}$ via linear interpolation. This applies to $\mathfrak{Y}^N$, for example. The main ingredients that we briefly overviewed above are stated precisely as follows.

**Proposition 8.2.** *Consider any compact set $\mathbb{K} \subseteq \mathbb{R}$ and define its microscopic coordinates $\mathbb{K}_N \overset{\bullet}{=} N\mathbb{K} \cap \mathbb{Z}$. There exist universal constants $\beta_{\mathrm{univ},1}, \beta_{\mathrm{univ},2} \in \mathbb{R}_{>0}$ such that outside an event of probability at most $N^{-\beta_{\mathrm{univ},1}}$ times uniformly bounded factors, we have*

$$
\|\mathfrak{Z}^N - \mathfrak{Y}^N\|_{1;\mathbb{K}_N} \lesssim_{\mathbb{K}} N^{-\beta_{\mathrm{univ},2}}. \quad (8.3)
$$

**Proposition 8.3.** *The process $\mathfrak{Y}_{\bullet,N\bullet}^N$ is tight in $\mathbf{D}_1$ and all limit points are the law of the solution of* SHE *with initial data $\mathfrak{Z}_{0,\bullet}$.*

**Lemma 8.4.** *Consider processes $\mathfrak{X}^{N,1}$ and $\mathfrak{X}^{N,2}$ in $\mathbf{D}_1$. Suppose $\mathfrak{X}^{N,2}$ is tight in $\mathbf{D}_1$ with limit $\mathfrak{X}^{\infty,2}$, and there are universal constants $\beta_{\mathrm{univ},1}, \beta_{\mathrm{univ},2} \in \mathbb{R}_{>0}$ so that for any compact set $\mathbb{K} \subseteq \mathbb{R}$, outside an event with probability at most $N^{-\beta_{\mathrm{univ},1}}$ times universal factors,*

$$
\|\mathfrak{X}^{N,1} - \mathfrak{X}^{N,2}\|_{1;\mathbb{K}} \lesssim_{\mathbb{K}} N^{-\beta_{\mathrm{univ},2}}. \quad (8.4)
$$

*The sequence $\mathfrak{X}^{N,1}$ is also tight and it converges to the same limit $\mathfrak{X}^{\infty,2}$. All limits are at the level of probability measures on $\mathbf{D}_1$.*

*Proof of Theorem 1.7.* Observe Proposition 8.2 gives a power-saving bound in $N \in \mathbb{Z}_{>0}$ on the difference of spatially-rescaled versions of $\mathfrak{Z}^N$ and $\mathfrak{Y}^N$ uniformly over a discretization of the compact space-time set $[0, 1] \times \mathbb{K}$. Because both processes extend *by linear interpolation* outside such a discretization to the whole space-time set $[0, 1] \times \mathbb{K}$, the difference between the spatially rescaled versions of $\mathfrak{Z}^N$ and $\mathfrak{Y}^N$ on the set $[0, 1] \times \mathbb{K}$ is controlled by the same difference but restricting to the aforementioned discretization. Thus we deduce Theorem 1.7 from the tightness and convergence of $\mathfrak{Y}^N$ within Proposition 8.3 combined with the comparison result in Lemma 8.4. This completes the proof. $\quad\square$

The rest of the section is organized as follows. We will focus on a proof of Proposition 8.2 because it takes up the bulk of the section since it requires a few auxiliary lemmas and an elementary but detailed "high-probability" pathwise argument. We will present the auxiliary lemmas, but we will defer their proofs to the end of this section in order to avoid obscuring key arguments and since their proofs are based on standard methods or are a lot of technical work. On the other hand, the proof of Proposition 8.3 is relatively short as it is the focus of [10]. Lemma 8.4 is an elementary topological result proved at the end of this section.

*8.1. Proof of Proposition 8.2.* Let us inherit the notation from Proposition 8.2 throughout this subsection. We observe that the intersection of any two events whose complements hold with probability at most $N^{-\beta_{\mathrm{univ},1}}$ times uniformly bounded factors also has complement holding with probability at most $N^{-\beta_{\mathrm{univ},1}}$ times uniformly bounded factors. Thus, Proposition 8.2 follows by the following pair of comparison estimates courtesy of this observation and the triangle inequality.

**Lemma 8.5.** *Consider a compact set $\mathbb{K} \subseteq \mathbb{R}$ and define its microscopic coordinates $\mathbb{K}_N \doteq N\mathbb{K} \cap \mathbb{Z}$. There exist universal constants $\beta_{\mathrm{univ},1}, \beta_{\mathrm{univ},2} \in \mathbb{R}_{>0}$ such that outside an event of probability at most $N^{-\beta_{\mathrm{univ},1}}$ times uniformly bounded factors, we have*

$$\|\bar{\mathfrak{Z}}^N - \mathfrak{Y}^N\|_{1;\mathbb{K}_N} \lesssim_{\mathbb{K}} \|\bar{\mathfrak{Z}}^N - \mathfrak{Y}^N\|_{1;\mathbb{T}_N} \lesssim N^{-\beta_{\mathrm{univ},2}}. \tag{8.5}$$

**Lemma 8.6.** *Lemma 8.5 holds upon replacing $\mathfrak{Y}^N$ by $\mathfrak{Z}^N$ and forgetting the middle inequality/$\|\|_{1;\mathbb{T}_N}$-term.*

Lemma 8.6 is almost direct from Proposition 4.6. The difference is that Lemma 8.6 uses $\|\|_{1;\mathbb{K}_N}$ and Proposition 4.6 uses the discretized norm $[]_{1;\mathbb{K}_N}$. We will use short-time stochastic continuity to bootstrap norms similar to the proof of Lemma 6.4. We also use Lemma 8.5 to prove Lemma 8.6; the proof of Lemma 8.5 will be independent of Lemma 8.6.

The proof of Lemma 8.6 will be deferred to a later section on technical estimates as it will require no preliminary ingredients. The proof of Lemma 8.5, however, is more involved and requires a few preliminaries. It is based on the following outline.

- The first estimate in Lemma 8.5 follows by noting $\mathbb{T}_N$ is centered at $0 \in \mathbb{Z}$ and $|\mathbb{T}_N| \gg N$ and thus $\mathbb{K}_N \subseteq \mathbb{T}_N$ if $N \gtrsim_{\mathbb{K}} 1$.
- Consider Proposition 7.1 and suppose that $\bar{\mathfrak{Z}}^N$ were bounded; a fortiori it resembles the continuous SHE solution. Since the stochastic equations for $\bar{\mathfrak{Z}}^N$ and $\mathfrak{Y}^N$ differ only in the $\bar{\Phi}^{N,2}$-term of interest in Proposition 7.1, with high-probability we would deduce this term is negligible in the large-$N$ limit and Lemma 8.5 would follow by linear theory for equations on $\mathbb{T}_N$.
- The previous remark depends on the assumption $\bar{\mathfrak{Z}}^N$ is uniformly bounded on $\mathbb{T}_N$. This is not clear, nor will we try to show it. Note it is actually enough for $\bar{\mathfrak{Z}}^N$ to be bounded by small powers of $N$ with high-probability. We will show this for $\mathfrak{Y}^N$ through moment estimates and union bound as in the proof of Proposition 6.2. As $\mathfrak{Y}^N$ and $\mathfrak{Z}^N$ are supposed to be close we expect this to hold true for $\mathfrak{Z}^N$ as well. To implement this, we use a continuity argument. In what follows $\varepsilon \in \mathbb{R}_{>0}$ is thought of as small.
- Consider the first time $\mathfrak{t}_{2\varepsilon}$ before 1 that $\bar{\mathfrak{Z}}^N$ exceeds $N^{2\varepsilon}$ in the $\|\|_{\mathfrak{t}_{2\varepsilon};\mathbb{T}_N}$-norm. Proposition 7.1 tells us that with high-probability the $\bar{\Phi}^{N,2}$-term is negligible, so

$\bar{\mathfrak{Z}}^N \approx \mathfrak{Y}^N$ until time $\mathfrak{t}_{2\varepsilon}$. With high-probability we know $\mathfrak{Y}^N$ is at most $N^\varepsilon$ in $\|\|_{\mathfrak{t}_{2\varepsilon};\mathbb{T}_N}$-norm, thus the same is true for $\bar{\mathfrak{Z}}^N$. This means we can wait *after* $\mathfrak{t}_{2\varepsilon}$ for $\bar{\mathfrak{Z}}^N$ to exceed $N^{2\varepsilon}$. This propagates $\mathfrak{t}_{2\varepsilon} \to 1$ with high-probability.

- We emphasize our estimates for the $\bar{\Phi}^{N,2}$-term in Proposition 7.1 and for $\mathfrak{Y}^N$ will not depend on the random time $\mathfrak{t}_{2\varepsilon}$, so even though this scheme may take a very large number of iterations, our comparison estimates for $\bar{\mathfrak{Z}}^N$ and $\mathfrak{Y}^N$ stay the same.

We now make the previous outline precise and prove Lemma 8.5. This begins with the following construction of random times.

**Definition 8.7.** Consider any $\varepsilon_{\mathrm{st}} \in \mathbb{R}_{>0}$ arbitrarily small but universal. Define the random time $\mathfrak{t}_{1,\mathrm{st}} \overset{\bullet}{=} \mathfrak{t}_{1,\mathrm{st},1} \wedge \mathfrak{t}_{1,\mathrm{st},2}$, where, for some universal constant $\beta_{\mathrm{univ}} \in \mathbb{R}_{>0}$, we define the following in which the implied constant for $\mathfrak{t}_{1,\mathrm{st},2}$ is large but universal:

$$\mathfrak{t}_{1,\mathrm{st},1} \overset{\bullet}{=} \inf\left\{ \mathfrak{t} \in [0, 1] : \|\bar{\mathfrak{Z}}^N\|_{\mathfrak{t};\mathbb{T}_N} \geqslant N^{2\varepsilon_{\mathrm{st}}} \right\} \wedge 1 \tag{8.6a}$$

$$\mathfrak{t}_{1,\mathrm{st},2} \overset{\bullet}{=} \inf\left\{ \mathfrak{t} \in [0, 1] : \|\bar{\mathfrak{P}}_{T,x}^N(\bar{\Phi}^{N,2})\|_{\mathfrak{t};\mathbb{T}_N} \gtrsim N^{-\beta_{\mathrm{univ}}} + N^{-\beta_{\mathrm{univ}}}\|\bar{\mathfrak{Z}}^N\|_{\mathfrak{t};\mathbb{T}_N}^2 \right\} \wedge 1. \tag{8.6b}$$

Define $\mathfrak{D}^N \overset{\bullet}{=} \mathfrak{Y}^N - \mathfrak{Q}^N$ where $\mathfrak{Q}^N$ solves the stochastic equation on $\mathbb{R}_{\geqslant 0} \times \mathbb{T}_N$ below with $\bar{\mathfrak{Z}}_{0,\bullet}^N = \chi_\bullet \mathfrak{Z}_{0,\bullet}^N$ from Definition 4.3:

$$\mathfrak{Q}_{T,x}^N = \bar{\mathfrak{P}}_{T,x}^{N,\mathbf{X}}(\bar{\mathfrak{Z}}_{0,\bullet}^N) + \bar{\mathfrak{P}}_{T,x}^N(\mathfrak{Q}^N \mathrm{d}\xi^N) + \bar{\mathfrak{P}}_{T,x}^N(\bar{\Phi}_{S,y}^{N,2}\mathbf{1}_{S\leqslant\mathfrak{t}_{1,\mathrm{st}}}) + \bar{\mathfrak{P}}_{T,x}^N(\bar{\Phi}^{N,5}). \tag{8.7}$$

See Remark 4.4 for definition of $\mathfrak{Q}^N \mathrm{d}\xi^N$. Recall $\bar{\Phi}^{N,2}$ in Definition 4.3. We have also introduced the following adaptation $\bar{\Phi}^{N,5}$ of $\bar{\Phi}^{N,3}$ in the defining equation for $\bar{\mathfrak{Z}}^N$ but catered to $\mathfrak{Q}^N$. Recall $\bar{\nabla}^! = N\bar{\nabla}$ where $\bar{\nabla}$ is gradient on the torus $\mathbb{T}_N$:

$$\bar{\Phi}_{T,x}^{N,5} \overset{\bullet}{=} \mathfrak{w}_{T,x}\mathfrak{Q}_{T,x}^N + \sum_{k=-2\mathfrak{m}}^{2\mathfrak{m}} c_k \bar{\nabla}_k^!(\mathfrak{w}_{T,x}^k \mathfrak{Q}_{T,x}^N). \tag{8.8}$$

We will eventually pick $\beta_{\mathrm{univ}} \in \mathbb{R}_{>0}$ small depending only on universal constants including $\beta_{\mathrm{univ},2} \in \mathbb{R}_{>0}$ from Proposition 7.1. We will then choose $\varepsilon_{\mathrm{st}} \in \mathbb{R}_{>0}$ sufficiently small depending only on $\beta_{\mathrm{univ}} \in \mathbb{R}_{>0}$. To be concrete, the reader is welcome to take $\varepsilon_{\mathrm{st}} = 99^{-99} \wedge 99^{-99}\beta_{\mathrm{univ}}$ and $\beta_{\mathrm{univ}} = 99^{-99}\beta_{\mathrm{univ},2}$ with $\beta_{\mathrm{univ},2} \in \mathbb{R}_{>0}$ from the statement of Proposition 7.1.

We first provide the following pathwise result which relates $\mathfrak{Q}^N$ to $\bar{\mathfrak{Z}}^N$ relative to the random time $\mathfrak{t}_{1,\mathrm{st}}$. Roughly speaking, it follows from the observation that $\bar{\mathfrak{Z}}^N$ and $\mathfrak{Q}^N$ solve the same defining stochastic linear equation before time $\mathfrak{t}_{1,\mathrm{st}}$.

**Lemma 8.8.** *We have* $\bar{\mathfrak{Z}}_{\mathfrak{s},x}^N = \mathfrak{Q}_{\mathfrak{s},x}^N$ *for all* $0 \leqslant \mathfrak{s} \leqslant \mathfrak{t}_{1,\mathrm{st}}$ *and* $x \in \mathbb{T}_N$ *with probability 1.*

Roughly speaking, the result Lemma 8.9 below provides $\|\|_{1;\mathbb{T}_N}$-estimates for the difference process $\mathfrak{D}^N = \mathfrak{Y}^N - \mathfrak{Q}^N$. Note $\mathfrak{D}^N$ solves a linear equation like that defining $\mathfrak{Y}^N$ but with the extra $\bar{\Phi}^{N,2}$-contribution in the $\mathfrak{Q}^N$ equation and vanishing initial condition since the initial data of $\mathfrak{Y}^N$ and $\mathfrak{Q}^N$ coincide. The proposed bound in the first result below is will be shown by standard moment bounds to get pointwise control on $\mathfrak{D}^N$ and a stochastic continuity argument like with the proof of Proposition 6.2. In the proof for Lemma 8.9 we rely heavily on the fact that we have cutoff the $\bar{\Phi}^{N,2}$-contribution in $\mathfrak{Q}^N$ by the upper bound on the RHS of the estimate in Proposition 7.1 so we have *deterministic* control on the $\bar{\Phi}^{N,2}$-contribution in $\mathfrak{Q}^N$. Afterwards, we record a second result Lemma 8.10 that shows the $\mathfrak{t}_{1,\mathrm{st}}$-cutoff in the $\mathfrak{Q}^N$-equation can be ignored with high-probability.

**Lemma 8.9.** *For some $\beta_{\mathrm{univ},2} \in \mathbb{R}_{>0}$ universal and any $C \in \mathbb{R}_{>0}$, with probability at least $1 - \kappa_C N^{-C}$ we have $\|\mathfrak{D}^N\|_{1;\mathbb{T}_N} \lesssim N^{-\beta_{\mathrm{univ},2}}$.*

**Lemma 8.10.** *There exists some $\beta_{\mathrm{univ},1} \in \mathbb{R}_{>0}$ universal so that with probability at least $1 - \kappa N^{-\beta_{\mathrm{univ},1}}$, we have $\mathfrak{t}_{1,\mathrm{st}} = 1$ provided $\beta_{\mathrm{univ}} \in \mathbb{R}_{>0}$ in the definition of $\mathfrak{t}_{1,\mathrm{st}}$ is sufficiently small and $\varepsilon_{\mathrm{st}} \in \mathbb{R}_{>0}$ is sufficiently small depending only on our choice of $\beta_{\mathrm{univ}}$.*

*Proof of Lemma 8.5.* Lemma 8.9 gives Lemma 8.5 if we replace $\bar{\mathfrak{Z}}^N$ by $\mathfrak{Q}^N$ in the statement of Lemma 8.5. But Lemma 8.10 tells us that with the required high-probability, such replacement can be removed as $\bar{\mathfrak{Z}}^N = \mathfrak{Q}^N$ with high-probability on $[0,1] \times \mathbb{T}_N$ by pathwise identification in Lemma 8.8. This completes the proof of Lemma 8.5 modulo proofs of Lemmas 8.8, 8.9, and 8.10. We defer these to the final subsection as they are on the technical side. $\qquad\square$

*8.2. Proof of Proposition 8.3.* Let us first observe the defining stochastic linear equation for $\mathfrak{Y}^N$ contains two terms which resemble a microscopic version of the continuum SHE and another that contains only weakly vanishing terms. In particular, the stochastic linear equation for $\mathfrak{Y}^N$ resembles that of the microscopic Cole-Hopf transform within [10]. We are now in a position to follow the proof of Proposition 2.2 in [10] and thus the proof of Theorem 1.1 in [10]. We explain this in more detail below.

- Tightness of the microscopic Cole-Hopf transform under the appropriate space-time scaling in [10] is stated in Proposition 1.4 in [10]. It uses two ingredients. The first is Corollary 3.3 in [10] and the second is boundedness of a Poisson clock speed. For the former, the list of moment estimates in Corollary 3.3 in [10] also holds for the $\mathfrak{Y}^N$-process here because it only needs the defining linear equation (3.2) in [10] for the microscopic Cole-Hopf transform. The defining linear equation for $\mathfrak{Y}^N$ has the same structure as that in (3.2) in [10]. We also use Lemma B.1 instead of Lemma 3.1 in [10] here. The second ingredient concerning Poisson clocks speeds also holds in our case; in [10] the maximal jump-length did not play an important role.
- To get convergence of the microscopic Cole-Hopf transform to the continuum SHE in [10], the approach taken therein is via the martingale problem. To use ideas in [10], we require two ingredients. The first is spatial regularity of $\mathfrak{Y}^N$ on microscopic length-scales to treat the correlations in $\mathrm{d}\xi^N$ at different spatial points that come from the non-simple nature of the particle system. For the microscopic Cole-Hopf transform this was done via explicit formulas in terms of the particle system in [10], but the regularity from Corollary 3.3 in [10], which holds for $\mathfrak{Y}^N$ as noted in the last bullet point, suffices for this as well. The second ingredient we require is a hydrodynamic limit argument for analysis of weakly vanishing functionals at macroscopic space-time scales; see Lemma 2.5 from [10]. Our version is Lemma 8.11.
- The analysis in [10] is done on $\mathbb{R}_{\geq 0} \times \mathbb{Z}$, while $\mathfrak{Y}^N$ is defined on $\mathbb{R}_{\geq 0} \times \mathbb{T}_N$. This does not matter as our heat kernel estimates in Lemma A.1 hold for $\bar{\mathfrak{P}}^N$. We may also follow the proof of Proposition 4.6 but for $\mathfrak{Y}^N$ to replace $\mathfrak{Y}^N$ by its $\mathbb{R}_{\geq 0} \times \mathbb{Z}$-version.

**Lemma 8.11.** *Consider a weakly vanishing functional $\mathfrak{w} : \Omega_{\mathbb{Z}} \to \mathbb{R}$. Provided any smooth test function $\varphi : \mathbb{R} \to \mathbb{R}$ with compact support, for $\mathfrak{u} \in \{1,2\}$ and for any fixed $T \in \mathbb{R}_{\geq 0}$ independent of $N \in \mathbb{Z}_{>0}$, we have the following convergence in probability:*

$$\lim_{N\to\infty} \frac{1}{N} \int_0^T \sum_{y\in\mathbb{Z}} \varphi_{N^{-1}y} \cdot \mathfrak{w}_{S,y} \cdot |\mathfrak{Y}_{S,y}^N|^{\mathfrak{u}} \, \mathrm{d}S \;=\; 0. \tag{8.9}$$

We defer the proof of Lemma 8.11 to the last subsection along with other technical results. We remark that the proof of Lemma 8.11 follows that of Lemma 2.5 in [10], as we have spatial regularity of $\mathfrak{Y}^N$ needed to replace $\mathfrak{w}$ by mesoscopic spatial average and we have the entropy production estimate in Proposition 3.6 needed to run the one-block/two-blocks steps in the proof for Lemma 2.5 in [10]. There is a caveat. In the proof of Lemma 2.5 in [10], density fluctuations are controlled by regularity of the microscopic Cole-Hopf transform. Here it is also given by the *original* microscopic Cole-Hopf transform $\mathfrak{Z}^N$, not $\mathfrak{Y}^N$, as only $\mathfrak{Z}^N$ is defined via the particle system. On the other hand, what we will need from $\mathfrak{Z}^N$ can be estimated via $\mathfrak{Y}^N$ by Proposition 8.2.

*8.3. Proofs of Technical Estimates.* It will be helpful to have the following a priori moment estimate for $\mathfrak{Y}^N$. Its proof was described in the subsection concerning the proof of Proposition 8.3 in the bullet point about tightness of $\mathfrak{Y}^N$ so we omit it.

**Lemma 8.12.** *Consider any* $p \in \mathbb{R}_{\geqslant 1}$. *We have the moment estimate* $\sup_{T \in [0,1]} \sup_{x \in \mathbb{T}_N}$ $\|\mathfrak{Y}^N_{T,x}\|_{\omega;2p} \lesssim_p 1$.

**Corollary 8.13.** *Consider any* $\varepsilon, C \in \mathbb{R}_{>0}$. *Outside an event with probability at most* $\kappa_{\varepsilon,C} N^{-C}$, *we have* $\|\mathfrak{Y}^N\|_{1;\mathbb{T}_N} \lesssim N^\varepsilon$.

*Proof.* Note in SDE-type equation formulation, for times in $[0, 1]$ we get $d\mathfrak{Y}^N = \mathcal{Q}\mathfrak{Y}^N dT$ with jumps of order $N^{-1/2}\|\mathfrak{Y}^N\|_{1;\mathbb{T}_N}$ where $\mathcal{Q}$ satisfies the same operator bounds as the operator $\mathcal{Q}$ in the proof of Lemma 6.6. Indeed, the only difference between these two operators is a term that is multiplicative in the solution $\mathfrak{Y}^N$ or $\bar{\mathfrak{Z}}^N$ with coefficient $\bar{\Phi}^{N,2}$ of order $N^{1/2}$. Thus, the proof of Lemma 6.6 applies to $\mathfrak{Y}^N$ in place of $\bar{\mathfrak{Z}}^N$. This controls $\|\cdot\|_{1;\mathbb{T}_N}$ by $[\cdot]_{1;\mathbb{T}_N}$ with the required high-probability as with Lemma 6.4:

$$\|\mathfrak{Y}^N\|_{1;\mathbb{T}_N} \lesssim [\mathfrak{Y}^N]_{1;\mathbb{T}_N}. \tag{8.10}$$

Indeed, the only way for (8.10) to fail is $\mathfrak{Y}^N$ to have large change between times in the discretization $\{jN^{-100}\}_{j=0}^{N^{100}}$. Lemma 6.6 though for $\mathfrak{Y}^N$ prevents that with high-probability. Thus it is left to show Corollary 8.13 upon replacing $\|\mathfrak{Y}^N\|_{1;\mathbb{T}_N}$ by $[\mathfrak{Y}^N]_{1;\mathbb{T}_N}$. To this end, we observe that a union bound allows us to control the probability for which $[\mathfrak{Y}^N]_{1;\mathbb{T}_N} \gtrsim N^\varepsilon$ by the probability for which $\mathfrak{Y}^N_{t,x} \gtrsim N^\varepsilon$, then multiplied by $N^{105}$. The one-point probabilities are at most $\kappa_{\varepsilon,C'} N^{-C'}$ for any $C' \in \mathbb{R}_{>0}$ by the Chebyshev inequality for high $p$-moments and the moment estimate in Lemma 8.12 as in the proofs of Proposition 4.6/Lemma 6.5.                                                            $\square$

*Proof of Lemma 8.6.* If $N \in \mathbb{Z}_{\geqslant 0}$ is sufficiently large depending only on $\mathbb{K} \subseteq \mathbb{R}$, note $\mathbb{K}_N = N\mathbb{K} \cap \mathbb{Z} \subseteq \widetilde{\mathbb{T}}_N$, where $\widetilde{\mathbb{T}}_N$ is defined in Proposition 4.6. Thus, it suffices to get the following comparison with high-probability in which $\beta_{\mathrm{univ}} \in \mathbb{R}_{>0}$ is universal:

$$\|\mathfrak{Z}^N - \bar{\mathfrak{Z}}^N\|_{1;\mathbb{K}_N} \lesssim [\mathfrak{Z}^N - \bar{\mathfrak{Z}}^N]_{1;\mathbb{K}_N} + N^{-\beta_{\mathrm{univ}}}. \tag{8.11}$$

Indeed, Proposition 4.6 tells us the first term on the RHS of (8.11) is at most $\kappa_C N^{-C}$ with high-probability as $\mathbb{K}_N \subseteq \widetilde{\mathbb{T}}_N$. Observe the LHS of (8.11) is controlled by the first term on the RHS of (8.11) and behavior of $\mathfrak{Z}^N$ and $\bar{\mathfrak{Z}}^N$ between times in $\{jN^{-100}\}_{j=0}^{N^{100}}$.

Precisely, we get a deterministic bound by the triangle inequality; recall time-gradients defined prior to Proposition 6.2:

$$\|\mathfrak{Z}^N - \bar{\mathfrak{Z}}^N\|_{1;\mathbb{K}_N} \lesssim [\mathfrak{Z}^N - \bar{\mathfrak{Z}}^N]_{1;\mathbb{K}_N} + \sup_{0 \leqslant t \leqslant 1} \sup_{0 \leqslant \mathfrak{s} \leqslant N^{-100}} \sup_{x \in \mathbb{K}_N} |\nabla_{\mathfrak{s}}^{\mathbf{T}} \mathfrak{Z}_{t,x}^N|$$
$$+ \sup_{0 \leqslant t \leqslant 1} \sup_{0 \leqslant \mathfrak{s} \leqslant N^{-100}} \sup_{x \in \mathbb{K}_N} |\nabla_{\mathfrak{s}}^{\mathbf{T}} \bar{\mathfrak{Z}}_{t,x}^N|. \tag{8.12}$$

To get (8.11) with the required high-probability, we estimate the second and third terms on the RHS of (8.12).

- We cite (2.3) in [10] for the following stochastic equation for the microscopic Cole-Hopf transform $\mathfrak{Z}^N$ that is a starting point for the proof of the stochastic equation within Proposition 2.7 we have been employing until this point. We emphasize that this holds for $\mathfrak{Z}^N$ in this paper because the proof of Eq. (2.3) in [10] does not depend on the maximal jump-length:

$$(\mathfrak{Z}_{T,x}^N)^{-1} \mathrm{d}\mathfrak{Z}_{T,x}^N = \mathfrak{f}_{T,x}^N \mathrm{d}T + \mathrm{d}\xi_{T,x}^N. \tag{8.13}$$

Above the functional $\mathfrak{f}^N$ satisfies $|\mathfrak{f}^N| \lesssim N^2$ and $\mathrm{d}\xi^N$ is the same martingale integrator that is in Proposition 2.7. In words, the dynamic of the original microscopic Cole-Hopf transform $\mathfrak{Z}^N$ is multiplicative in $\mathfrak{Z}^N$ *at the same point* with a continuous part that has speed of order $N^2$ and jumps of speed $N^2$ that are of order $N^{-1/2}$. This is in contrast to $\bar{\mathfrak{Z}}^N$, because the only dynamic equation we have for $\bar{\mathfrak{Z}}^N$ has the discrete-type Laplacian $\bar{\mathscr{L}}^{!!}$ that relates the growth of $\bar{\mathfrak{Z}}^N$ to its values at neighboring points. We use reasoning in the proof of Lemma 6.6 to get the following high-probability short-time $\|\,\|_{1;\mathbb{K}_N}$-bound for any $\varepsilon \in \mathbb{R}_{>0}$:

$$\sup_{0 \leqslant t \leqslant 1} \sup_{0 \leqslant \mathfrak{s} \leqslant N^{-100}} \sup_{x \in \mathbb{K}_N} |\nabla_{\mathfrak{s}}^{\mathbf{T}} \mathfrak{Z}_{t,x}^N| \lesssim N^{-\frac{1}{2}+\varepsilon} \|\mathfrak{Z}^N\|_{1;\mathbb{K}_N}. \tag{8.14}$$

- We now use Lemma 6.6 as written to get, also with the required high-probability, a short-time bound now in terms of $\|\,\|_{1;\mathbb{T}_N}$:

$$\sup_{0 \leqslant t \leqslant 1} \sup_{0 \leqslant \mathfrak{s} \leqslant N^{-100}} \sup_{x \in \mathbb{K}_N} |\nabla_{\mathfrak{s}}^{\mathbf{T}} \bar{\mathfrak{Z}}_{t,x}^N| \lesssim N^{-\frac{1}{2}+\varepsilon} \|\bar{\mathfrak{Z}}^N\|_{1;\mathbb{T}_N}. \tag{8.15}$$

Via Lemma 8.5 and Corollary 8.13 and the triangle inequality, we estimate $\bar{\mathfrak{Z}}^N$ by comparing it to $\mathfrak{Y}^N$ and then by estimating $\mathfrak{Y}^N$. Ultimately, with the required high-probability we get the following for some universal constant $\beta_{\mathrm{univ},2}$:

$$\|\bar{\mathfrak{Z}}^N\|_{1;\mathbb{T}_N} \leqslant \|\bar{\mathfrak{Z}}^N - \mathfrak{Y}^N\|_{1;\mathbb{T}_N} + \|\mathfrak{Y}^N\|_{1;\mathbb{T}_N} \lesssim N^{-\beta_{\mathrm{univ},2}} + N^{\varepsilon}. \tag{8.16}$$

We now combine (8.12), (8.14), (8.15), and (8.16) to deduce the following for which we give more explanation after:

$$\|\mathfrak{Z}^N - \bar{\mathfrak{Z}}^N\|_{1;\mathbb{K}_N} \lesssim [\mathfrak{Z}^N - \bar{\mathfrak{Z}}^N]_{1;\mathbb{K}_N} + N^{-\frac{1}{2}+\varepsilon} \|\mathfrak{Z}^N\|_{1;\mathbb{K}_N} + N^{-\frac{1}{2}+\varepsilon} \|\bar{\mathfrak{Z}}^N\|_{1;\mathbb{T}_N} \tag{8.17}$$

$$\lesssim [\mathfrak{Z}^N - \bar{\mathfrak{Z}}^N]_{1;\mathbb{K}_N} + N^{-\frac{1}{2}+\varepsilon} \|\mathfrak{Z}^N - \bar{\mathfrak{Z}}^N\|_{1;\mathbb{K}_N} + N^{-\frac{1}{2}+\varepsilon} \|\bar{\mathfrak{Z}}^N\|_{1;\mathbb{T}_N} \tag{8.18}$$

$$\lesssim [\mathfrak{Z}^N - \bar{\mathfrak{Z}}^N]_{1;\mathbb{K}_N} + N^{-\frac{1}{2}+\varepsilon} \|\mathfrak{Z}^N - \bar{\mathfrak{Z}}^N\|_{1;\mathbb{K}_N}$$
$$+ N^{-\frac{1}{2}+\varepsilon} N^{-\beta_{\mathrm{univ},2}} + N^{-\frac{1}{2}+2\varepsilon}. \tag{8.19}$$

The first line in the above calculation follows from directly applying (8.14) and (8.15) to (8.12). The second line follows from the triangle inequality $\|\mathfrak{Z}^N\|_{1;\mathbb{K}_N} \leqslant \|\mathfrak{Z}^N - \bar{\mathfrak{Z}}^N\|_{1;\mathbb{K}_N} + \|\bar{\mathfrak{Z}}^N\|_{1;\mathbb{K}_N}$ combined with the bound $\|\bar{\mathfrak{Z}}^N\|_{1;\mathbb{K}_N} \leqslant \|\bar{\mathfrak{Z}}^N\|_{1;\mathbb{T}_N}$. The last bound (8.19) follows from employing (8.16). We now move the second term within (8.19) to the LHS of the first line in this previous set of estimates and choose $\varepsilon \in \mathbb{R}_{>0}$ small to deduce (8.11) and complete the proof. $\qquad\square$

*Proof of Lemma 8.8.* Observe until time $\mathfrak{t}_{1,\mathrm{st}} \in \mathbb{R}_{\geqslant 0}$, the difference $\Phi = \bar{\mathfrak{Z}}^N - \mathfrak{Q}^N$ solves the same equation as $\bar{\mathfrak{Z}}^N$ upon replacing every factor of $\bar{\mathfrak{Z}}^N$ with $\Phi$. We emphasize $\Phi$ has vanishing initial condition. Thus, because $\Phi$ solves a linear evolution equation, whose solutions are unique with probability 1, and a vanishing initial condition until $\mathfrak{t}_{1,\mathrm{st}}$, we have $\Phi = 0$ until $\mathfrak{t}_{1,\mathrm{st}}$. $\qquad\square$

*Proof of Lemma 8.9.* The first step we take to control $\mathfrak{D}^N$ is the following stochastic equation for $\mathfrak{D}^N$:

$$\mathfrak{D}_{T,x}^N = \bar{\mathfrak{P}}_{T,x}^N(\mathfrak{D}^N \mathrm{d}\xi^N) - \bar{\mathfrak{P}}_{T,x}^N(\bar{\Phi}_{S,y}^{N,2}\mathbf{1}_{S \leqslant \mathfrak{t}_{1,\mathrm{st}}}) + \bar{\mathfrak{P}}_{T,x}^N(\bar{\Phi}^{N,6}). \tag{8.20}$$

The last term in the $\mathfrak{D}^N$-equation is the following relevant weakly-vanishing content from Proposition 2.7 but catered to $\mathfrak{D}^N$:

$$\bar{\Phi}_{T,x}^{N,6} \overset{\bullet}{=} \mathfrak{w}_{T,x}\mathfrak{D}_{T,x}^N + \sum_{k=-2\mathfrak{m}}^{2\mathfrak{m}} c_k \bar{\nabla}_k^!(\mathfrak{w}_{T,x}^k \mathfrak{D}_{T,x}^N). \tag{8.21}$$

This equation for $\mathfrak{D}^N$ follows by subtracting stochastic equations for $\mathfrak{Q}^N$ and $\mathfrak{Y}^N$. We now proceed with the following steps.

- Observe $\mathrm{d}\mathfrak{D}^N = \mathcal{Q}\mathfrak{D}^N \mathrm{d}T + \mathbf{1}_{T \leqslant \mathfrak{t}_{1,\mathrm{st}}}\bar{\Phi}^{N,2}\mathrm{d}T$ with jumps of order $N^{-1/2}\mathfrak{D}^N$. Here the operator $\mathcal{Q}$ satisfies the same operator-estimates as that within the proof for Lemma 6.6. Indeed, the operator $\mathcal{Q}$ here is that from the proof of Corollary 8.13, so the only difference between $\mathcal{Q}$ in here and that in the proof for Lemma 6.6 is the lack of a multiplicative term in the solution with pseudo-gradient coefficients of order $N^{1/2}$. In particular, Lemma 6.6 holds for $\mathfrak{D}^N$ as well except now in short time-intervals we also get short-time contribution of $\bar{\Phi}^{N,2}$ for times before $\mathfrak{t}_{1,\mathrm{st}}$. However, before time $\mathfrak{t}_{1,\mathrm{st}}$, we control $\bar{\Phi}^{N,2}$ by $\bar{\mathfrak{Z}}^N$ times small powers of $N$, which is then controlled via small powers of $N$. Thus, the contribution of $\bar{\Phi}^{N,2}$ before time $\mathfrak{t}_{1,\mathrm{st}}$ for short times is at most $N^{10}$ times the length $N^{-100}$ of these short time-windows. With the required high-probability, as in the proof of (8.10),

$$\|\mathfrak{D}^N\|_{1;\mathbb{T}_N} \lesssim [\mathfrak{D}^N]_{1;\mathbb{T}_N} + N^{-90}. \tag{8.22}$$

By (8.22) it is left to get Lemma 8.9 upon replacing $\|\mathfrak{D}^N\|_{1;\mathbb{T}_N}$ by $[\mathfrak{D}^N]_{1;\mathbb{T}_N}$. As in proofs for Proposition 4.6/Corollary 8.13, it suffices to bound pointwise moments of $\mathfrak{D}^N$. Precisely, it suffices to show, for any $p \in \mathbb{R}_{\geqslant 1}$ and some $\beta \in \mathbb{R}_{>0}$ universal, that

$$\sup_{\mathfrak{t} \in [0,1]} \sup_{x \in \mathbb{T}_N} \|\mathfrak{D}_{\mathfrak{t},x}^N\|_{\omega;2p} \lesssim_p N^{-\beta}. \tag{8.23}$$

We turn to the stochastic heat operator equation for $\mathfrak{D}^N$ at the beginning of this proof. First, this stochastic equation gives

$$\|\mathfrak{D}_{T,x}^N\|_{\omega;2p}^2 \lesssim \|\bar{\mathfrak{P}}_{T,x}^N(\bar{\Phi}_{S,y}^{N,2}\mathbf{1}_{S \leqslant \mathfrak{t}_{1,\mathrm{st}}})\|_{\omega;2p}^2$$
$$+ \|\bar{\mathfrak{P}}_{T,x}^N(\mathfrak{D}^N \mathrm{d}\xi^N)\|_{\omega;2p}^2 + \|\bar{\mathfrak{P}}_{T,x}^N(\bar{\Phi}^{N,6})\|_{\omega;2p}^2. \tag{8.24}$$

We now estimate the first term on the RHS of (8.24). To this end, we apply the deterministic bound in Lemma A.3 to get

$$\|\bar{\mathfrak{P}}^N_{T,x}(\bar{\Phi}^{N,2}_{S,y}\mathbf{1}_{S\leqslant t_{1,\mathrm{st}}})\|_{1;\mathbb{T}_N} \leqslant \|\bar{\mathfrak{P}}^N_{T,x}(\bar{\Phi}^{N,2}_{S,y})\|_{t_{1,\mathrm{st}};\mathbb{T}_N} \lesssim N^{-\beta_{\mathrm{univ}}+2\varepsilon_{\mathrm{st}}} \lesssim N^{-\frac{1}{2}\beta_{\mathrm{univ}}}. \tag{8.25}$$

The second bound in (8.25) follows by definition of $t_{1,\mathrm{st}}$ in Definition 8.7. The third bound in (8.25) follows by recalling $\varepsilon_{\mathrm{st}}$ can be chosen small depending only on $\beta_{\mathrm{univ}}$, for example $\varepsilon_{\mathrm{st}} \leqslant 99^{-99}\beta_{\mathrm{univ}}$. Using the deterministic bound (8.25) with (8.24),

$$\|\mathfrak{D}^N_{T,x}\|^2_{\omega;2p} \lesssim N^{-\beta_{\mathrm{univ}}} + \|\bar{\mathfrak{P}}^N_{T,x}(\mathfrak{D}^N\mathrm{d}\xi^N)\|^2_{\omega;2p} + \|\bar{\mathfrak{P}}^N_{T,x}(\bar{\Phi}^{N,6})\|^2_{\omega;2p}. \tag{8.26}$$

At this point, we follow the proof for (3.12) in Proposition 3.2 in [10], as we did in the proof of Lemma 4.13, to get (8.23) from (8.26); we again use the martingale inequality in Lemma B.1 instead of Lemma 3.1 in [10]. This completes the proof. $\qquad\square$

*Proof of Lemma 8.10.* We first start with constructing some random times in addition to the previous $t_{1,\mathrm{st}}$-random time.

**Definition 8.14.** Recall $\varepsilon_{\mathrm{st}} \in \mathbb{R}_{\geqslant 0}$ from the construction of $t_{1,\mathrm{st}}$. Define the random time $t_{2,\mathrm{st}} = t_{2,\mathrm{st},1} \wedge t_{2,\mathrm{st},2}$, where

$$t_{2,\mathrm{st},1} \overset{\bullet}{=} \inf\left\{t \in [0,1]: \|\mathfrak{Y}^N\|_{t;\mathbb{T}_N} \geqslant N^{\varepsilon_{\mathrm{st}}}\right\} \wedge 1 \tag{8.27a}$$

$$t_{2,\mathrm{st},2} \overset{\bullet}{=} \inf\left\{t \in [0,1]: \sup_{0\leqslant\mathfrak{s}\leqslant N^{-100}}\sup_{0\leqslant t_0\leqslant t}\sup_{x\in\mathbb{T}_N}\left(1+\|\bar{\mathfrak{z}}^N\|^2_{t_0;\mathbb{T}_N}\right)^{-1}|\nabla^{\mathbf{T}}_{\pm\mathfrak{s}}\bar{\mathfrak{z}}^N_{t_0,x}| \gtrsim N^{-\frac{1}{2}+\varepsilon_{\mathrm{st}}}\right\} \wedge 1. \tag{8.27b}$$

We additionally define the random time $t_{3,\mathrm{st}} \overset{\bullet}{=} \inf\{t \in [0,1]: \|\mathfrak{D}^N\|_{t;\mathbb{T}_N} \geqslant N^{\varepsilon_{\mathrm{st}}}\} \wedge 1$.

The time $t_{2,\mathrm{st}}$ will both control $\mathfrak{Y}^N$ and the maximal change in $\bar{\mathfrak{z}}^N$ over any short time-interval. The time $t_{3,\mathrm{st}}$ will control the difference between $\mathfrak{Y}^N$ and $\mathfrak{Q}^N$. Thus the minimum of $t_{2,\mathrm{st}}$ and $t_{3,\mathrm{st}}$ will control $\mathfrak{Q}^N$. Moreover, with the pathwise identification in Lemma 8.8, this will control $\bar{\mathfrak{z}}^N$ until time $t_{1,\mathrm{st}}$. The short-time continuity estimate within $t_{2,\mathrm{st},2}$ will then help us to propagate this control on $\bar{\mathfrak{z}}^N$. We first show $t_{2,\mathrm{st}}$ and $t_{3,\mathrm{st}}$ are equal to 1 with high probability. We also do this for $t_{1,\mathrm{st},2}$ from Definition 8.7.

- For $t_{2,\mathrm{st}}$, we first note that Lemma 6.6 guarantees $\mathbf{P}[t_{2,\mathrm{st},2} \neq 1] \lesssim_{\varepsilon_{\mathrm{st}},C} N^{-C}$. Moreover, Corollary 8.13 guarantees the same is true for $t_{2,\mathrm{st},1}$. Thus, because the event $t_{2,\mathrm{st}} \neq 1$ implies either $t_{2,\mathrm{st},1} \neq 1$ or $t_{2,\mathrm{st},2} \neq 1$ by definition, we get

$$\mathbf{P}[t_{2,\mathrm{st}} \neq 1] \leqslant \mathbf{P}[t_{2,\mathrm{st},1} \neq 1] + \mathbf{P}[t_{2,\mathrm{st},2} \neq 1] \lesssim_{\varepsilon_{\mathrm{st}},C} N^{-C}. \tag{8.28}$$

- Consequence of Lemma 8.9, we gain very high-probability control of $\|\mathfrak{D}^N\|_{1;\mathbb{T}_N}$, and thus

$$\mathbf{P}[t_{3,\mathrm{st}} \neq 1] \lesssim_{\varepsilon_{\mathrm{st}},C} N^{-C}. \tag{8.29}$$

- We use Proposition 7.1 with the random time $t_{\mathrm{st}} = t_{1,\mathrm{st},2}$. As $t_{1,\mathrm{st},2} \leqslant 1$, if $t_{1,\mathrm{st},2} \neq 1$ then $t_{1,\mathrm{st},2} < 1$. If $\beta_{\mathrm{univ}} \in \mathbb{R}_{>0}$ in definition of $t_{1,\mathrm{st},2}$ is sufficiently small but still universal depending only on $\beta_{\mathrm{univ},2} \in \mathbb{R}_{>0}$ in Proposition 7.1 and $t_{1,\mathrm{st},2} < 1$, then (7.1) in Proposition 7.1 fails for $t_{\mathrm{st}} = t_{1,\mathrm{st},2}$. By Proposition 7.1, this happens with probability at most $N^{-\beta_{\mathrm{univ},1}}$ times universal factors:

$$\mathbf{P}[t_{1,\mathrm{st},2} \neq 1] \lesssim N^{-\beta_{\mathrm{univ},1}}. \tag{8.30}$$

Thus, for us to prove Lemma 8.10, it suffices to show $t_{1,st,1} = 1$ conditioning on the event on which $t_{1,st,2}, t_{2,st}, t_{3,st} = 1$. Precisely, because $t_{1,st} = t_{1,st,1} \wedge t_{1,st,2}$, we get a union bound inequality that says if $t_{1,st} \neq 1$ then $t_{1,st,2} \neq 1$ or if $t_{1,st,2} = 1$ then $t_{1,st,1} \neq 1$:

$$\mathbf{P}[t_{1,st} \neq 1] \leqslant \mathbf{P}[t_{1,st,2} \neq 1] + \mathbf{P}[t_{1,st,2} = 1, t_{1,st,1} \neq 1]. \qquad (8.31)$$

We decompose the second probability on the RHS of (8.31) according to a similar observation for $t_{2,st}, t_{3,st}$:

$$\mathbf{P}[t_{1,st,2} = 1, t_{1,st,1} \neq 1] \leqslant \mathbf{P}[t_{2,st} \neq 1] + \mathbf{P}[t_{3,st} \neq 1] + \mathbf{P}[t_{1,st,2}, t_{2,st}, t_{3,st} \\ = 1, t_{1,st,1} \neq 1]. \qquad (8.32)$$

We combine the union bound inequalities (8.31) and (8.32) with (8.28), (8.29), and (8.30) to get

$$\mathbf{P}[t_{1,st} \neq 1] \lesssim N^{-\beta_{univ,1}} + \mathbf{P}[t_{1,st,2}, t_{2,st}, t_{3,st} = 1, t_{1,st,1} \neq 1]. \qquad (8.33)$$

Thus, to prove Lemma 8.10/estimate the LHS of (8.33) above, it suffices to show the last probability is equal to 0 at least when $N$ is sufficiently large depending only on $\varepsilon_{st}, \beta_{univ,1} \in \mathbb{R}_{>0}$. We do this with a continuity argument that is classical in PDEs.

Suppose $t_{1,st,1} \neq 1$ while $t_{1,st,2}, t_{2,st}, t_{3,st} = 1$. In this case we observe $t_{1,st,1} < 1$ since $t_{1,st,1} \leqslant 1$ with probability 1. In this case, we can pick $0 < t_N \leqslant N^{-100}$ such that $t_{1,st,1} + t_N \leqslant 1$. We will show that $\|\bar{\mathfrak{Z}}^N\|_{t_{1,st,1}+t_N; \mathbb{T}_N} \lesssim N^{\varepsilon_{st}}$ with universal implied constant with probability 1. If $N \in \mathbb{Z}_{>0}$ is sufficiently large depending only on $\varepsilon_{st} \in \mathbb{R}_{>0}$, then we get from this $\|\bar{\mathfrak{Z}}^N\|_{t_{1,st,1}+t_N; \mathbb{T}_N} \leqslant N^{3\varepsilon_{st}/2}$. This would contradict the definition of $t_{1,st,1}$ conditioning on $t_{1,st,1} \neq 1$; we could push $t_{1,st,1}$ past whatever we condition it to be if $t_{1,st,1} \neq 1$ and $t_{1,st,2}, t_{2,st}, t_{3,st} = 1$. This shows $t_{1,st,1} \neq 1$ and $t_{1,st,2}, t_{2,st}, t_{3,st} = 1$ is impossible as we would have $t_{1,st,1} > t_{1,st,1}$.

Because $t_{2,st} = 1$, from the definitions of $t_{2,st} = 1$ and $t_{2,st,2}$ we get $t_{2,st,2} = 1$ so $\bar{\mathfrak{Z}}^N$ cannot change very much in short time, so given any $\mathfrak{s} \in [0, t_N]$ we establish the following continuity estimate. We recall that $t_N \leqslant N^{-100}$ and $\varepsilon_{st} \in \mathbb{R}_{>0}$ is arbitrarily small but universal. For example, we may pick $\varepsilon_{st} = 99^{-99} \wedge 99^{-99}\beta_{univ}$ where $\beta_{univ} \in \mathbb{R}_{>0}$ is the universal exponent defining $t_{1,st,2}$:

$$|\bar{\mathfrak{Z}}^N_{t_{1,st,1}+\mathfrak{s},x}| \lesssim |\bar{\mathfrak{Z}}^N_{t_{1,st,1},x}| + N^{-\frac{1}{2}+\varepsilon_{st}} + N^{-\frac{1}{2}+\varepsilon_{st}}\|\bar{\mathfrak{Z}}^N\|^2_{t_{1,st,1}; \mathbb{T}_N} \\ \lesssim \|\bar{\mathfrak{Z}}^N\|_{t_{1,st,1}; \mathbb{T}_N} + N^{-\frac{1}{2}+\varepsilon_{st}} + N^{-\frac{1}{2}+\varepsilon_{st}}\|\bar{\mathfrak{Z}}^N\|^2_{t_{1,st,1}; \mathbb{T}_N}. \qquad (8.34)$$

The last bound getting the far RHS of (8.34) follows by trivially bounding the first term in the previous bound by its space-time supremum over $[0, t_{1,st,1}] \times \mathbb{T}_N$. We now observe that because $t_{1,st} = t_{1,st,1} \wedge t_{1,st,2}$ while $t_{1,st,1} < 1$ and $t_{1,st,2} = 1$ by assumption on the event we are working on, we have $t_{1,st} = t_{1,st,1}$. Thus the pathwise identification in Lemma 8.8 gives

$$\|\bar{\mathfrak{Z}}^N\|_{t_{1,st,1}; \mathbb{T}_N} = \|\mathfrak{Q}^N\|_{t_{1,st,1}; \mathbb{T}_N} \leqslant \|\mathfrak{D}^N\|_{t_{1,st,1}; \mathbb{T}_N} + \|\mathfrak{Y}^N\|_{t_{1,st,1}; \mathbb{T}_N} \lesssim N^{\varepsilon_{st}}. \qquad (8.35)$$

The first identity follows from Lemma 8.8 and $t_{1,st} = t_{1,st,1}$ as we noted earlier. The second statement follows from definition of $\mathfrak{D}^N = \mathfrak{Y}^N - \mathfrak{Q}^N$ and the triangle inequality. The final estimate follows by the assumption that $t_{2,st}, t_{3,st} = 1$ which gives a priori

control on $\mathfrak{D}^N$ and $\mathfrak{Y}^N$ in the $\|\|_{1;\mathbb{T}_N}$-norm; recall $\mathfrak{t}_{1,\mathrm{st},1} \leqslant 1$. We now combine (8.34) and (8.35) to deduce the uniform bound

$$|\bar{\mathfrak{Z}}^N_{\mathfrak{t}_{1,\mathrm{st},1}+\mathfrak{s},x}| \lesssim N^{\varepsilon_{\mathrm{st}}} + N^{-1/2+3\varepsilon_{\mathrm{st}}} \lesssim N^{\varepsilon_{\mathrm{st}}}. \qquad (8.36)$$

This final estimate (8.36) is deterministic and the implied constant is universal. Moreover, (8.34) and (8.35) are uniform over $\mathfrak{s} \in [0, \mathfrak{t}_N]$ and $x \in \mathbb{T}_N$, and thus so is the estimate (8.36). Let us now observe that $\|\bar{\mathfrak{Z}}^N\|_{\mathfrak{t}_{1,\mathrm{st},1}+\mathfrak{t}_N;\mathbb{T}_N}$ is controlled by $\|\bar{\mathfrak{Z}}^N\|_{\mathfrak{t}_{1,\mathrm{st},1};\mathbb{T}_N}$ plus the supremum over $\mathfrak{s} \in [0, \mathfrak{t}_N]$ and $x \in \mathbb{T}_N$ of the LHS of (8.36), so $\|\bar{\mathfrak{Z}}^N\|_{\mathfrak{t}_{1,\mathrm{st},1}+\mathfrak{t}_N;\mathbb{T}_N} \lesssim N^{\varepsilon_{\mathrm{st}}}$ with universal implied constant; see (8.35) for control on $\|\bar{\mathfrak{Z}}^N\|_{\mathfrak{t}_{1,\mathrm{st},1};\mathbb{T}_N}$. As indicated in the discussion preceding (8.34), this completes the proof. $\qquad \square$

*Proof of Lemma 8.11.* Like with the proof of Lemma 2.5 in [10], we have the entropy production in Proposition 3.6 and spatial regularity of $\mathfrak{Y}^N$, so the one-block, two-blocks scheme implies that it suffices to prove the following expectation estimate which is pointwise on compact space-time sets; the following LHS/term we bound resembles $\widetilde{U}_{13}$ in the proof of Lemma 2.5 in [10]:

$$\sup_{\mathfrak{u}=1,2} \sup_{0\leqslant S\leqslant 1} \sup_{|y|\lesssim_\varphi N} \mathbf{E}|\mathfrak{A}^N_{S,y,\delta}||\mathfrak{Y}^N_{S,y}|^{\mathfrak{u}}$$
$$\overset{\bullet}{=} \sup_{\mathfrak{u}=1,2} \sup_{0\leqslant S\leqslant 1} \sup_{|y|\lesssim_\varphi N} \mathbf{E}|\left(\widetilde{\sum_{j=1,\ldots,\delta N^{1/2}} \eta_{S,y+j}}\right)||\mathfrak{Y}^N_{S,y}|^{\mathfrak{u}} \lesssim_{\delta,\varphi} N^{-\beta_{\mathrm{univ}}}. \qquad (8.37)$$

All estimates in this argument will be allowed to depend on $\delta$ and $\varphi$. For $\sigma \in \mathbb{R}_{>0}$ arbitrarily small but universal, as $|\mathfrak{A}^N| \lesssim 1$, we consider the following decomposition which is similar to the $D_1$ and $D_2$ events from the proof of Lemma 2.5 in [10] but with refinements to address the relevance of both of $\mathfrak{Z}^N$ and $\mathfrak{Y}^N$ as well as refinements in the cutoff-exponents for technical reasons; see the final estimate (8.57) of this proof for why we will need to choose the cutoff-exponent $\sigma$ appearing below carefully:

$$|\mathfrak{A}^N_{S,y,\delta}||\mathfrak{Y}^N_{S,y}|^{\mathfrak{u}} \leqslant |\mathfrak{A}^N_{S,y,\delta}||\mathfrak{Y}^N_{S,y}|^{\mathfrak{u}}\mathbf{1}_{|\mathfrak{Z}^N_{S,y}|\leqslant N^{-\sigma}}\mathbf{1}_{|\mathfrak{Y}^N_{S,y}|\leqslant N^{-\sigma/2}}$$
$$+ |\mathfrak{A}^N_{S,y,\delta}||\mathfrak{Y}^N_{S,y}|^{\mathfrak{u}}\mathbf{1}_{|\mathfrak{Z}^N_{S,y}|\leqslant N^{-\sigma}}\mathbf{1}_{|\mathfrak{Y}^N_{S,y}|\geqslant N^{-\sigma/2}}$$
$$+ |\mathfrak{A}^N_{S,y,\delta}||\mathfrak{Y}^N_{S,y}|^{\mathfrak{u}}\mathbf{1}_{|\mathfrak{Z}^N_{S,y}|\geqslant N^{-\sigma}}$$
$$\lesssim N^{-\frac{1}{2}\sigma}$$
$$+ |\mathfrak{A}^N_{S,y,\delta}||\mathfrak{Y}^N_{S,y}|^{\mathfrak{u}}\mathbf{1}_{|\mathfrak{Z}^N_{S,y}|\leqslant N^{-\sigma}}\mathbf{1}_{|\mathfrak{Y}^N_{S,y}|\geqslant N^{-\sigma/2}}$$
$$+ |\mathfrak{A}^N_{S,y,\delta}||\mathfrak{Y}^N_{S,y}|^{\mathfrak{u}}\mathbf{1}_{|\mathfrak{Z}^N_{S,y}|\geqslant N^{-\sigma}}. \qquad (8.38)$$

We first analyze the second term on the RHS of (8.38). To this end, we first observe the following for $N \in \mathbb{Z}_{>0}$ sufficiently large:

$$|\mathfrak{A}^N_{S,y,\delta}||\mathfrak{Y}^N_{S,y}|^{\mathfrak{u}}\mathbf{1}_{|\mathfrak{Z}^N_{S,y}|\leqslant N^{-\sigma}}\mathbf{1}_{|\mathfrak{Y}^N_{S,y}|\geqslant N^{-\sigma/2}}$$
$$\leqslant |\mathfrak{A}^N_{S,y,\delta}||\mathfrak{Y}^N_{S,y}|^{\mathfrak{u}}\mathbf{1}_{|\mathfrak{Z}^N_{S,y}|\leqslant N^{-\sigma}}\mathbf{1}_{|\mathfrak{Y}^N_{S,y}-\mathfrak{Z}^N_{S,y}|\gtrsim N^{-\sigma/2}}. \qquad (8.39)$$

Indeed, if $|\mathfrak{Z}^N| \leqslant N^{-\sigma}$ and $|\mathfrak{Y}^N| \geqslant N^{-\sigma/2}$, then the difference, in absolute value, is at least order $N^{-\sigma/2}$. As $|\mathfrak{A}^N| \lesssim 1$, we get

$$\mathbf{E}|\mathfrak{A}_{S,y,\delta}^N||\mathfrak{Y}_{S,y}^N|^{\mathfrak{u}}\mathbf{1}_{|\mathfrak{Z}_{S,y}^N|\leqslant N^{-\sigma}}\mathbf{1}_{|\mathfrak{Y}_{S,y}^N-\mathfrak{Z}_{S,y}^N|\gtrsim N^{-\sigma/2}}$$
$$\lesssim \left(\mathbf{E}|\mathfrak{Y}_{S,y}^N|^{2\mathfrak{u}}\right)^{1/2}\left(\mathbf{P}\left(|\mathfrak{Y}_{S,y}^N-\mathfrak{Z}_{S,y}^N|\gtrsim N^{-\frac{1}{2}\sigma}\right)\right)^{1/2}. \qquad (8.40)$$

Lemma 8.12 shows the $\mathfrak{Y}^N$-factor from the RHS of (8.40) is uniformly bounded. Proposition 8.2 applied using $\mathbb{K}$ any compact neighborhood of the support of our test function $\varphi$ controls the probability in the second factor on the RHS of (8.40). Thus we get the following if $\sigma \in \mathbb{R}_{>0}$ is sufficiently small but still universal and for $\beta_{\mathrm{univ},1} \in \mathbb{R}_{>0}$ universal and $|y| \lesssim_\varphi N$:

$$\left(\mathbf{E}|\mathfrak{Y}_{S,y}^N|^{2\mathfrak{u}}\right)^{1/2}\left(\mathbf{P}\left(|\mathfrak{Y}_{S,y}^N-\mathfrak{Z}_{S,y}^N|\gtrsim N^{-\frac{1}{2}\sigma}\right)\right)^{1/2} \lesssim \left(\mathbf{P}\left(|\mathfrak{Y}_{S,y}^N-\mathfrak{Z}_{S,y}^N|\gtrsim N^{-\frac{1}{2}\sigma}\right)\right)^{1/2}$$
$$\lesssim_\varphi N^{-\frac{1}{2}\beta_{\mathrm{univ},1}}. \qquad (8.41)$$

The implied constant in the final inequality of (8.41) depends on $\varphi$ through its support because the estimate in Proposition 8.2 depends on the compact set $\mathbb{K} \subseteq \mathbb{R}$ taken therein. Combining (8.39), (8.40), and (8.41) gives

$$\mathbf{E}|\mathfrak{A}_{S,y,\delta}^N||\mathfrak{Y}_{S,y}^N|^{\mathfrak{u}}\mathbf{1}_{|\mathfrak{Z}_{S,y}^N|\leqslant N^{-\sigma}}\mathbf{1}_{|\mathfrak{Y}_{S,y}^N|\geqslant N^{-\sigma/2}} \lesssim N^{-\frac{1}{2}\beta_{\mathrm{univ},1}}. \qquad (8.42)$$

We now analyze the last term on the RHS of (8.38) in expectation. For $\beta \in \mathbb{R}_{>0}$ small but to be determined, we decompose

$$|\mathfrak{A}_{S,y,\delta}^N||\mathfrak{Y}_{S,y}^N|^{\mathfrak{u}}\mathbf{1}_{|\mathfrak{Z}_{S,y}^N|\geqslant N^{-\sigma}} \leqslant |\mathfrak{A}_{S,y,\delta}^N||\mathfrak{Y}_{S,y}^N|^{\mathfrak{u}}\mathbf{1}_{|\mathfrak{Z}_{S,y}^N|\geqslant N^{-\sigma}}\mathbf{1}_{|\mathfrak{Y}_{S,\star}^N-\mathfrak{Z}_{S,\star}^N|\geqslant N^{-\beta}}$$
$$+ |\mathfrak{A}_{S,y,\delta}^N||\mathfrak{Y}_{S,y}^N|^{\mathfrak{u}}\mathbf{1}_{|\mathfrak{Z}_{S,y}^N|\geqslant N^{-\sigma}}\mathbf{1}_{|\mathfrak{Y}_{S,\star}^N-\mathfrak{Z}_{S,\star}^N|\leqslant N^{-\beta}}. \qquad (8.43)$$

In the newly introduced event $\star \in \{y, y + \delta N^{1/2}\}$ is allowed to take two values, so this event is defined by two constraints. Like the first term on the RHS of (8.43) and the estimate (8.40), for the first term on the RHS of (8.43) we employ Lemma 8.12 and Proposition 8.2 with $\mathbb{K}$ any compact neighborhood of the support of $\varphi$, and we choose $\beta \in \mathbb{R}_{>0}$ sufficiently small to get

$$\mathbf{E}|\mathfrak{A}_{S,y,\delta}^N||\mathfrak{Y}_{S,y}^N|^{\mathfrak{u}}\mathbf{1}_{|\mathfrak{Z}_{S,y}^N|\geqslant N^{-\sigma}}\mathbf{1}_{|\mathfrak{Y}_{S,\star}^N-\mathfrak{Z}_{S,\star}^N|\geqslant N^{-\beta}}$$
$$\lesssim \mathbf{E}|\mathfrak{Y}_{S,y}^N|^{\mathfrak{u}}\mathbf{1}_{|\mathfrak{Y}_{S,\star}^N-\mathfrak{Z}_{S,\star}^N|\geqslant N^{-\beta}} \qquad (8.44)$$
$$\lesssim \left(\mathbf{E}|\mathfrak{Y}_{S,y}^N|^{2\mathfrak{u}}\right)^{1/2}\left(\mathbf{P}\left(|\mathfrak{Y}_{S,\star}^N-\mathfrak{Z}_{S,\star}^N|\geqslant N^{-\beta}\right)\right)^{1/2} \qquad (8.45)$$
$$\lesssim_\varphi N^{-\frac{1}{2}\beta_{\mathrm{univ},1}}. \qquad (8.46)$$

The last bound (8.46) follows by observing $\star \in N\mathbb{K}$ from which we obtain high-probability comparison via Proposition 8.2.

We now address second term on the RHS of (8.43). By definition of $\mathfrak{h}^N$ and $\mathfrak{Z}^N$, as in the proof of Lemma 2.5 in [10] we get

$$|\mathfrak{A}_{S,y,\delta}^N| \lesssim \delta^{-1}|\log(1 + (\mathfrak{Z}_{S,y}^N)^{-1}\nabla_{\delta N^{1/2}}\mathfrak{Z}_{S,y}^N)|. \qquad (8.47)$$

At this point, we differ slightly from the proof of Lemma 2.5 in [10]. In particular, we observe that the term $(\mathfrak{Z}_{S,y}^N)^{-1}\nabla_{\delta N^{1/2}}\mathfrak{Z}_{S,y}^N$ is at most the power series for the exponential evaluated at something that is at most $\delta$ and without the constant term. This is a consequence of Taylor expansion and definition of $\mathfrak{Z}^N$. More precisely, we have the deterministic inequality

$$|(\mathfrak{Z}_{S,y}^N)^{-1}\nabla_{\delta N^{1/2}}\mathfrak{Z}_{S,y}^N| \lesssim \sum_{j=1}^\infty |\lambda|^j (j!)^{-1}|\delta|^j \lesssim_\lambda |\delta|. \tag{8.48}$$

The second estimate (8.48) follows from choosing $\delta \in \mathbb{R}_{>0}$ sufficiently small but depending only on $\lambda \in \mathbb{R}$. We now note, like in the proof of Lemma 2.5 in [10], that $|\log(1+x)| \lesssim_\delta |x|$ for $|x| \lesssim \delta$. Thus, combining this with (8.47) and (8.48), we have

$$|\mathfrak{A}_{S,y,\delta}^N||\mathfrak{Y}_{S,y}^N|^{\mathfrak{u}}\mathbf{1}_{|\mathfrak{Z}_{S,y}^N|\geqslant N^{-\sigma}}\mathbf{1}_{|\mathfrak{Y}_{S,\star}^N-\mathfrak{Z}_{S,\star}^N|\leqslant N^{-\beta}}$$

$$\lesssim_\lambda \delta^{-1}|(\mathfrak{Z}_{S,y}^N)^{-1}\nabla_{\delta N^{1/2}}\mathfrak{Z}_{S,y}^N||\mathfrak{Y}_{S,y}^N|^{\mathfrak{u}}\mathbf{1}_{|\mathfrak{Z}_{S,y}^N|\geqslant N^{-\sigma}}\mathbf{1}_{|\mathfrak{Y}_{S,\star}^N-\mathfrak{Z}_{S,\star}^N|\leqslant N^{-\beta}} \tag{8.49}$$

$$\lesssim \delta^{-1}N^\sigma|\nabla_{\delta N^{1/2}}\mathfrak{Z}_{S,y}^N||\mathfrak{Y}_{S,y}^N|^{\mathfrak{u}}\mathbf{1}_{|\mathfrak{Z}_{S,y}^N|\geqslant N^{-\sigma}}\mathbf{1}_{|\mathfrak{Y}_{S,\star}^N-\mathfrak{Z}_{S,\star}^N|\leqslant N^{-\beta}}. \tag{8.50}$$

On the event corresponding to the second indicator function in (8.50), the gradient in question of $\mathfrak{Z}^N$ in (8.50) can be bounded in terms of the same gradient but of $\mathfrak{Y}^N$ up to an admissible error. More precisely, we have

$$|\nabla_{\delta N^{1/2}}\mathfrak{Z}_{S,y}^N|\mathbf{1}_{|\mathfrak{Z}_{S,y}^N|\geqslant N^{-\sigma}}\mathbf{1}_{|\mathfrak{Y}_{S,\star}^N-\mathfrak{Z}_{S,\star}^N|\leqslant N^{-\beta}} \lesssim |\nabla_{\delta N^{1/2}}\mathfrak{Z}_{S,y}^N|\mathbf{1}_{|\mathfrak{Y}_{S,\star}^N-\mathfrak{Z}_{S,\star}^N|\leqslant N^{-\beta}}$$

$$\lesssim |\nabla_{\delta N^{1/2}}\mathfrak{Y}_{S,y}^N|\mathbf{1}_{|\mathfrak{Y}_{S,\star}^N-\mathfrak{Z}_{S,\star}^N|\leqslant N^{-\beta}} + N^{-\beta}. \tag{8.51}$$

This step is the precise implementation of transferring relevant $\mathfrak{Z}^N$-data in the proof of Lemma 8.11 into $\mathfrak{Y}^N$-estimates that we discussed in the proof of Proposition 8.3. By (8.50), (8.51), the Cauchy-Schwarz inequality, and Lemma 8.12, we get

$$\mathbf{E}|\mathfrak{A}_{S,y,\delta}^N||\mathfrak{Y}_{S,y}^N|^{\mathfrak{u}}\mathbf{1}_{|\mathfrak{Z}_{S,y}^N|\geqslant N^{-\sigma}}\mathbf{1}_{|\mathfrak{Y}_{S,\star}^N-\mathfrak{Z}_{S,\star}^N|\leqslant N^{-\beta}}$$

$$\lesssim \delta^{-1}N^\sigma\mathbf{E}|\nabla_{\delta N^{1/2}}\mathfrak{Z}_{S,y}^N||\mathfrak{Y}_{S,y}^N|^{\mathfrak{u}}\mathbf{1}_{|\mathfrak{Z}_{S,y}^N|\geqslant N^{-\sigma}}\mathbf{1}_{|\mathfrak{Y}_{S,\star}^N-\mathfrak{Z}_{S,\star}^N|\leqslant N^{-\beta}} \tag{8.52}$$

$$\lesssim \delta^{-1}N^\sigma\mathbf{E}|\nabla_{\delta N^{1/2}}\mathfrak{Y}_{S,y}^N||\mathfrak{Y}_{S,y}^N|^{\mathfrak{u}}\mathbf{1}_{|\mathfrak{Y}_{S,\star}^N-\mathfrak{Z}_{S,\star}^N|\leqslant N^{-\beta}} + \delta^{-1}N^\sigma N^{-\beta}\mathbf{E}|\mathfrak{Y}_{S,y}^N|^{\mathfrak{u}} \tag{8.53}$$

$$\lesssim \delta^{-1}N^\sigma(\mathbf{E}|\nabla_{\delta N^{1/2}}\mathfrak{Y}_{S,y}^N|^2\mathbf{1}_{|\mathfrak{Y}_{S,\star}^N-\mathfrak{Z}_{S,\star}^N|\leqslant N^{-\beta}})^{\frac{1}{2}}(\mathbf{E}|\mathfrak{Y}_{S,y}^N|^{2\mathfrak{u}})^{\frac{1}{2}} + \delta^{-1}N^{\sigma-\beta} \tag{8.54}$$

$$\lesssim \delta^{-1}N^\sigma(\mathbf{E}|\nabla_{\delta N^{1/2}}\mathfrak{Y}_{S,y}^N|^2)^{\frac{1}{2}} + \delta^{-1}N^{\sigma-\beta}. \tag{8.55}$$

To estimate the first term, provided Lemma 8.12 note that the spatial-regularity estimate (3.13) from Proposition 3.2 of [10] still holds for $\mathfrak{Y}^N$, again because the estimate therein for the $Z$-field therein comes just from its mild form (3.2) in [10], which is sufficiently identical to the defining equation for $\mathfrak{Y}^N$. We used this spatial regularity of $\mathfrak{Y}^N$ in qualitative fashion earlier at the beginning to reduce proof of Lemma 8.11 into proof of (8.37). Ultimately we get the following Holder bound for any $\vartheta < 1$:

$$\delta^{-1}N^\sigma(\mathbf{E}|\nabla_{\delta N^{1/2}}\mathfrak{Y}_{S,y}^N|^2)^{\frac{1}{2}} \lesssim_\vartheta \delta^{-1}N^\sigma \cdot \delta^{\frac{1}{2}\vartheta}N^{\frac{1}{4}\vartheta} \cdot N^{-\frac{1}{2}\vartheta}. \tag{8.56}$$

We combine (8.55) and (8.56) to deduce

$$\mathbf{E}|\mathfrak{A}_{S,y,\delta}^{N}||\mathfrak{Y}_{S,y}^{N}|^{\mathrm{u}}\mathbf{1}_{|3_{S,y}^{N}|\geqslant N^{-\sigma}}\mathbf{1}_{|\mathfrak{Y}_{S,\star}^{N}-3_{S,\star}^{N}|\leqslant N^{-\beta}} \lesssim_{\vartheta} \delta^{-1+\frac{1}{2}\vartheta}N^{\sigma-\frac{1}{4}\vartheta}+\delta^{-1}N^{\sigma-\beta}. \tag{8.57}$$

The estimate (8.37) follows from combining (8.38), (8.42), (8.43), (8.46), and (8.57) with the last of these bounds employed having chosen $\sigma \in \mathbb{R}_{>0}$ sufficiently small depending on $\beta \in \mathbb{R}_{>0}$ and $\vartheta < 1$. This completes the proof. $\qquad\square$

*Proof of Lemma 8.4.* We begin with the following general properties of $\mathbf{D}_1$ in [3].

- The set $\mathbf{D}_1$ is metrizable with metric bounded by the uniform norm; in the definition (12.13) on page 124 in [3] but replacing $\mathbb{R}$-valued paths with $\mathbf{C}(\mathbb{K})$-valued paths, we choose the identity function in the infimum and recover the uniform metric.

Thus $\mathfrak{X}^{N,1}-\mathfrak{X}^{N,2}$ converges to the 0 process in probability in $\mathbf{D}_1$. As $\mathfrak{X}^{N,1}=\mathfrak{X}^{N,2}+(\mathfrak{X}^{N,1}-\mathfrak{X}^{N,2})$ and $\mathfrak{X}^{N,2}$ converges in law to $\mathfrak{X}^{\infty,2}$ while $\mathfrak{X}^{N,1}-\mathfrak{X}^{N,2}$ converges *in probability* to 0, standard probability on separable metric spaces finishes the proof. $\quad\square$

## Appendix A. Heat Operator Estimates

We collect heat kernel estimates for $\mathfrak{P}^N$ and $\bar{\mathfrak{P}}^N$ based on Proposition A.1/Corollary A.2 from [10]. The following estimates are effectively the heat kernel estimates that are used in [10] as well as versions for the compactified heat operator $\bar{\mathfrak{P}}^N$ though with a few elementary additional ingredients.

**Lemma A.1.** *Recall* Definition *4.15 for the entire statement of this lemma. Provided any* $\kappa \in \mathbb{R}_{\geqslant 0}$, *uniformly bounded* $k \in \mathbb{Z}$, *and uniformly bounded* $S, T \in \mathbb{R}_{\geqslant 0}$ *with* $S \leqslant T$, *we have the following pointwise off-diagonal heat kernel estimates for* $\mathfrak{P}^N$ *and* $\bar{\mathfrak{P}}^N$:

$$\sup_{x,y\in\mathbb{Z}}\left(e_{S,T,x,y}^{N,\kappa}\mathfrak{P}_{S,T,x,y}^{N}\right)+\sup_{x,y\in\mathbb{T}_N}\left(e_{S,T,x,y}^{N,\kappa,\mathrm{cpt}}\bar{\mathfrak{P}}_{S,T,x,y}^{N}\right)\lesssim_{\kappa} N^{-1}\mathfrak{s}_{S,T}^{-\frac{1}{2}}\wedge 1 \tag{A.1}$$

$$\sup_{x,y\in\mathbb{Z}}\left(e_{S,T,x,y}^{N,\kappa}|\nabla_k^!\mathfrak{P}_{S,T,x,y}^{N}|\right)+\sup_{x,y\in\mathbb{T}_N}\left(e_{S,T,x,y}^{N,\kappa,\mathrm{cpt}}|\bar{\nabla}_k^!\bar{\mathfrak{P}}_{S,T,x,y}^{N}|\right)\lesssim_{\kappa} |k|\left(N^{-1}\mathfrak{s}_{S,T}^{-1}\wedge N\right). \tag{A.2}$$

*Provided $\varphi : \mathbb{R}_{\geqslant 0} \times \mathbb{Z} \to \mathbb{R}$ and $\bar{\varphi} : \mathbb{R}_{\geqslant 0} \times \mathbb{T}_N \to \mathbb{R}$, any subset $I \subseteq \mathbb{R}_{\geqslant 0}$ with measure $|I|$, any uniformly bounded $\mathfrak{t}_{st} \in \mathbb{R}_{\geqslant 0}$, any uniformly bounded $\mathfrak{t}_{reg} \in \mathbb{R}_{\leqslant 0}$, any $\delta \in \mathbb{R}_{>0}$, any $k \in \mathbb{Z}$ uniformly bounded, and any $\mathfrak{l} \in \mathbb{Z}$, we have*

$$\|\varphi\|_{\mathfrak{t}_{st};\mathbb{Z}}^{-1}\|\mathfrak{P}^N(\varphi_{S,y}\mathbf{1}_{S\in I})\|_{\mathfrak{t}_{st};\mathbb{Z}} + \|\bar{\varphi}\|_{\mathfrak{t}_{st};\mathbb{T}_N}^{-1}\|\bar{\mathfrak{P}}^N(\bar{\varphi}_{S,y}\mathbf{1}_{S\in I})\|_{\mathfrak{t}_{st};\mathbb{T}_N} \lesssim |I| \wedge \mathfrak{t}_{st} \tag{A.3}$$

$$\|\varphi\|_{\mathfrak{t}_{st};\mathbb{Z}}^{-1}\|\mathbf{1}_{T\geqslant|\mathfrak{t}_{reg}|}\mathfrak{s}_{|\mathfrak{t}_{reg}|,T}^{1-\delta}\nabla_{\mathfrak{t}_{reg}}^{\mathbf{T}}\mathfrak{P}_{T,x}^{N,\mathbf{X}}(\varphi)\|_{\mathfrak{t}_{st};\mathbb{Z}}$$
$$+ \|\bar{\varphi}\|_{\mathfrak{t}_{st};\mathbb{T}_N}^{-1}\|\mathbf{1}_{T\geqslant|\mathfrak{t}_{reg}|}\mathfrak{s}_{|\mathfrak{t}_{reg}|,T}^{1-\delta}\nabla_{\mathfrak{t}_{reg}}^{\mathbf{T}}\bar{\mathfrak{P}}_{T,x}^{N,\mathbf{X}}(\bar{\varphi})\|_{\mathfrak{t}_{st};\mathbb{T}_N} \lesssim_\delta N^{2\delta}|\mathfrak{t}_{reg}| \tag{A.4}$$

$$\|\varphi\|_{\mathfrak{t}_{st};\mathbb{Z}}^{-1}\|\mathbf{1}_{T\geqslant|\mathfrak{t}_{reg}|}\nabla_{\mathfrak{t}_{reg}}^{\mathbf{T}}\mathfrak{P}_{T,x}^{N}(\varphi)\|_{\mathfrak{t}_{st};\mathbb{Z}} + \|\bar{\varphi}\|_{\mathfrak{t}_{st};\mathbb{T}_N}^{-1}\|\mathbf{1}_{T\geqslant|\mathfrak{t}_{reg}|}\nabla_{\mathfrak{t}_{reg}}^{\mathbf{T}}\bar{\mathfrak{P}}_{T,x}^{N}(\bar{\varphi})\|_{\mathfrak{t}_{st};\mathbb{T}_N}$$
$$\lesssim_\delta N^{2\delta}|\mathfrak{t}_{reg}| \tag{A.5}$$

$$\|\varphi\|_{\mathfrak{t}_{st};\mathbb{Z}}^{-1}\|\mathbf{1}_{T\geqslant|\mathfrak{t}_{reg}|}\nabla_{\mathfrak{t}_{reg}}^{\mathbf{T}}\mathfrak{P}_{T,x}^{N}(\nabla_k^!\varphi)\|_{\mathfrak{t}_{st};\mathbb{Z}} + \|\bar{\varphi}\|_{\mathfrak{t}_{st};\mathbb{T}_N}^{-1}\|\mathbf{1}_{T\geqslant|\mathfrak{t}_{reg}|}\nabla_{\mathfrak{t}_{reg}}^{\mathbf{T}}\bar{\mathfrak{P}}_{T,x}^{N}(\bar{\nabla}_k^!\bar{\varphi})\|_{\mathfrak{t}_{st};\mathbb{T}_N}$$
$$\lesssim_\delta |\mathfrak{t}_{reg}|^{\frac{1}{4}-\delta} \tag{A.6}$$

$$\|\varphi\|_{\mathfrak{t}_{st};\mathbb{Z}}^{-1}\|\mathfrak{P}^N(\nabla_\mathfrak{l}^!\varphi)\|_{\mathfrak{t}_{st};\mathbb{Z}} + \|\bar{\varphi}\|_{\mathfrak{t}_{st};\mathbb{T}_N}^{-1}\|\bar{\mathfrak{P}}^N(\bar{\nabla}_\mathfrak{l}^!\bar{\varphi})\|_{\mathfrak{t}_{st};\mathbb{T}_N}$$
$$\lesssim |\mathfrak{l}|. \tag{A.7}$$

*Proof.* We first discuss $\mathfrak{P}^N$-estimates. The $\mathfrak{P}^N$ estimates in (A.1) and (A.2) follow from Proposition A.1 in [10]. The $\mathfrak{P}^N$-bound in (A.3) follows by the fact that the spatial heat operator $\mathfrak{P}^{N,\mathbf{X}}$ is a contraction with respect to the supremum-norm on functions on $\mathbb{Z}$ and $\mathfrak{P}^N$ integrates $\mathfrak{P}^{N,\mathbf{X}}$-operators on $I \cap [0, \mathfrak{t}_{st}]$ on the LHS of (A.3). The other $\mathfrak{P}^N$-bounds start with the observation that by definition of $\mathfrak{P}^N$, we have a fundamental-theorem-of-calculus estimate where $\delta \in \mathbb{R}_{>0}$ is arbitrarily small and $S + |\mathfrak{t}_{reg}| \leqslant T$. Below, $\nabla^{\mathbf{T}}$ on the far LHS below acts on the heat kernel via its forwards time-variable and $\|\varphi\| = \|\varphi\|_{T;\mathbb{Z}}$ for convenience:

$$\sum_{y\in\mathbb{Z}}|\nabla_{\mathfrak{t}_{reg}}^{\mathbf{T}}\mathfrak{P}_{S,T,x,y}^{N}||\varphi_{S,y}| \lesssim \|\varphi\|\int_{\mathfrak{t}_{reg}}^{0}\sum_{y\in\mathbb{Z}}|\mathscr{L}^{!!}\mathfrak{P}_{S,T+\mathfrak{r},x,y}^{N}|d\mathfrak{r} \lesssim_m N^{2\delta}\|\varphi\|\int_{\mathfrak{t}_{reg}}^{0}\mathfrak{s}_{S,T+\mathfrak{r}}^{-1+\delta}d\mathfrak{r}$$
$$\lesssim N^{2\delta}|\mathfrak{t}_{reg}|\mathfrak{s}_{S+|\mathfrak{t}_{reg}|,T}^{-1+\delta}\|\varphi\|. \tag{A.8}$$

The second estimate above follows by (A.29) in Corollary A.2 in [10] for $\upsilon = 1 - 2\delta$, and the last estimate follows by observing the integrand is decreasing in $\mathfrak{r} \in [\mathfrak{t}_{reg}, 0]$; the heat kernel here is scaled by $N^2$ in time but the heat kernel in [10] is not.

- As $T \leqslant \mathfrak{t}_{st}$, the norm $\|\varphi\|$ in (A.8) can be extended to $\|\varphi\|_{\mathfrak{t}_{st};\mathbb{Z}}$. We get from this and (A.8) for $S = 0$ the $\mathfrak{P}^N$-bound in (A.4).
- To obtain the $\mathfrak{P}^N$ estimate in (A.5), we decompose the space-time heat operator $\nabla^{\mathbf{T}}\mathfrak{P}^N$ into a time-integral of the 1-norm on $\mathbb{Z}$ of the time-gradient of the heat kernel, which we denote by $\Phi_1$, and a short-time integral of the spatial heat operator, which we denote by $\Phi_2$. This was the decomposition used to get $I_{31}$ and $I_{32}$ in the proof of (3.14) in Proposition 3.2 in [10]. For the heat-kernel-time-gradient term $\Phi_1$, we integrate (A.8) in $S \in [0, T - |\mathfrak{t}_{reg}|]$. For the short-time integral $\Phi_2$, we employ (A.3) for $I = T + [\mathfrak{t}_{reg}, 0]$. Again we refer to $I_{31}$ and $I_{32}$ terms and their estimates in the proof of (3.14) in Proposition 3.2 in [10].
- The proof of the $\mathfrak{P}^N$-estimate in (A.6) follows via the proof of the estimates for $I_{41}$ and $I_{42}$ in the proof of (3.14) in Proposition 3.2 in [10]. Roughly speaking, we move the gradient from $\varphi$ onto the heat kernel and employ space-time regularity estimates for the heat kernel given in (A.10) from Proposition A.1 in [10] and (A.27) from Corollary A.2 in [10].

- The proof of the $\mathfrak{P}^N$ estimate in (A.7) follows by summation-by-parts to move the gradient from $\varphi$ to the heat kernel, pulling out the $\varphi$-factor in the resulting spatial gradient of the $\mathfrak{P}^N$ heat operator via its $\|\|\|_{\mathsf{t}_{\mathrm{st}};\mathbb{Z}}$-norm, applying (A.27) in Corollary A.2 in [10] with $v = 1$ to control the 1-norm/sum of the absolute values of the spatial gradient of the heat kernel, and integrating the resulting integrable singularity in time. The summation-by-parts is the same giving (3.2) in [10], and the estimate noted in this last sentence is the same that gives the $I_4$ estimate after (3.23) in the proof of (3.12) in Proposition 3.2 in [10].

The proof of the proposed $\bar{\mathfrak{P}}^N$-bounds follows by identical considerations; see Remark 4.2. This completes the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

We now collect a time-regularity estimate for $\bar{\mathfrak{P}}^N$ which requires an additional integration-by-parts-type ingredient, though it is still an integrated time-smoothness estimate for the heat operator like (A.5) and requires only a little more gymnastics.

**Lemma A.2.** *For any possibly random* $\bar{\varphi} : \mathbb{R}_{\geqslant 0} \times \mathbb{T}_N \to \mathbb{R}$, *any* $\mathsf{t}_{\mathrm{st}}, \mathsf{t}_{\mathrm{reg}} \in \mathbb{R}_{\geqslant 0}$ *uniformly bounded such that* $\mathsf{t}_{\mathrm{st}}$ *is possibly random and with probability 1 we have* $\mathsf{t}_{\mathrm{st}} \leqslant 1$, *and any* $\delta \in \mathbb{R}_{>0}$, *we have the following forwards-time-gradient estimate:*

$$\|\bar{\varphi}\|_{\mathsf{t}_{\mathrm{st}}+\mathsf{t}_{\mathrm{reg}};\mathbb{T}_N}^{-1} \|\bar{\mathfrak{P}}_{T,x}^N(\nabla_{\mathsf{t}_{\mathrm{reg}}}^{\mathbf{T}}\bar{\varphi})\|_{\mathsf{t}_{\mathrm{st}};\mathbb{T}_N} \lesssim_\delta N^{2\delta}|\mathsf{t}_{\mathrm{reg}}|. \qquad (A.9)$$

*Proof.* Similar to (7.40) we have the following time-discrete Leibniz rule in which $\nabla^{\mathbf{T}}$ acts only on $S$-variables:

$$\begin{aligned}
\bar{\mathfrak{P}}_{S,T,x,y}^N \mathbf{1}_{S\in[0,T]} \nabla_{\mathsf{t}_{\mathrm{reg}}}^{\mathbf{T}}\bar{\varphi}_{S,y} &= \nabla_{\mathsf{t}_{\mathrm{reg}}}^{\mathbf{T}}\left(\bar{\mathfrak{P}}_{S-\mathsf{t}_{\mathrm{reg}},T,x,y}^N \mathbf{1}_{S-\mathsf{t}_{\mathrm{reg}}\in[0,T]}\bar{\varphi}_{S,y}\right) \\
&\quad + \bar{\varphi}_{S,y}\nabla_{-\mathsf{t}_{\mathrm{reg}}}^{\mathbf{T}}\left(\bar{\mathfrak{P}}_{S,T,x,y}^N \mathbf{1}_{S\in[0,T]}\right). \qquad (A.10)
\end{aligned}$$

The identity (A.10) is (7.40) after replacing the $\bar{\mathfrak{Z}}^N$-process with the product of the $\bar{\mathfrak{P}}^N$-heat kernel and the indicator function. We proceed with expansion of the time-gradient in the second term on the RHS of (A.10) which also resembles a time-discrete Leibniz rule similar to (A.10). The following can be checked by an elementary calculation; again $\nabla^{\mathbf{T}}$ acts only on $S$-variables:

$$\begin{aligned}
\bar{\varphi}_{S,y}\nabla_{-\mathsf{t}_{\mathrm{reg}}}^{\mathbf{T}}\left(\bar{\mathfrak{P}}_{S,T,x,y}^N \mathbf{1}_{S\in[0,T]}\right) &= \bar{\varphi}_{S,y}\left(\nabla_{-\mathsf{t}_{\mathrm{reg}}}^{\mathbf{T}}\bar{\mathfrak{P}}_{S,T,x,y}^N\right)\mathbf{1}_{S\in[0,T]} \\
&\quad + \bar{\varphi}_{S,y}\bar{\mathfrak{P}}_{S-\mathsf{t}_{\mathrm{reg}},T,x,y}^N\left(\nabla_{-\mathsf{t}_{\mathrm{reg}}}^{\mathbf{T}}\mathbf{1}_{S\in[0,T]}\right) \overset{\bullet}{=} \Phi_1 + \Phi_2. \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (A.11)
\end{aligned}$$

We will now sum each term on the LHS/RHS of (A.10) over $y \in \mathbb{T}_N$ and additionally integrate over $S \in \mathbb{R}$.

- Summation over $y \in \mathbb{T}_N$ and integration over $S \in \mathbb{R}$ of the LHS of (A.10) gives the heat operator $\bar{\mathfrak{P}}_{T,x}^N(\nabla_{\mathsf{t}_{\mathrm{reg}}}^{\mathbf{T}}\bar{\varphi})$ by definition.
- Integrating the first term on the RHS of (A.10) over $S \in \mathbb{R}$ yields 0 because the integral of the time-discrete gradient of any compactly supported function vanishes; note that the first term on the RHS of (A.10) is compactly supported because of the indicator function. This general observation about integrating time-discrete gradients is analogous to the fact that summing a compactly supported discrete gradient on $\mathbb{Z}$ gives 0, or that integrating a compactly supported derivative on $\mathbb{R}$ gives 0.

- We now treat the second term on the RHS of (A.10) via (A.11). We may let the $\nabla^{\mathbf{T}}$-operator in $\Phi_1$ act instead on the forwards time-variable of the heat kernel if we change $-t_{reg} \to t_{reg}$. This is because the heat kernel is time-homogeneous. Elementary calculation lets us *then* replace $t_{reg} \to -t_{reg}$ if we replace $T \to T + t_{reg}$ and introduce a factor of $-1$. With these changes,

$$\left\| \left( \int_{\mathbb{R}} \sum_{y \in \mathbb{T}_N} \left( \nabla^{\mathbf{T}}_{-t_{reg}} \bar{\mathfrak{P}}^N_{S,T+t_{reg},x,y} \right) \mathbf{1}_{S \in [0,T]} \bar{\varphi}_{S,y} dS \right) \right\|_{t_{st}; \mathbb{T}_N}$$

$$\leqslant \|\bar{\varphi}\|_{t_{st}+t_{reg}; \mathbb{T}_N} \left\| \left( \int_0^T \sum_{y \in \mathbb{T}_N} |\nabla^{\mathbf{T}}_{-t_{reg}} \bar{\mathfrak{P}}^N_{S,T+t_{reg},x,y}| \, dS \right) \right\|_{t_{st}; \mathbb{T}_N}. \quad (A.12)$$

Note the $t_{reg}$-parameter in (A.8) is non-positive whereas $t_{reg}$ here is non-negative. Thus we may apply (A.8), with $t_{reg}$ therein equal to $-t_{reg}$ here and $T$ therein equal to $T + t_{reg}$ here, to the RHS of (A.12). Recalling the term in the integration/summation on the LHS of (A.12) is $\Phi_1$ from (A.11) by definition, after applying (A.8) as just described to (A.12) we get, for any $\delta \in \mathbb{R}_{>0}$,

$$\left\| \left( \int_{\mathbb{R}} \sum_{y \in \mathbb{T}_N} \Phi_1 dS \right) \right\|_{t_{st}; \mathbb{T}_N} \lesssim N^{2\delta} |t_{reg}| \|\bar{\varphi}\|_{t_{st}+t_{reg}; \mathbb{T}_N} \left\| \left( \int_0^T \mathfrak{s}^{-1+\delta}_{S+|t_{reg}|, T+|t_{reg}|} dS \right) \right\|_{t_{st}; \mathbb{T}_N}$$

$$\lesssim_\delta N^{2\delta} |t_{reg}| \|\bar{\varphi}\|_{t_{st}+t_{reg}; \mathbb{T}_N}. \quad (A.13)$$

The last estimate in (A.13) follows by integrating the integral in the middle term in (A.13) to get a space-time uniform bound.

- Finally for the $\Phi_2$-term in (A.11), the time-gradient therein is controlled by the indicator function for two intervals of length $t_{reg}$. Thus $\Phi_2$ gives a space-time heat operator acting on $\bar{\varphi}$ times the indicator function $\mathbf{1}_{S \in I}$ with $|I| \lesssim t_{reg}$. Technically, when we integrate $\Phi_2$ in $S \in \mathbb{R}$ we may employ change-of-variables $S - t_{reg} \to S$ to remove the cosmetic time-shift $-t_{reg}$ in the heat kernel in $\Phi_2$. Thus we employ (A.3) to deduce the following estimate where the $\|\bar{\varphi}\|_{t_{st}+t_{reg}; \mathbb{T}_N}$-factor comes from the fact that $\Phi_2$ and every other "heat-term" in this proof only sees the test function $\bar{\varphi}$ until the time $t_{st} + t_{reg}$:

$$\left\| \left( \int_{\mathbb{R}} \sum_{y \in \mathbb{T}_N} \Phi_2 dS \right) \right\|_{t_{st}; \mathbb{T}_N} \lesssim \|\bar{\varphi}\|_{t_{st}+t_{reg}; \mathbb{T}_N} |I| \lesssim \|\bar{\varphi}\|_{t_{st}+t_{reg}; \mathbb{T}_N} |t_{reg}|. \quad (A.14)$$

The proof of Lemma A.2 follows immediately upon combining the previous bullet points with (A.10) and (A.11). $\qquad \square$

**Lemma A.3.** *Consider any function $\bar{\varphi} : \mathbb{R}_{\geqslant 0} \times \mathbb{T}_N \to \mathbb{R}$ and possibly random time $t_{1,st} \in \mathbb{R}_{\geqslant 0}$ satisfying $t_{1,st} \leqslant 1$ with probability 1. We emphasize $\bar{\varphi}$ may be random. We have the following deterministic estimate for a space-time heat operator with time-cutoff:*

$$\|\bar{\mathfrak{P}}^N_{T,x}(\bar{\varphi}_{S,y} \mathbf{1}_{S \leqslant t_{1,st}})\|_{1; \mathbb{T}_N} \leqslant \|\bar{\mathfrak{P}}^N_{T,x}(\bar{\varphi}_{S,y})\|_{t_{1,st}; \mathbb{T}_N}. \quad (A.15)$$

*Proof.* We consider the following two cases distinguished by $t_{1,st}$. In what follows we assume $T \in \mathbb{R}_{\geqslant 0}$ satisfies $T \leqslant 1$.

- Take any $T \leqslant \mathfrak{t}_{1,\mathrm{st}}$. We have the following two bounds which certainly suffice for this case. The first bound is straightforward and the second bound follows from the cutoff $\mathbf{1}_{S \leqslant \mathfrak{t}_{1,\mathrm{st}}}$ being already built into the heat operator $\bar{\mathfrak{P}}^N$ for $T \leqslant \mathfrak{t}_{1,\mathrm{st}}$:

$$|\bar{\mathfrak{P}}^N_{T,x}(\bar{\varphi}_{S,y}\mathbf{1}_{S \leqslant \mathfrak{t}_{1,\mathrm{st}}})| \;\leqslant\; \|\bar{\mathfrak{P}}^N_{T,x}(\bar{\varphi}_{S,y}\mathbf{1}_{S \leqslant \mathfrak{t}_{1,\mathrm{st}}})\|_{\mathfrak{t}_{1,\mathrm{st}};\mathbb{T}_N} \;=\; \|\bar{\mathfrak{P}}^N_{T,x}(\bar{\varphi}_{S,y})\|_{\mathfrak{t}_{1,\mathrm{st}};\mathbb{T}_N}.$$
(A.16)

- Take $T \geqslant \mathfrak{t}_{1,\mathrm{st}}$. By definition of the space-time heat operator and the Chapman-Kolmogorov equation for the heat kernel,

$$\begin{aligned}
\bar{\mathfrak{P}}^N_{T,x}(\bar{\varphi}_{S,y}\mathbf{1}_{S \leqslant \mathfrak{t}_{1,\mathrm{st}}}) \;&=\; \int_0^{\mathfrak{t}_{1,\mathrm{st}}} \sum\nolimits_{y \in \mathbb{T}_N} \bar{\mathfrak{P}}^N_{S,T,x,y} \bar{\varphi}_{S,y} \, \mathrm{d}S \\
&=\; \sum\nolimits_{w \in \mathbb{T}_N} \bar{\mathfrak{P}}^N_{\mathfrak{t}_{1,\mathrm{st}},T,x,w} \left( \int_0^{\mathfrak{t}_{1,\mathrm{st}}} \sum\nolimits_{y \in \mathbb{T}_N} \bar{\mathfrak{P}}^N_{S,\mathfrak{t}_{1,\mathrm{st}},w,y} \bar{\varphi}_{S,y} \, \mathrm{d}S \right).
\end{aligned}$$
(A.17)

We observe the parenthetical term on the RHS of (A.17) is a space-time heat operator acting on $\bar{\varphi}$ evaluated at time $\mathfrak{t}_{1,\mathrm{st}}$. Thus the far RHS of (A.17) is the *spatial* heat operator acting on this time-$\mathfrak{t}_{1,\mathrm{st}}$ space-time heat operator. Because the spatial heat operator is a contraction with respect to the supremum-norm on functions on $\mathbb{T}_N$, we get from (A.17) the bound

$$\|\mathbf{1}_{T \geqslant \mathfrak{t}_{1,\mathrm{st}}} \bar{\mathfrak{P}}^N_{T,x}(\bar{\varphi}_{S,y}\mathbf{1}_{S \leqslant \mathfrak{t}_{1,\mathrm{st}}})\|_{1;\mathbb{T}_N} \;\leqslant\; \|\bar{\mathfrak{P}}^N_{T,x}(\bar{\varphi}_{S,y})\|_{\mathfrak{t}_{1,\mathrm{st}};\mathbb{T}_N}. \tag{A.18}$$

Combining the previous two bullet points gives the required bound for all $T \leqslant 1$ so we are done. □

## Appendix B. Martingale Inequalities

We provide here a generalization of the martingale estimate in Lemma 3.1 in [10]; the proof of the result in [10] is specific to the microscopic Cole-Hopf transform while for this paper we will need it for other processes as well. Roughly speaking, Lemma 3.1 in [10] requires a short-time bound for the microscopic Cole-Hopf transform that holds at the level of $\|\,\|_{\omega;2p}$-norms; results like Lemma 6.6, which holds for processes we are interested in, are for very a different norm and thus not engineered for this.

**Lemma B.1.** *Retain the setting of* Lemma 3.1 *in [10] and consider the fundamental solutions* $\mathfrak{J}^N, \bar{\mathfrak{J}}^N$ *defined prior to the statement of* Proposition *4.11. Recall* Definition *4.10 and consider the following from* Lemma 3.1 *in [10] in which* $0 \leqslant \mathfrak{t}_1 \leqslant \mathfrak{t}_2$ *are fixed:*

- *We define the following local quadratic function of* $\varphi : \mathbb{R}_{\geqslant 0} \times \mathbb{Z}^2 \to \mathbb{R}$ *where* $\mathfrak{m}$ *is the uniformly bounded maximal jump-length:*

$$\widetilde{\varphi}^{\mathfrak{t}_1,\mathfrak{t}_2}_{R,x,w} \;\overset{\bullet}{=}\; \sup_{\mathfrak{r}' \in [\mathfrak{t}_1,\mathfrak{t}_2):\ \lfloor \mathfrak{r}' \rfloor = \lfloor R \rfloor} \sup_{|\mathfrak{j}| \leqslant \mathfrak{m}} |\varphi_{\mathfrak{r}',x,w+\mathfrak{j}} \varphi_{\mathfrak{r}',x,w}|. \tag{B.1}$$

*Provided any* $p \in \mathbb{R}_{\geqslant 1}$ *and* $\mathfrak{t}_1 \leqslant \mathfrak{t}_2$ *and* $\varphi : \mathbb{R}_{\geqslant 0} \times \mathbb{Z}^2 \to \mathbb{R}$, *all deterministic, we have, again recalling notation in* Definition *4.10,*

$$\left\| \int_{\mathfrak{t}_1}^{\mathfrak{t}_2} \sum_{w \in \mathbb{Z}} \varphi_{R,x,w} \cdot \mathfrak{J}^N_{S,T+R,w,y} \mathrm{d}\xi^N_{R,w} \right\|^2_{\omega;2p}$$

$$\lesssim_{p,\mathfrak{m}} \int_{\mathfrak{t}_1}^{\mathfrak{t}_2} N \left( \sup_{\mathfrak{t}_1 \leqslant \mathfrak{r} \leqslant \mathfrak{t}_2} \sup_{w \in \mathbb{Z}} e^2_{N,w,y} \|\mathfrak{J}^N_{S,T+\mathfrak{r},w,y}\|^2_{\omega;2p} \right) \sum_{w \in \mathbb{Z}} \widetilde{\varphi}^{\mathfrak{t}_1,\mathfrak{t}_2}_{R,x,w} e^{-2}_{N,w,y} \, \mathrm{d}R. \tag{B.2}$$

*We have the same bound after replacing $\mathfrak{I}^N$ by $\bar{\mathfrak{I}}^N$, $\mathbb{Z}$ by $\mathbb{T}_N$, $e_{N,w,y}$ by $e_{N,w,y}^{\mathrm{cpt}}$, $\varphi$ by $\bar{\varphi} : \mathbb{R}_{\geqslant 0} \times \mathbb{T}_N^2 \to \mathbb{R}$, and replacing $\widetilde{\varphi}^{\mathfrak{t}_1, \mathfrak{t}_2}$ with*

$$\widetilde{\varphi}_{R,x,w}^{\mathfrak{t}_1, \mathfrak{t}_2} \;\stackrel{\bullet}{=}\; \sup\nolimits_{\mathfrak{r}' \in [\mathfrak{t}_1, \mathfrak{t}_2): \; \lfloor \mathfrak{r}' \rfloor = \lfloor R \rfloor} \sup\nolimits_{|\mathfrak{j}| \leqslant \mathfrak{m}} \mathbf{1}_{w + \mathfrak{j} \in \mathbb{T}_N} |\bar{\varphi}_{\mathfrak{r}', x, w + \mathfrak{j}} \bar{\varphi}_{\mathfrak{r}', x, w}|. \tag{B.3}$$

*Recall $\mathfrak{Y}^N$ from the beginning of* Sect. 8 *and recall $\mathfrak{D}^N$ in* Definition 8.7. *We also have the following estimates for $\mathfrak{X}^N = \mathfrak{Y}^N, \mathfrak{D}^N$ and for any uniformly bounded adapted process $\mathfrak{R}^N : \mathbb{R}_{\geqslant 0} \times \mathbb{T}_N \to \mathbb{R}$:*

$$\|\int_{\mathfrak{t}_1}^{\mathfrak{t}_2} \sum_{w \in \mathbb{T}_N} \bar{\varphi}_{R,x,w} \cdot \mathfrak{X}_{R,w}^N \mathrm{d}\xi_{R,w}^N\|_{\omega;2p}^2 \;\lesssim_{p,\mathfrak{m}}\; \int_{\mathfrak{t}_1}^{\mathfrak{t}_2} N \left( \sup_{w \in \mathbb{T}_N} \|\mathfrak{X}_{\lfloor R \rfloor, w}^N\|_{\omega;2p}^2 \right) \sum_{w \in \mathbb{T}_N} \widetilde{\varphi}_{R,x,w}^{\mathfrak{t}_1, \mathfrak{t}_2} \, \mathrm{d}R \tag{B.4a}$$

$$\|\int_{\mathfrak{t}_1}^{\mathfrak{t}_2} \sum_{w \in \mathbb{T}_N} \bar{\varphi}_{R,x,w} \cdot \mathfrak{R}_{R,w}^N \mathrm{d}\xi_{R,w}^N\|_{\omega;2p}^2 \;\lesssim_{p,\mathfrak{m}}\; \int_{\mathfrak{t}_1}^{\mathfrak{t}_2} N \sum_{w \in \mathbb{T}_N} \widetilde{\varphi}_{R,x,w}^{\mathfrak{t}_1, \mathfrak{t}_2} \, \mathrm{d}R. \tag{B.4b}$$

We first define the kernel for the total mass of a branching random walk to study short-time behavior of processes in Lemma B.1. We will first recall the discrete approximations of Laplacians given by $\mathscr{L}^{!!}$ in Definition 2.6 and $\bar{\mathscr{L}}^{!!}$ in Definition 4.1.

**Definition B.2.** Define the kernel $\mathfrak{L}^N$ as the function on $\mathbb{R}_{\geqslant 0}^2 \times \mathbb{Z}^2$ satisfying $\mathfrak{L}_{S,S,x,y}^N = \mathbf{1}_{x=y}$ and the *deterministic* equation

$$\partial_T \mathfrak{L}_{S,T,x,y}^N \;=\; \mathscr{L}^{!!} \mathfrak{L}_{S,T,x,y}^N + N^{\frac{3}{2}} \sum_{|k| \leqslant N^{\beta_X}} \mathfrak{L}_{S,T,x+k,y}^N. \tag{B.5}$$

We define $\bar{\mathfrak{L}}^N$ to be the solution to the same equation after replacing $\mathfrak{L}^N$ by $\bar{\mathfrak{L}}^N$ and $\mathscr{L}^{!!}$ by $\bar{\mathscr{L}}^{!!}$ and $\mathbb{Z}$ by $\mathbb{T}_N$. We will additionally define the maximal processes $\mathfrak{L}_{S,T,x,y}^{N,\max} \stackrel{\bullet}{=} \sup_{0 \leqslant \mathfrak{r} \leqslant N^{-2}} \mathfrak{L}_{S,T+\mathfrak{r},x,y}^N$ and $\bar{\mathfrak{L}}_{S,T,x,y}^{N,\max} \stackrel{\bullet}{=} \sup_{0 \leqslant \mathfrak{r} \leqslant N^{-2}} \bar{\mathfrak{L}}_{S,T+\mathfrak{r},x,y}^N$.

The following auxiliary estimate for the two maximal kernels defined immediately above will help control their short-time behavior and, as we later explain in the proof of Lemma B.1, the short-time behavior of the processes of interest in the statement of Lemma B.1 too. The proof of the following is an elementary deterministic idea for linear ODEs; it is a substitute for short-time estimates on the microscopic Cole-Hopf transform which are key to the martingale estimate in Lemma 3.1 in [10] but which are inaccessible without a formula in terms of the particle system (there is such formula for the microscopic Cole-Hopf transform).

**Lemma B.3.** *Recall* Definition 4.10 *and $\beta_X = \frac{1}{3} + \varepsilon_{X,1}$. Define $\widetilde{e}_{N,x,y} = \mathbf{Exp}(N^{-\beta_X} \mathbf{d}_{x,y})$ and $\widetilde{e}_{N,x,y}^{\mathrm{cpt}} = \mathbf{Exp}(N^{-\beta_X} \mathbf{d}_{x,y}^{\mathrm{cpt}})$. We have*

$$\sum\nolimits_{y \in \mathbb{Z}} \mathfrak{L}_{0,0,x,y}^{N,\max} e_{N,x,y}^2 + \sum\nolimits_{y \in \mathbb{T}_N} \bar{\mathfrak{L}}_{0,0,x,y}^{N,\max} (e_{N,x,y}^{\mathrm{cpt}})^2$$
$$\leqslant \sum\nolimits_{y \in \mathbb{Z}} \mathfrak{L}_{0,0,x,y}^{N,\max} \widetilde{e}_{N,x,y}^2 + \sum\nolimits_{y \in \mathbb{T}_N} \bar{\mathfrak{L}}_{0,0,x,y}^{N,\max} (\widetilde{e}_{N,x,y}^{\mathrm{cpt}})^2 \;\lesssim_{\mathfrak{m}}\; 1. \tag{B.6}$$

*Proof.* The first bound follows via $e_{N,x,y} \leqslant \widetilde{e}_{N,x,y}$, and the same for cpt-versions, as $N^{-3/4} \leqslant N^{-\beta_X}$; see Definition 4.10. We first prove the claimed bound for $\mathfrak{L}^N_{0,\mathfrak{t},x,y}$ for $\mathfrak{t} \leqslant N^{-2}$ instead of $\mathfrak{L}^{N,\max}$. We use the PDE for $\mathfrak{L}^N$ to get, with more explanation after,

$$
\begin{aligned}
\sum_{y \in \mathbb{Z}} |\partial_T \mathfrak{L}^N_{0,T,x,y}| \widetilde{e}^2_{N,x,y} \\
\lesssim N^2 \sum_{y \in \mathbb{Z}} \Big( \sum_{|k| \leqslant \mathfrak{m}} \mathfrak{L}^N_{0,T,x+k,y} \widetilde{e}^2_{N,x+k,y} + N^{-\frac{1}{2}} \sum_{|k| \leqslant N^{\beta_X}} \mathfrak{L}^N_{0,T,x+k,y} \widetilde{e}^2_{N,x+k,y} \Big) \\
\lesssim_{\mathfrak{m}} N^2 (1 + N^{-\frac{1}{2}+\beta_X}) \sup_{x \in \mathbb{Z}} \sum_{y \in \mathbb{Z}} \mathfrak{L}^N_{0,T,x,y} \widetilde{e}^2_{N,x,y} \\
\lesssim N^2 \sup_{x \in \mathbb{Z}} \sum_{y \in \mathbb{Z}} \mathfrak{L}^N_{0,T,x,y} \widetilde{e}^2_{N,x,y} \overset{\bullet}{=} N^2 \Phi_T.
\end{aligned}
\tag{B.7}
$$

We used the inequality $\widetilde{e}_{N,x,y} \lesssim \widetilde{e}_{N,x+k,y}$ for any $|k| \leqslant N^{\beta_X}$ as well as $N^{\beta_X} \leqslant N^{1/2}$. As the last display is uniform in $x \in \mathbb{Z}$, we get

$$
\begin{aligned}
\Phi_{\mathfrak{t}} &\lesssim \Phi_0 + \sup_{x \in \mathbb{Z}} \int_0^{\mathfrak{t}} \Big| \Big( \sum_{y \in \mathbb{Z}} \partial_T \mathfrak{L}^N_{0,T,x,y} \widetilde{e}^2_{N,x,y} \Big) \Big| \, \mathrm{d}T \\
&\leqslant \Phi_0 + \sup_{x \in \mathbb{Z}} \int_0^{\mathfrak{t}} \sum_{y \in \mathbb{Z}} |\partial_T \mathfrak{L}^N_{0,T,x,y}| \widetilde{e}^2_{N,x,y} \, \mathrm{d}T \\
&\lesssim_{\mathfrak{m}} \Phi_0 + N^2 \int_0^{\mathfrak{t}} \Phi_T \, \mathrm{d}T.
\end{aligned}
$$

Above we bounded each term in the sup defining $\Phi_{\mathfrak{t}}$ by initial data plus the integral of the absolute value of its time-derivative, and we took a sup over $x \in \mathbb{Z}$ of that bound to control $\Phi_{\mathfrak{t}}$ itself. If $\mathfrak{t} \leqslant N^{-2}$, the Gronwall inequality and the simple observation $\Phi_0 \lesssim 1$ (note $\mathfrak{L}^N_{0,0,x,y} = \mathbf{1}_{x=y}$) give the claim for $\mathfrak{L}^N_{0,\mathfrak{t},x,y}$. For $\mathfrak{L}^{N,\max}$, we have, with explanation after, a similar integral bound:

$$
\begin{aligned}
|\mathfrak{L}^{N,\max}_{0,0,x,y}| &\lesssim \mathbf{1}_{x=y} + \sup_{0 \leqslant \mathfrak{t} \leqslant N^{-2}} \Big( N^2 \int_0^{\mathfrak{t}} \sum_{|k| \leqslant \mathfrak{m}} |\mathfrak{L}^N_{0,\mathfrak{r},x+k,y}| \mathrm{d}\mathfrak{r} \\
&\quad + N^{\frac{3}{2}} \int_0^{\mathfrak{t}} \sum_{|k| \leqslant N^{\beta_X}} |\mathfrak{L}^N_{0,\mathfrak{r},x+k,y}| \mathrm{d}\mathfrak{r} \Big)
\end{aligned}
\tag{B.8}
$$

$$
\leqslant \mathbf{1}_{x=y} + N^2 \int_0^{N^{-2}} \sum_{|k| \leqslant \mathfrak{m}} |\mathfrak{L}^N_{0,\mathfrak{r},x+k,y}| \mathrm{d}\mathfrak{r} + N^{\frac{3}{2}} \int_0^{N^{-2}} \sum_{|k| \leqslant N^{\beta_X}} |\mathfrak{L}^N_{0,\mathfrak{r},x+k,y}| \mathrm{d}\mathfrak{r}.
\tag{B.9}
$$

The first bound follows by writing the $\mathfrak{L}^N$-kernel as initial data plus the integral of the absolute value of its time-derivative and then applying its PDE. The second line follows from noting the integrals on the RHS of the first line are of non-negative terms. Let us now recall $\widetilde{e}_{N,x,y} \lesssim \widetilde{e}_{N,x+k,y}$ given any $|k| \leqslant N^{\beta_X}$. Thus, we multiply (B.9) and the LHS of (B.8) by these $\widetilde{e}^2_{N,x,y}$-weights, sum over $y \in \mathbb{Z}$, apply the result for $\mathfrak{L}^N$ instead of $\mathfrak{L}^{N,\max}$ established in the above paragraph, and then integrate-in-time on the interval $[0, N^{-2}]$ to get the desired bound for $\mathfrak{L}^{N,\max}$. A similar argument works to prove the $\widetilde{\mathfrak{L}}^{N,\max}$-estimates. □

*Proof of Lemma B.1.* Let us first prove (B.2). As $\mathfrak{J}^N$ is adapted in the forwards time-variable, we can follow the proof of Lemma 3.1 in [10] starting with (3.5) except we replace the supremum over $|j| \leqslant \mathfrak{m}$ with a sum, and we estimate each term in this sum.

We now find a replacement for (3.7) in [10] for short-time control on $\mathfrak{J}^N$. Recall $\mathfrak{J}^N$ from before Proposition 4.11.

- The process $\mathfrak{J}^N$ solves a linear stochastic equation with a discrete-type Laplacian $\mathscr{L}^{!!}$, some gradient terms with coefficients that are at most order $N^{3/2}$, a random multiplicative potential of order $N^{3/2}$, and jumps of order $N^{-1/2}$. Note this equation is time-inhomogeneous. However, we may still use a Feynman-Kac representation for $\mathfrak{J}^N$ where the potential is the sum of the aforementioned order $N^{3/2}$-potential and order $N^{-1/2}$ jumps, while the underlying random walk is the one whose generator is $\mathscr{L}^{!!}$ plus gradient-operators in the $\mathfrak{J}^N$-equation. It is easy to see the transition probability of this random walk is bounded by the $\mathfrak{L}^N$-kernel, since gradient terms in the $\mathfrak{J}^N$-equation are controlled by non-$\mathscr{L}^{!!}$-operator terms in the $\mathfrak{L}^N$-equation.
- We make the bullet point above precise. First, we let $\mathfrak{y}^\star$ denote the random walk in the previous bullet point independent of the particle system and we denote expectation with respect to $\mathfrak{y}^\star$ by $\mathbb{E}$. We also define the accumulation of order $N^{-1/2}$-jumps and the order $N^{3/2}$-potential in the $\mathfrak{J}^N$-equation along a sample $\mathfrak{y}^\star$-path below in which $\kappa \in \mathbb{R}_{>0}$ is a universal constant:

$$\mathfrak{J}_{\lfloor R \rfloor, \mathfrak{t}} \overset{\bullet}{=} |\left( \int_0^{\mathfrak{t}} \mathrm{d}\xi^N_{T+\lfloor R \rfloor + \mathfrak{t} - \mathfrak{s}, \mathfrak{y}^\star_{T+\lfloor R \rfloor + \mathfrak{s}}} \right)| + \kappa N^{3/2} \mathfrak{t}. \tag{B.10}$$

The integral on the RHS above, which is with respect to $\mathfrak{s}$, is a sum of jumps plus the integral of the order $N^{3/2}$ compensation drift in $\mathrm{d}\xi^N$; see (2.4) in [10]. By the Feynman-Kac formula given by Proposition 7.1 in Appendix 1 of [21], which holds for time-inhomogeneous processes after straightforward replacements in its proof, for $\mathfrak{t} \leqslant N^{-2}$ we have, with explanation after,

$$(\mathfrak{J}^N_{S,T+\lfloor R \rfloor + \mathfrak{t}, w, y})^2 \lesssim \left( \mathbb{E}_{\mathfrak{y}^\star_{T+\lfloor R \rfloor} = w} \mathbf{Exp}\left( \mathfrak{J}_{\lfloor R \rfloor, \mathfrak{t}} \right) \mathfrak{J}^N_{S, T+\lfloor R \rfloor, \mathfrak{y}^\star_{T+\lfloor R \rfloor + \mathfrak{t}}, y} \right)^2 \tag{B.11}$$

$$\leqslant \left( \mathbb{E}_{\mathfrak{y}^\star_{T+\lfloor R \rfloor} = w} \mathbf{Exp}\left( 2\mathfrak{J}_{\lfloor R \rfloor, \mathfrak{t}} \right) \right) \left( \mathbb{E}_{\mathfrak{y}^\star_{T+\lfloor R \rfloor} = w} |\mathfrak{J}^N_{S, T+\lfloor R \rfloor, \mathfrak{y}^\star_{T+\lfloor R \rfloor + \mathfrak{t}}, y}|^2 \right)$$

$$\lesssim \left( \mathbb{E}_{\mathfrak{y}^\star_{T+\lfloor R \rfloor} = w} \mathbf{Exp}\left( 2\mathfrak{J}_{\lfloor R \rfloor, \mathfrak{t}} \right) \right) \sum_{z \in \mathbb{Z}} \mathfrak{L}^{N,\max}_{0,0,w,z} |\mathfrak{J}^N_{S, T+\lfloor R \rfloor, z, y}|^2. \tag{B.12}$$

The first estimate is the Feynman-Kac formula because $\mathfrak{J}_{\lfloor R \rfloor, \mathfrak{t}}$ bounds all jumps and potentials in the $\mathfrak{J}^N$-equation. The second bound follows by the Cauchy-Schwarz inequality. The third/final bound follows by recalling the transition probability of $\mathfrak{y}^\star$ is controlled by the $\mathfrak{L}^N$-kernel, as noted in the first bullet point, which is then controlled by the $\mathfrak{L}^{N,\max}$-kernel. This will be our replacement for (3.7) in [10] because $\mathfrak{L}^{N,\max}$ is deterministic so beyond the first exponential factor in (B.12) the randomness in (B.12) is $\lfloor R \rfloor$-measurable. We also have the following conditional expectation estimate, where $\mathfrak{n}_{\lfloor R \rfloor, \mathfrak{t}, w}$ denotes the number of times that any Poisson clock across $w \in \mathbb{Z}$ rings in the interval $T + \lfloor R \rfloor + [0, \mathfrak{t}]$ and $\mathbf{E}_{\lfloor R \rfloor}$ conditions on $\lfloor R \rfloor$-data. Here $\kappa \in \mathbb{R}_{>0}$ is, again, universal, and $\mathfrak{t} \leqslant N^{-2}$, and $p \in \mathbb{R}_{\geqslant 1}$. We give more explanation for each bound below after:

$$\mathbf{E}_{\lfloor R\rfloor}|\mathfrak{n}_{\lfloor R\rfloor,\mathfrak{t},w}|^{2p}\mathbb{E}_{\mathfrak{y}^{\star}_{T+\lfloor R\rfloor}=w}\mathbf{Exp}\left(2p\mathfrak{J}_{\lfloor R\rfloor,\mathfrak{t}}\right) \tag{B.13}$$

$$\lesssim \sum_{\mathfrak{k}\in\mathbb{Z}_{\geqslant 1}}e^{-\mathfrak{k}}\mathbf{E}_{\lfloor R\rfloor}\left(|\mathfrak{n}_{\lfloor R\rfloor,\mathfrak{t},w}|^{2p}\mathbf{Exp}\left(2p\kappa N^{-1/2}\sum_{z\in w+\llbracket -10N^{\beta_X}\mathfrak{k},\,10N^{\beta_X}\mathfrak{k}\rrbracket}\mathfrak{n}_{\lfloor R\rfloor,\mathfrak{t},z}\right)\right) \tag{B.14}$$

$$\lesssim_p \mathfrak{t}\sum_{\mathfrak{k}\in\mathbb{Z}_{\geqslant 1}}e^{-\frac{1}{2}\mathfrak{k}} \lesssim \mathfrak{t}. \tag{B.15}$$

First recall the $\mathfrak{y}^{\star}$-walk has transition probability bounded by $\mathfrak{L}^{N,\max}$. The first inequality follows by conditioning on $\mathfrak{y}^{\star}$ to hit $w+\llbracket -10N^{\beta_X}\mathfrak{k},\,10N^{\beta_X}\mathfrak{k}\rrbracket\setminus\llbracket -10N^{\beta_X}(\mathfrak{k}-1),\,10N^{\beta_X}(\mathfrak{k}-1)\rrbracket$, which happens with probability $\mathscr{O}(e^{-\mathfrak{k}})$ because of the exponential tail estimate on $\mathfrak{L}^{N,\max}$ in Lemma B.3, but stay in $w+\llbracket -10N^{\beta_X}\mathfrak{k},\,10N^{\beta_X}\mathfrak{k}\rrbracket$, and then bounding $\mathfrak{J}_{\lfloor R\rfloor,\mathfrak{t}}$ for this $\mathfrak{y}^{\star}$-walk-trajectory by the total number of ringings $\mathfrak{n}_{\lfloor R\rfloor,\mathfrak{t},z}$ across all points $z\in w+\llbracket -10N^{\beta_X}\mathfrak{k},\,10N^{\beta_X}\mathfrak{k}\rrbracket$. The second bound follows by elementary Poisson calculations as in (3.8) in [10]. We note the exponential in $\mathbf{E}_{\lfloor R\rfloor}$, per-sum-index $\mathfrak{k}\in\mathbb{Z}_{\geqslant 1}$, is the exponential of a Poisson variable, with speed of order $N^{\beta_X}\mathfrak{k}$, scaled by order $N^{-1/2}$. The Laplace transform of such a Poisson variable is exponential in $\mathscr{O}_p(N^{-1/2+\beta_X}\mathfrak{k})$, which is then dwarfed by the $e^{-\mathfrak{k}}$-factors in the $\mathfrak{k}\in\mathbb{Z}_{\geqslant 1}$-sum. The last bound is a simple geometric series.

We proceed after (3.5) in [10] but having replaced the supremum over $|\mathfrak{j}|\leqslant\mathfrak{m}$ with a sum and skipping (3.6a) in [10]. Namely, by (B.12), the conditional expectation estimate (B.15), and the $\lfloor R\rfloor$-measurable nature of the second factor in (B.12), we show the following parallel to the bound after (3.8) in [10]; the $N$-factor comes by change-of-variables to account for the $N^2$-speed:

$$\|\int_{\mathfrak{t}_1}^{\mathfrak{t}_2}\sum_{w\in\mathbb{Z}}\varphi_{R,x,w}\cdot\mathfrak{J}^N_{S,T+R,w,y}\mathrm{d}\xi^N_{R,w}\|^2_{\omega;2p}$$

$$\lesssim_{p,\mathfrak{m}}\sum_{|\mathfrak{j}|\leqslant\mathfrak{m}}\int_{\mathfrak{t}_1}^{\mathfrak{t}_2}N\sum_{w\in\mathbb{Z}}\widetilde{\varphi}^{\mathfrak{t}_1,\mathfrak{t}_2}_{R,x,w}\cdot\|\sum_{z\in\mathbb{Z}}\mathfrak{L}^{N,\max}_{0,0,w+\mathfrak{j},z}|\mathfrak{J}^N_{S,T+\lfloor R\rfloor,z,y}|^2\|_{\omega;p}\,\mathrm{d}R \tag{B.16}$$

$$\lesssim_{p,\mathfrak{m}}\sum_{|\mathfrak{j}|\leqslant\mathfrak{m}}\int_{\mathfrak{t}_1}^{\mathfrak{t}_2}N\sum_{w\in\mathbb{Z}}\widetilde{\varphi}^{\mathfrak{t}_1,\mathfrak{t}_2}_{R,x,w}\cdot\sum_{z\in\mathbb{Z}}\mathfrak{L}^{N,\max}_{0,0,w+\mathfrak{j},z}\|\mathfrak{J}^N_{S,T+\lfloor R\rfloor,z,y}\|^2_{\omega;2p}\,\mathrm{d}R \tag{B.17}$$

$$\lesssim\sum_{|\mathfrak{j}|\leqslant\mathfrak{m}}\int_{\mathfrak{t}_1}^{\mathfrak{t}_2}N\sum_{w\in\mathbb{Z}}\widetilde{\varphi}^{\mathfrak{t}_1,\mathfrak{t}_2}_{R,x,w}e^{-2}_{N,w,y}\cdot e^2_{N,w,y}\sum_{z\in\mathbb{Z}}\mathfrak{L}^{N,\max}_{0,0,w+\mathfrak{j},z}\|\mathfrak{J}^N_{S,T+\lfloor R\rfloor,z,y}\|^2_{\omega;2p}\,\mathrm{d}R. \tag{B.18}$$

The bound (B.17) follows from the previous line combined with the triangle inequality for the $\|\|_{\omega;2p}$-norm; for this step recall the $\mathfrak{L}^{N,\max}$-kernel is deterministic. The final bound (B.18) follows from multiplying by $e^{-2}_{N,w,y}$ and $e^2_{N,w,y}$. At this point, it suffices to use the following in which we observe $e^2_{N,w,y}\lesssim e^2_{N,w,z}e^2_{N,z,y}\lesssim e^2_{N,w+\mathfrak{j},z}e^2_{N,z,y}$ for any $|\mathfrak{j}|\leqslant\mathfrak{m}$ which follows by definition of these exponential weights in Definition 4.10, the bound $\mathfrak{m}\lesssim 1$, and elementary calculations with the triangle inequality. We deduce

$$e^2_{N,w,y}\sum_{z\in\mathbb{Z}}\mathfrak{L}^{N,\max}_{0,0,w+\mathfrak{j},z}\cdot\|\mathfrak{J}^N_{S,T+\lfloor R\rfloor,z,y}\|^2_{\omega;2p}$$

$$\lesssim\sum_{z\in\mathbb{Z}}\mathfrak{L}^{N,\max}_{0,0,w+\mathfrak{j},z}e^2_{N,w+\mathfrak{j},z}\cdot e^2_{N,z,y}\|\mathfrak{J}^N_{S,T+\lfloor R\rfloor,z,y}\|^2_{\omega;2p}$$

$$\lesssim \sup_{\substack{\mathfrak{t}_1\leqslant\mathfrak{r}\leqslant\mathfrak{t}_2\\z\in\mathbb{Z}}}e^2_{N,z,y}\|\mathfrak{J}^N_{S,T+\mathfrak{r},z,y}\|^2_{\omega;2p}. \tag{B.19}$$

The second estimate in the last display follows via pulling out the space-time supremum of the $\|\cdot\|^2$-factor in the $\mathbb{Z}$-summation, and then estimating the remaining $\mathbb{Z}$-summation above using Lemma B.3. If we plug the above display into (B.18) and multiply by 3$\mathfrak{m}$ to bound the sum over $|\mathfrak{j}| \leqslant \mathfrak{m}$, then we get (B.2). To establish the estimate (B.2) but with compactification replacements, the same argument works if we replace $\mathbb{Z}$ by $\mathbb{T}_N$, $\mathscr{L}^{!!}$ by $\bar{\mathscr{L}}^{!!}$, gradients on $\mathbb{Z}$ by those on $\mathbb{T}_N$, and $\mathfrak{L}^{N,\max}$ by $\bar{\mathfrak{L}}^{N,\max}$.

To establish (B.4a), observe that $\mathfrak{X}^N$ solves a similar stochastic equation as $\tilde{\mathfrak{J}}^N$, so that (B.17) holds upon replacing $\mathfrak{J}^N$ by $\mathfrak{X}^N$ including the necessary replacements $\widetilde{\varphi}$ by $\widetilde{\widetilde{\varphi}}$ and $\mathbb{Z}$ by $\mathbb{T}_N$ and $\mathfrak{L}^{N,\max}$ by $\bar{\mathfrak{L}}^{N,\max}$. At that point we pull out the supremum over $z \in \mathbb{Z}$ of $\|\mathfrak{X}^N_{\lfloor R \rfloor, z}\|^2_{\omega; 2p}$ and estimate the 1-norm of the $\bar{\mathfrak{L}}^{N,\max}$-kernel with Lemma B.3. This provides (B.4a) after again multiplying the resulting estimate by 3$\mathfrak{m}$. To get (B.4b), we actually instead follow the proof for Lemma 3.1 in [10], but with a replacement of $\mathbb{Z}$ by $\mathbb{T}_N$ until (3.5) therein, so we have (3.5) therein but replacing the $Z$-process with $\mathfrak{R}$. At that point, we drop the $\mathfrak{R}$-factors because they are uniformly bounded and proceed like in the rest of the proof for Lemma 3.1 in [10] starting at (3.8), though for our version of (3.8) here we can do without the conditioning. This proves (B.4b) and therefore finishes the proof. □

## Appendix C. Notation

- "Arbitrarily small/large but universal" constants will be taken as small/large as we want but independent of $N \in \mathbb{Z}_{>0}$. Also, we will always take $N \in \mathbb{Z}_{>0}$ sufficiently large though depending only on some choices of universal constants just to make certain statements true, such as $100 \leqslant N^\varepsilon$ where $\varepsilon \in \mathbb{R}_{>0}$ is universal.
- We always let $\mathfrak{m} \in \mathbb{Z}_{>0} \cup \{+\infty\}$ denote the maximal jump-length in the particle system.
- Provided any $a, b \in \mathbb{R}$, we will define the discretized interval $[\![a, b]\!] \overset{\bullet}{=} [a, b] \cap \mathbb{Z}$.
- We use big-Oh notation $\mathscr{O}$, and $\lesssim$ is synonymous with $\mathscr{O}$. Dependences of implied constants in $\lesssim$ and $\mathscr{O}$ will be recorded as subscripts. For example $\lesssim_I$ means the implied constant depends only on a set $I$. We also use the pre-factor $\kappa_I$ synonymously with $\lesssim_I$ or $\mathscr{O}_I$ so that the constants $\kappa_I$ and $\kappa_{i_1,\ldots,i_\mathfrak{n}}$ depend only on the set $I = \{i_1, \ldots, i_\mathfrak{n}\}$.
- For any $X \in \mathbb{Z}$ define the shift-map $\tau_X : \Omega_{\mathbb{Z}} \to \Omega_{\mathbb{Z}}$ via $(\tau_X \eta)_Z = \eta_{Z+X}$. For any $\varphi : \Omega_{\mathbb{Z}} \to \mathbb{R}$ define $\tau_X \varphi = \varphi \circ \tau_X : \Omega_{\mathbb{Z}} \to \mathbb{R}$.
- For any $\varphi : \Omega_{\mathbb{Z}} \to \mathbb{R}$, its support is defined as the smallest subset $\mathfrak{I} \subseteq \mathbb{Z}$ such that $\varphi$ depends only on spins $\eta_x$ for $x \in \mathfrak{I}$.
- Provided any *finite set* $I$, we define the averaged sum $\widetilde{\sum}_{i \in I} = |I|^{-1} \sum_{i \in I}$ where $|I|$ denotes the cardinality/size of $I$.
- For any operator $\mathscr{T}$, each superscript ! denotes scaling by $N$. In particular, define the discrete gradient $\nabla_k \varphi_x = \varphi_{x+k} - \varphi_x$ for any $k \in \mathbb{Z}$ and the discrete Laplacian $\Delta_k \varphi_x = \varphi_{x+k} + \varphi_{x-k} - 2\varphi_x$ for any $k \in \mathbb{Z}$. We then define $\nabla_k^! = N\nabla_k$ and $\Delta_k^{!!} = N^2 \Delta_k$.
- Define the *torus* $\mathbb{T}_N \overset{\bullet}{=} [\![-N^{5/4+\varepsilon_{\mathrm{cpt}}}, N^{5/4+\varepsilon_{\mathrm{cpt}}}]\!] \subseteq \mathbb{Z}$ with $\varepsilon_{\mathrm{cpt}} \in \mathbb{R}_{>0}$ arbitrarily small but universal.
- For $k \in \mathbb{Z}$ define $\bar{\nabla}_k \varphi_x = \varphi_{x \widetilde{+} k} - \varphi_x$ and $\bar{\Delta}_k \varphi_x = \varphi_{x \widetilde{+} k} + \varphi_{x \widetilde{-} k} - 2\varphi_x$. Here $\widetilde{+}/\widetilde{-}$ denote the induced addition/subtraction on the *torus* $\mathbb{T}_N$. This addition/subtraction notation is specific to this bullet point. Define $\bar{\nabla}_k^! = N\bar{\nabla}_k$ and $\bar{\Delta}_k^{!!} = N^2 \bar{\Delta}_k$.
- The macroscopically smooth cutoff function $\chi$ is catered to the torus $\mathbb{T}_N$; it is defined at the beginning of Sect. 4.

- For any $\mathfrak{t} \in \mathbb{R}_{\geqslant 0}$ and any subset $\mathfrak{X} \subseteq \mathbb{R}$ we define a discretization $\mathfrak{I}_{\mathfrak{t}} \overset{\bullet}{=} \{\mathfrak{t} \cdot \mathfrak{j} N^{-100}\}_{\mathfrak{j}=0}^{N^{100}}$ and the norms/norm-type quantities

$$\|\varphi\|_{\mathfrak{t};\mathfrak{X}} \overset{\bullet}{=} \sup_{0 \leqslant \mathfrak{s} \leqslant \mathfrak{t}} \sup_{x \in \mathfrak{X}} |\varphi_{\mathfrak{s},x}|$$

$$[\varphi]_{\mathfrak{t};\mathfrak{X}} \overset{\bullet}{=} \sup_{\mathfrak{s} \in \mathfrak{I}_{\mathfrak{t}}} \sup_{x \in \mathfrak{X}} |\varphi_{\mathfrak{s},x}|$$

$$\langle\varphi\rangle_{\mathfrak{t};\mathfrak{X}} \overset{\bullet}{=} 1 + \|\varphi\|_{\mathfrak{t};\mathfrak{X}}^{2}.$$

- Define $\mathbf{d}_{x,y} = |x - y|$ as the usual absolute value distance for $x, y \in \mathbb{Z}$ and $\mathbf{d}_{x,y}^{\mathrm{cpt}}$ as geodesic distance on the torus $\mathbb{T}_{N}$.
- For any $p \in \mathbb{R}_{\geqslant 1}$, define $\|\varphi\|_{\omega;p}$ as the $p$-norm *over all randomness in the exclusion process* including the initial measure.
- Provided any $S, T \in \mathbb{R}_{\geqslant 0}$, we define $\mathfrak{s}_{S,T} \overset{\bullet}{=} |T - S|$ usually as an integrable singularity in a time-integral.

# References

1. Amir, G., Corwin, I., Quastel, J.: Probability distribution of the free energy of the continuum directed polymer model in $(1 + 1)$-dimensions. Commun. Pure Appl. Math. **64**, 466–537 (2011)
2. Bertini, L., Giacomin, G.: Stochastic Burgers and KPZ equations from particle systems. Commun. Math. Phys. **183**(3), 571–606 (1997)
3. Billingsley, P.: Convergence of Probability Measures. New York: John Wiley & Sons, Inc
4. Brox, T., Rost, H.: Equilibrium fluctuations of stochastic particle systems: the role of conserved quantities. Ann. Probab. **12**(3), 742–759 (1984)
5. Chang, C.C., Yau, H.T.: Fluctuations of one-dimensional Ginzburg–Landau models in nonequilibrium. Commun. Math. Phys. **145**, 209 (1992)
6. Corwin, I., Ghosal, P., Shen, H., Tsai, L.-C.: Stochastic PDE limit of the six vertex model. Commun. Math. Phys. **375**, 1945–2038 (2020)
7. Corwin, I.: The Kardar–Parisi–Zhang equation and universality class. Random Matrices Theory Appl. **01**(01), 1130001 (2012)
8. Corwin, I., Shen, H., Tsai, L.-C.: ASEP(q, j) converges to the KPZ equation. Ann. Inst. H. Poincare **54**, 995–1012 (2018)
9. Corwin, I., Tsai, L.-C.: KPZ equation limit of higher spin exclusion processes. Ann. Probab. **45**(3), 1771–1798 (2017)
10. Dembo, A., Tsai, L.-C.: Weakly asymmetric non-simple exclusion process and the KPZ equation. Commun. Math. Phys. **341**(1), 219–261 (2016)
11. Erhard, D., Hairer, M.: "A scaling limit of the parabolic Anderson model with exclusion interaction". (2021) arXiv:2103.13479
12. Goncalves, P., Jara, M.: Nonlinear fluctuations of weakly asymmetric interacting particle systems. Arch. Ration. Mech. Anal. **212**, 597 (2014)
13. Goncalves, P., Jara, M.: Stochastic Burgers equation from long range exclusion interactions. Stoch. Process. Appl. **127**(12), 4029–4052 (2017)
14. Guo, M.Z., Papnicolaou, G.C., Varadhan, S.R.S.: Nonlinear diffusion limit for a system with nearest neighbor interactions. Commun. Math. Phys. **118**, 31 (1988)
15. Hairer, M.: Solving the KPZ equation. Ann. Math. **178**(2), 559–664 (2013)
16. Hairer, M.: A theory of regularity structures. Invent. Math. **198**(2), 269–504 (2014)
17. Hairer, M., Shen, H.: A central limit theorem for the KPZ equation. Ann. Probab. **45**(6N), 4167–4221 (2017)
18. Hairer, M., Quastel, J.: A class of growth models rescaling to KPZ. Forum Math. Pi **6**, E3 (2018). https://doi.org/10.1017/fmp.2018.2
19. Kardar, M., Parisi, G., Zhang, Y.-C.: Dynamic scaling of growing interfaces. Phys. Rev. Lett. **56**, 9, 889 (1986)
20. Komorowski, T., Landim, C., Olla, S.: Fluctuations of Markov Processes, "Time Symmetry and Martingale Approximation". Springer, Berlin (2012)

21. Kipnis, C., Landim, C.: Scaling Limits of Interacting Particle Systems, Springer-Verlig Berlin Heidelberg, Volume 320, Volume 346, 2019. Pages 1137–1332 (1999)
22. Landim, C., Mourragui, M.: Hydrodynamic limit of mean zero asymmetric zero range processes infinite volume. Ann. Inst. H. Poincare 65–82 (1997)
23. Matetski, K.: "Martingale-driven approximations of singular stochastic PDEs" (2018). arXiv:1808.09429
24. Mueller, C.: On the support of solutions to the heat equation with noise. Stoch. Rep. (1991)
25. Yau, H.T.: Logarithmic Sobolev inequality for generalized simple exclusion processes. Probab. Theory Relat. Fields **109**, 507 (1997)