Local KPZ behavior under arbitrary scaling limits

Sourav Chatterjee
The KPZ equation

The Kardar–Parisi–Zhang (KPZ) equation models the growth of a generic randomly growing surface.

If \( f(t, x) \) is the height of a \( d \)-dimensional surface at time \( t \in \mathbb{R}_{\geq 0} \) and location \( x \in \mathbb{R}^d \), the KPZ equation says

\[
\partial_t f = \nu \Delta f + \frac{\lambda}{2} |\nabla f|^2 + \sqrt{D} \xi,
\]

where \( \xi \) is a random field known as space-time white noise, and \( \nu, \lambda \) and \( D \) are three strictly positive parameters.

Formally, space-time white noise is a distribution-valued centered Gaussian random field, with covariance

\[
\mathbb{E}(\xi(t, x)\xi(t', x')) = \delta(t - t')\delta^{(d)}(x - x'),
\]

where \( \delta \) and \( \delta^{(d)} \) are the Dirac delta functions on \( \mathbb{R} \) and \( \mathbb{R}^d \), respectively. (Can be rigorously defined.)
The KPZ equation was introduced by Kardar, Parisi, and Zhang in 1986 as a ‘universal’ model for random surface growth, in the same way that Brownian motion is a universal model for many 1D random processes.

The study of the KPZ equation and related topics is one of the most active areas in probability today.

The literature is huge. Will skip over this, due to time constraint. (See in the paper.)

Let me only briefly mention that KPZ is now well-understood in 1D, but remains mysterious in $d \geq 2$. 
The problem of scaling limits

- A fundamental roadblock in constructing nontrivial solutions of the KPZ equation in \( d \geq 2 \) is that *we do not know how to take scaling limits* of approximate solutions to reach a nontrivial limit.

- Even in 1D, there can be many different scaling limits.

- But in many 1D models, we know at least one way of taking a scaling limit that leads to a nontrivial solution.

- In higher dimensions, the question becomes less tractable.

- Physicists believe that for 2D models, the celebrated **Family–Vicsek scaling** is the correct one, and leads to a *function-valued*, rather than *distribution-valued*, solution of the 2D KPZ equation. (Distribution-valued solutions have been rigorously constructed in \( d \geq 2 \) in recent years.)

- This has been verified in numerical simulations, but remains out of the reach of rigorous mathematics.
I will present a small step towards understanding KPZ in $d \geq 2$ without confronting the issue of constructing scaling limits.

The approach is based on a framework introduced recently in Chatterjee '21a '21b and Chatterjee & Souganidis '21.

Since the ‘correct’ way to scale is still mysterious, the following workaround is proposed.

Consider a general class of growth models, which contains at least one model of widespread interest.

Then show that, irrespective of how we take a scaling limit, the growth is always locally like the KPZ equation, breaking up as the sum of a Laplacian term, a gradient squared term, a noise term, and a residual term that is negligible compared to the other three terms and their sum.

Surprisingly, this turns out to be doable. The details are in the following slides.
Defining local KPZ growth: Step 1

- Take any $d \geq 1$. Suppose that we have a collection of random functions $\{f_\varepsilon\}_{\varepsilon>0}$ from $\mathbb{Z}_{\geq 0} \times \mathbb{Z}^d$ into $\mathbb{R}$.

- A general ‘rescaling’ of $f_\varepsilon$ is defined as follows.

- Let $\alpha(\varepsilon)$, $\beta(\varepsilon)$ and $\gamma(\varepsilon)$ be positive real numbers depending on $\varepsilon$, such that $\alpha(\varepsilon)$ and $\beta(\varepsilon)$ tend to zero as $\varepsilon \to 0$.

- Based on these coefficients, the rescaled version of $f_\varepsilon$ is the function $f^{(\varepsilon)} : \mathbb{R}_{>0} \times \mathbb{R}^d \to \mathbb{R}$ defined as

$$f^{(\varepsilon)}(t, x) := \gamma(\varepsilon)f_\varepsilon(\lceil \alpha(\varepsilon)^{-1}t \rceil, \lceil \beta(\varepsilon)^{-1}x \rceil).$$

- This means we are rescaling space and time so that successive time points are separated by $\alpha(\varepsilon)$ and neighboring points in space are separated by $\beta(\varepsilon)$.

- The factor $\gamma(\varepsilon)$ is just a multiplicative factor meant to ensure that the limit of $f^{(\varepsilon)}$ as $\varepsilon \to 0$ (on some appropriate space of functions or distributions) is nontrivial.
Defining local KPZ growth: Step 2

- Let \( A = \{0, \pm e_1, \ldots, \pm e_d\} \) be the set consisting of the origin and its nearest neighbors in \( \mathbb{Z}^d \).
- Define the ‘local average’ of \( f^{(\varepsilon)} \) at \((t, x) \in \mathbb{R}_{>0} \times \mathbb{R}^d\) as

\[
\overline{f}^{(\varepsilon)}(t, x) := \frac{1}{2d + 1} \sum_{a \in A} f^{(\varepsilon)}(t, x + \beta(\varepsilon) a).
\]

- Define the ‘approximate time derivative’ of \( f^{(\varepsilon)} \) as

\[
\tilde{\partial}_t f^{(\varepsilon)}(t, x) := \frac{f^{(\varepsilon)}(t + \alpha(\varepsilon), x) - f^{(\varepsilon)}(t, x)}{\alpha(\varepsilon)},
\]

- Define the ‘approximate Laplacian’ as

\[
\tilde{\Delta} f^{(\varepsilon)}(t, x) := (2d + 1) \left( \frac{\overline{f}^{(\varepsilon)}(t, x) - f^{(\varepsilon)}(t, x)}{\beta(\varepsilon)^2} \right).
\]
Define the ‘approximate squared gradient’ as

\[
|\tilde{\nabla} f^{(\varepsilon)}(t, x)|^2 := \frac{1}{2} \sum_{a \in A} \left( \frac{f^{(\varepsilon)}(t, x + \beta^{(\varepsilon)}a) - \bar{f}^{(\varepsilon)}(t, x)}{\beta^{(\varepsilon)}} \right)^2.
\]

The above definitions are inspired by the fact that if \( \alpha^{(\varepsilon)} \to 0, \beta^{(\varepsilon)} \to 0, \) and \( f^{(\varepsilon)} \) converges in some strong sense to a smooth function \( f \) as \( \varepsilon \to 0, \) then the approximate time derivative, the approximate Laplacian, and the approximate squared gradient converge to \( \partial_t f, \Delta f \) and \( |\nabla f|^2. \)

Of course, we do not expect \( f^{(\varepsilon)} \) to converge to a smooth limit in general.
Definition of local KPZ growth

Definition

We will say that \( f(\varepsilon) \) has ‘local KPZ behavior’ as \( \varepsilon \to 0 \) if for some strictly positive \( \nu(\varepsilon) \), \( \lambda(\varepsilon) \), \( D(\varepsilon) \), and some random maps \( \xi(\varepsilon) \) and \( R(\varepsilon) \) from \( \mathbb{R}_{>0} \times \mathbb{R}^d \) into \( \mathbb{R} \), we have that for all \( (t, x) \in \mathbb{R}_{>0} \times \mathbb{R}^d \),

\[
\tilde{\partial}_t f^{(\varepsilon)}(t, x) = \nu(\varepsilon)\tilde{\Delta}f^{(\varepsilon)}(t, x) + \frac{\lambda(\varepsilon)}{2}|\tilde{\nabla}f^{(\varepsilon)}(t, x)|^2 + \sqrt{D(\varepsilon)}\xi^{(\varepsilon)}(t, x) + R^{(\varepsilon)}(t, x),
\]

such that the following hold:

1. The noise field \( \xi^{(\varepsilon)} \) converges in law to white noise on \( \mathbb{R}_{>0} \times \mathbb{R}^d \) as \( \varepsilon \to 0 \).
2. The remainder term \( R^{(\varepsilon)}(t, x) \) is \( o_P \) of the first three terms on the right and their sum, meaning that \( R^{(\varepsilon)}(t, x) \) divided by any of the first three terms, or by their sum, tends to zero in probability as \( \varepsilon \to 0 \).
Remarks

- It may seem as if $\nu(\varepsilon)$, $\lambda(\varepsilon)$ and $D(\varepsilon)$ should not be allowed to vary with $\varepsilon$, since the coefficients of $\Delta f$, $|\nabla f|^2$, and $\xi$ in the KPZ equation are constants.

- However, this is not true. In the KPZ literature, it is understood that the coefficients can be allowed to vary when taking a scaling limit, and even be allowed to tend to zero or blow up to infinity.

- This is especially true in dimensions higher than one. For example, the Family–Vicsek scaling for 2D surfaces requires this.
Motivation

The motivation for this ‘local’ definition of KPZ growth comes from a desire to create a universal framework for scaling limits of discrete surface growth, analogous to diffusion approximations of Markov chains.

Recall that a diffusion process $\{X_t\}_{t \geq 0}$ adapted to a filtration $\{\mathcal{F}_t\}_{t \geq 0}$ is characterized by the ‘local growth conditions’

$$
\mathbb{E}(X_{t+h} - X_t | \mathcal{F}_t) = a(X_t)h + o(h),
$$

$$
\mathbb{E}((X_{t+h} - X_t)^2 | \mathcal{F}_t) = b(X_t)h + o(h),
$$

$$
\mathbb{E}|X_{t+h} - X_t|^3 = o(h)
$$

as $h \to 0$, for any fixed $t$. This gives an informal representation of the stochastic differential equation

$$
dX_t = a(X_t)dt + \sqrt{b(X_t)}dB_t
$$

without actually defining it rigorously. The purpose of local KPZ growth is to do the same for surface growth.
A class of growing random surfaces

- Recall that $A = \{0, \pm e_1, \ldots, \pm e_d\}$.
- Let $\phi : \mathbb{R}^A \rightarrow \mathbb{R}$ be a function.
- Let $z = \{z_{t,x}\}_{t \in \mathbb{Z}_{\geq 0}, x \in \mathbb{Z}^d}$ be a collection of i.i.d. random variables, which we will refer to as the ‘noise field’.
- Given $\varepsilon > 0$, consider a function $f_\varepsilon : \mathbb{Z}_{\geq 0} \times \mathbb{Z}^d \rightarrow \mathbb{R}$ defined as follows: $f_\varepsilon(0, x) = 0$ for all $x$, and for each $t \geq 0$,
  $$f_\varepsilon(t + 1, x) := \phi((f_\varepsilon(t, x + a))_{a \in A}) + \varepsilon z_{t+1,x}.$$  
- Imagine $f_\varepsilon(t, x)$ to be the height of a $d$-dimensional random surface at time $t$ and location $x$. The above recursion says that the height at time $t + 1$ is a function of the heights at $x$ and its neighbors at time $t$, plus an independent random fluctuation.
- This setup was introduced in Chatterjee ’21a, ’21b and Chatterjee & Souganidis ’21.
- We will later see that the model of directed polymers in a random environment can be put in this framework.
Assumptions about $\phi$

- **Notation:** $1 \in \mathbb{R}^A$ denotes the vector of all 1’s. For $u \in \mathbb{R}^A$, $\overline{u}$ denotes the average of the coordinates of $u$. For $u, v \in \mathbb{R}^A$, we write $u \geq v$ if $u_a \geq v_a$ for each $a \in A$.

- We make the following assumptions about $\phi$.

- **Equivariance under constant shifts:** For all $u \in \mathbb{R}^A$ and $c \in \mathbb{R}$, $\phi(u + c1) = \phi(u) + c$.

- **Zero at the origin:** $\phi(0) = 0$.

- **Monotonicity:** $\phi(u) \geq \phi(v)$ whenever $u \geq v$.

- **Symmetry:** $\phi(u)$ remains unchanged under any permutation of the coordinates of $u$.

- **Regularity:** $\phi$ is $C^2$ in a neighborhood of the origin.

- **Nondegeneracy:** $\text{Hess} \phi(0) \neq 0$.

- There is one more crucial assumption, stated in the next slide...
The Edwards–Wilkinson surface growth model fits into our framework with \( \phi(u) = \bar{u} \). We assume that our surface grows at least as fast as the Edwards–Wilkinson surface, meaning that \( \phi(u) \geq \bar{u} \) for all \( u \).

Moreover, we assume that this domination is strict, in the following sense: If \( \{u_n\}_{n \geq 1} \) is a sequence such that \( \phi(u_n) - \bar{u}_n \to 0 \), then \( u_n - \bar{u}_n1 \to 0 \).

This is the assumption of strict Edwards–Wilkinson domination.

We will later see that the model of directed polymers in a random environment satisfies all of the above assumptions.
We make the following assumptions about the noise field $z$.

- **Zero mean.** $\mathbb{E}(z_{t,x}) = 0$.
- **Boundedness.** There is some constant $B$ such that $|z_{t,x}| \leq B$ with probability one.
- **Absolute continuity.** The law of $z_{t,x}$ is absolutely continuous with respect to Lebesgue measure.
Let $f_\varepsilon$ be defined using $\phi$ and the noise field $z$, as in the preceding slides.

Let $\alpha(\varepsilon)$, $\beta(\varepsilon)$ and $\gamma(\varepsilon)$ be any positive real numbers such that $\alpha(\varepsilon)$ and $\beta(\varepsilon)$ tend to zero as $\varepsilon \to 0$.

Define $f^{(\varepsilon)} : \mathbb{R}_{>0} \times \mathbb{R}^d \to \mathbb{R}$ as

$$f^{(\varepsilon)}(t, x) := \gamma(\varepsilon)f_\varepsilon([\alpha(\varepsilon)^{-1}t], [\beta(\varepsilon)^{-1}x]).$$
Theorem (C., 2021)

Suppose that all the stated assumptions about $\phi$ and $z$ are satisfied. Let $f^{(\varepsilon)}$ be the rescaled surface defined in the previous slide, with any $\alpha(\varepsilon)$, $\beta(\varepsilon)$ and $\gamma(\varepsilon)$, the sole condition being that $\alpha(\varepsilon), \beta(\varepsilon) \to 0$ as $\varepsilon \to 0$. Then as $\varepsilon \to 0$, $f^{(\varepsilon)}$ has local KPZ behavior, with

\[
\begin{align*}
\nu(\varepsilon) &= \frac{\beta(\varepsilon)^2}{(2d + 1)\alpha(\varepsilon)}, \\
\lambda(\varepsilon) &= \frac{2(q - r)\beta(\varepsilon)^2}{\alpha(\varepsilon)\gamma(\varepsilon)}, \\
D(\varepsilon) &= \frac{\sigma^2\varepsilon^2 \beta(\varepsilon)^d \gamma(\varepsilon)^2}{\alpha(\varepsilon)},
\end{align*}
\]

where $\sigma^2$ is the variance of the noise variables, $q$ is the value of the diagonal elements of $\text{Hess } \phi(0)$ (which are all equal due to the symmetry of $\phi$), and $r$ is the value of the off-diagonal elements of $\text{Hess } \phi(0)$. 
Remark

What if we want $\nu$, $\lambda$ and $D$ to remain bounded above and below by constants as $\varepsilon \to 0$?

Recall that up to constants, $\nu \sim \beta^2 \alpha^{-1}$, $\lambda \sim \beta^2 \alpha^{-1} \gamma^{-1}$, and $D \sim \varepsilon^2 \beta^d \gamma^2 \alpha^{-1}$.

So if we want $\nu \sim 1$, we need $\beta^2 \sim \alpha$, which gives $\lambda \sim \gamma^{-1}$.

Thus, to enforce, $\lambda \sim 1$, we would need $\gamma \sim 1$.

Finally, to have $D \sim 1$, we use $\beta^2 \sim \alpha$ and $\gamma \sim 1$ to get $\varepsilon^2 \beta^{d-2} \sim 1$.

But since $\beta \to 0$ as $\varepsilon \to 0$, this is impossible unless $d = 1$. Thus, any local KPZ limit in $d \geq 2$ must necessarily have some of the coefficients tending to zero or infinity as $\varepsilon \to 0$.

When $d = 1$, the above argument shows that there is exactly one way of getting $\nu$, $\lambda$ and $D$ to be bounded above and below by constants, which is to take $\beta \sim \varepsilon^2$, $\alpha \sim \varepsilon^4$, and $\gamma \sim 1$.

We will see later that for directed polymers, this is the intermediate disorder regime of Alberts, Khanin & Quastel ’14.
Recall that \( A = \{0, \pm e_1, \ldots, \pm e_d\} \).

Let \( \phi : \mathbb{R}^A \to \mathbb{R} \) be defined as

\[
\phi(u) := \log \left( \frac{1}{2d + 1} \sum_{a \in A} e^{u_a} \right).
\]

Easy to check: \( \phi \) is equivariant under constant shifts, zero at the origin, monotone, symmetric, \( C^2 \), and \( \text{Hess} \phi(0) \neq 0 \).

Moreover, \( \text{Hess} \phi \) positive semidefinite everywhere, which shows that \( \phi \) is convex.

Thus, the following lemma shows that \( \phi \) strictly dominates Edwards–Wilkinson growth.

**Lemma (C., 2021)**

If \( \phi : \mathbb{R}^A \to \mathbb{R} \) is equivariant under constant shifts, zero at the origin, monotone, symmetric, \( C^2 \) in a neighborhood of the origin, \( \text{Hess} \phi(0) \neq 0 \), and \( \phi \) is convex, then \( \phi \) satisfies the strict Edwards–Wilkinson domination condition.
Let $f_\varepsilon$ be the discrete random surface generated using the $\phi$ displayed in the previous slide and a random field $z = \{z_{t,x}\}_{t \in \mathbb{Z}_0, x \in \mathbb{Z}^d}$.

By induction on $t$, it is easy to show that

$$f_\varepsilon(t, x) = \log \left[ \frac{1}{(2d + 1)^{t-1}} \sum_{P \in \mathcal{P}_t} \exp \left( \varepsilon \sum_{i=0}^{t-1} z_{t-i,x+p_i} \right) \right],$$

where $\mathcal{P}_t$ is the set of all $P = (p_0, \ldots, p_{t-1}) \in (\mathbb{Z}^d)^t$ such that $p_0 = 0$ and $|p_i - p_{i-1}| \leq 1$ for each $i$, where $|\cdot|$ is the Euclidean norm.

This is the log-partition function of the $(d + 1)$-dimensional directed polymer model on lazy random walk paths of length $t - 1$ at inverse temperature $\varepsilon$, in the random environment $z$.

Thus, we arrive at the result displayed in the next slide.
Theorem (C., 2021)

Consider the random surface generated by the \((d + 1)\)-dimensional directed polymer model at inverse temperature \(\varepsilon\). Then the surface behaves locally like

\[
\tilde{\partial}_t f = \nu \tilde{\Delta} f + \frac{\lambda}{2} |\tilde{\nabla} f|^2 + \sqrt{D} \xi + \text{lower order error term},
\]

as we send \(\varepsilon \to 0\) and scale space and time \textit{in any way we like}, where \(\tilde{\partial}_t f, \tilde{\Delta} f\) and \(|\tilde{\nabla} f|^2\) are discrete approximations to the time-derivative, spatial Laplacian, and the squared gradient of \(f\), \(\xi\) is an approximation of white noise, and \(\nu, \lambda,\) and \(D\) are functions of \(\varepsilon\) and our choice of space-time scaling (that were displayed in an earlier slide).
Intermediate disorder regime

- Alberts, Khanin & Quastel ’14 showed that a nontrivial (global) KPZ limit can be obtained for the log-partition function of the (1 + 1)-dimensional directed polymer model by taking the inverse temperature proportional to $n^{-1/4}$, where $n$ is the length of the polymer, and looking at polymers in a spatial window of width $n^{1/2}$. They called this the intermediate disorder regime.

- In our language, the length of the polymer is $\alpha(\varepsilon)^{-1}$, the spatial window has width $\beta(\varepsilon)^{-1}$, and the inverse temperature is $\varepsilon$.

- So the intermediate disorder regime is equivalent to taking $\alpha(\varepsilon) \propto \varepsilon^4$, $\beta(\varepsilon) \propto \varepsilon^2$, and $\gamma(\varepsilon) = \text{constant}$.

- As we saw earlier, this is the only way to have that the coefficients $\nu(\varepsilon)$, $\lambda(\varepsilon)$ and $D(\varepsilon)$ in our equation do not depend on $\varepsilon$.

- Moreover, there is no way to do this when $d \geq 2$. 
Proof of the main result: Step 0

Define the ‘local average’ of \( f_\varepsilon \) as

\[
\bar{f}_\varepsilon(t, x) := \frac{1}{2d + 1} \sum_{a \in A} f_\varepsilon(t, x + a),
\]

If \( f_\varepsilon(t, x) \approx f_\varepsilon(t, x + a) \) for all \( a \in A \), then by Taylor expansion and using the properties of \( \phi \), one can show that

\[
f_\varepsilon(t + 1, x) - f_\varepsilon(t, x) = \underbrace{\bar{f}_\varepsilon(t, x) - f_\varepsilon(t, x)}_{\text{Laplacian term}} + K \underbrace{\sum_{a \in A} (f_\varepsilon(t, x + a) - \bar{f}_\varepsilon(t, x))^2}_{\text{Gradient squared term}} + \varepsilon z_{t+1, x} + \text{a remainder term of smaller order},
\]

where \( K \) is a constant depending on \( \phi \).

Thus, to get local KPZ behavior under arbitrary scaling limits, we need to have \( f_\varepsilon(t, x + a) \approx f_\varepsilon(t, x) \) for \( a \in A \) even if \( t \) and \( x \) are allowed to vary arbitrarily as \( \varepsilon \to 0 \).
Recall: \( f_\varepsilon(0, x) = 0 \) for all \( x \), and

\[
f_\varepsilon(t + 1, x) = \phi((f_\varepsilon(t, x + a))_{a \in A}) + \varepsilon z_{t+1, x},
\]

where \( A = \{0, \pm e_1, \ldots, \pm e_d\} \).

Let \( \partial_a \phi \) denote the derivative of \( \phi \) in coordinate \( a \in A \).

By equivariance under constant shifts, \( \phi(u + c1) = \phi(u) + c \).

Taking derivative in \( c \) at \( c = 0 \) shows that \( \sum_{a \in A} \partial_a \phi(u) = 1 \).

By monotonicity, \( \partial_a \phi(u) \geq 0 \).

Thus, at any \( u \), the derivatives of \( \phi \) are nonnegative and sum to 1.
Proof of the main result: Step 2

- Fix a realization of $f_\varepsilon$, and some $t \geq 1$ and $x \in \mathbb{Z}^d$.
- Define a random walk on $\mathbb{Z}^d$ as follows.
- The walk starts at $x$ at time $t$, and goes backwards in time, until reaching time $0$.
- If the walk is at $y \in \mathbb{Z}^d$ at time $s \geq 1$, then at time $s - 1$ it moves to $y + a$ with probability $\partial_a \phi((f_\varepsilon(s - 1, y + a))_{a \in A})$, for $a \in A$.
- By the observation from the previous slide, these numbers are nonnegative and sum to 1. Therefore, this describes a legitimate random walk on $\mathbb{Z}^d$, moving backwards in time.
- Inductively, one can show that

$$\frac{\partial}{\partial z_{s,y}} f_\varepsilon(t, x) = \varepsilon \mathbb{P}(S_s = y)$$

for any $1 \leq s \leq t$ and $y \in \mathbb{Z}^d$. Note that the probability on the right is conditional probability given $f_\varepsilon$. 

Sourav Chatterjee
Local KPZ behavior under arbitrary scaling limits
Proof of the main result: Step 3

- In particular, for any $1 \leq s \leq t$,

$$\sum_{y \in \mathbb{Z}^d} \left| \frac{\partial}{\partial z_{s,y}} f_\varepsilon(t, x) \right| = \varepsilon.$$

- By this identity and the multivariate mean-value theorem, it follows that if $z_{1,y}$ is replaced by 0 for each $y$, then the value of $f_\varepsilon(t, x)$ changes by at most $B\varepsilon$, where $B$ is a constant upper bound on the magnitude of the noise variables.

- Let $g_\varepsilon(t, x)$ be the value of $f_\varepsilon(t, x)$ after replacing all $z_{1,y}$ by 0.

- Note that $g_\varepsilon(1, x) = 0$ for each $x$. Thus, $g_\varepsilon$ is just like $f_\varepsilon$, except that instead of starting with an all zero initial condition at time 0, we start with an all zero initial condition at time 1.

- Thus, $g_\varepsilon(t + 1, x)$ has the same law as $f_\varepsilon(t, x)$.

- This gives us:

$$\mathbb{E}(f_\varepsilon(t + 1, x) - f_\varepsilon(t, x)) = \mathbb{E}(f_\varepsilon(t + 1, x) - g_\varepsilon(t + 1, x)) \leq B\varepsilon.$$
Proof of the main result: Step 3

Since the law of $f_\varepsilon(t, x)$ is the same for all $x$, this gives

$$
\mathbb{E}(f_\varepsilon(t + 1, x) - \bar{f}_\varepsilon(t, x)) = \mathbb{E}(f_\varepsilon(t + 1, x) - f_\varepsilon(t, x)) \leq B\varepsilon.
$$

Since the noise variables have mean zero,

$$
\mathbb{E}(f_\varepsilon(t + 1, x)) = \mathbb{E}(\phi((f_\varepsilon(t, x + a))_{a \in A}) + \varepsilon z_{t+1,x})
= \mathbb{E}(\phi((f_\varepsilon(t, x + a))_{a \in A})).
$$

Combining, and using equivariance of $\phi$ under constant shifts, we get

$$
\mathbb{E}(\phi(q_\varepsilon(t, x))) \leq B\varepsilon,
$$

where

$$
q_\varepsilon(t, x) := (f_\varepsilon(t, x + a) - \bar{f}_\varepsilon(t, x))_{a \in A}.
$$
Proof of the main result: Step 4

- Note that the vector \( q_\varepsilon(t, x) \) belongs to the hyperplane \( H := \{ u \in \mathbb{R}^A : u = 0 \} \).
- By Edwards–Wilkinson domination, \( \phi \) is nonnegative everywhere on this hyperplane. Thus, for any \( \eta > 0 \),

\[
\mathbb{P}(\phi(q_\varepsilon(t, x)) > \eta) \leq \frac{\mathbb{E}(\phi(q_\varepsilon(t, x)))}{\eta} \leq \frac{B_\varepsilon}{\eta}.
\]

- By strict EW domination, \( \phi(u_n) \to 0 \) implies \( u_n \to 0 \) on \( H \).
- This implies that for any \( \delta > 0 \), there exists \( \eta(\delta) > 0 \) such that \( u \in H \) and \( \phi(u) \leq \eta(\delta) \) implies \( |u| \leq \delta \).
- Thus,

\[
\mathbb{P}(|q_\varepsilon(t, x)| > \delta) \leq \mathbb{P}(\phi(q_\varepsilon(t, x)) > \eta(\delta)) \leq \frac{B_\varepsilon}{\eta(\delta)}.
\]

- Note: The bound has no dependence on \( t \) and \( x \).
- Thus, if \( \varepsilon \to 0 \) and \( t_\varepsilon, x_\varepsilon \) vary arbitrarily with \( \varepsilon \), the above inequality shows that \( |q_\varepsilon(t_\varepsilon, x_\varepsilon)| \to 0 \) in probability.
Since $q_\varepsilon(t, x) = (f_\varepsilon(t, x + a) - \overline{f}_\varepsilon(t, x))_{a \in A}$, this already shows that we can apply Taylor expansion even if $t$ and $x$ vary arbitrarily as $\varepsilon \to 0$.

Some more work is needed to establish that the remainder term is negligible compared to the other terms. This requires the assumption that $\text{Hess } \phi(0) \neq 0$. I will omit this part.