The Allen–Cahn equation with generic initial datum

Martin Hairer[1,2] · Khoa Lê[3,4] · Tommaso Rosati[5]

Abstract
We consider the Allen–Cahn equation $\partial_t u - \Delta u = u - u^3$ with a rapidly mixing Gaussian field as initial condition. We show that provided that the amplitude of the initial condition is not too large, the equation generates fronts described by nodal sets of the Bargmann–Fock Gaussian field, which then evolve according to mean curvature flow.

1 Introduction

The aim of the present article is to prove that nodal sets of a smooth Gaussian field, known as the Bargmann–Fock field, arise naturally as a random initial condition to evolution by mean curvature flow.

This problem is related to understanding the long-time behaviour of mean curvature flow: for a sufficiently generic initial datum composed of clusters with typical lengthscale of order 1, one expects that at time $t \gg 1$ the clusters have coarsened in

✉ Tommaso Rosati
T.Rosati@warwick.ac.uk

Martin Hairer
m.hairer@imperial.ac.uk; martin.hairer@epfl.ch

Khoa Lê
k.le@leeds.ac.uk

1   Imperial College London, London, UK

2   EPFL, Lausanne, Switzerland

3   TU Berlin, Berlin, Germany

4   Present Address: School of Mathematics, University of Leeds, Leeds, UK

5   University of Warwick, Coventry, UK

🖄 Springer

such a way that the typical lengthscale is of order $\sqrt{t}$. In fact, one would even expect that upon rescaling by $\sqrt{t}$, the clusters become self-similar/stationary at large times [3].

An understanding of such a behaviour remains currently beyond the reach of rigorous analysis although upper bounds on the coarsening rate can be proven via deterministic arguments, see e.g. [9] for the related Cahn–Hilliard dynamics. On the other hand, lower bounds on the rate of coarsening can only be expected to hold for sufficiently generic initial conditions, since for degenerate initial conditions one may not see any coarsening at all.

This motivates the question, addressed in the present work, of what such generic initial condition should look like. Notably, the correlation structure we obtain is the same as the first order approximation of the longtime two-point correlation predicted by Ohta, Jasnow and Kawasaki [10].

A natural way to construct random initial conditions to mean curvature flow is to consider the fronts formed by the dynamics of the Allen–Cahn equation with white noise initial data. Unfortunately this is not feasible, since the scaling exponent $-\frac{d}{2}$ of white noise on $\mathbf{R}^d$, with $d \geq 2$, lies below or at the critical exponent -1 (or $-\frac{2}{3}$ depending on what one really means by "critical" in this context), under which one does not expect any form of local well-posedness result for the equation.

Instead, we consider the following setting. Let η be a white noise on $\mathbf{R}^d$ with $d \geq 2$ and let φ be a Schwartz test function integrating to 1. Fix an exponent $\alpha \in (0, 1)$, and for each $\varepsilon \in (0, 1)$ and $x \in \mathbf{R}^d$ define

$$\varphi_x^\varepsilon(y) = \varepsilon^{-d}\varphi(\varepsilon^{-1}(x - y)), \qquad \eta_\varepsilon(x) = \varepsilon^{\frac{d}{2}-\alpha}\eta(\varphi_x^\varepsilon). \tag{1.1}$$

Here, the exponents are chosen in such a way that φ_x^ε converges to a Dirac distribution and typical values of η_ε are of order $\varepsilon^{-\alpha}$. Our aim is to describe the limit as $\varepsilon \to 0$ of the solutions to the Allen–Cahn equation with initial datum η_ε

$$(\partial_t - \Delta)u_\varepsilon = u_\varepsilon - u_\varepsilon^3, \quad u_\varepsilon(0, x) = \eta_\varepsilon(x). \tag{1.2}$$

The reason for restricting ourselves to $\alpha < 1$ is that 1 is precisely the critical exponent for which $\Delta\eta_\varepsilon$ and η_ε^3 are both of the same order of magnitude, i.e. when $\alpha + 2 = 3\alpha$.

For a fixed terminal time t, the behaviour of (1.2) is not very interesting since one has $\eta_\varepsilon \to 0$ weakly, so one would expect to also have $u_\varepsilon \to 0$. This is trivially the case for $\alpha < \frac{2}{3}$ since one has $\eta_\varepsilon \to 0$ in $\mathcal{C}^\beta$ for every $\beta < -\alpha$ and the Allen–Cahn equation is well-posed for arbitrary initial data in $\mathcal{C}^\beta$ if (and actually only if) $\beta > -\frac{2}{3}$. As a consequence of Proposition 3.15 below, we will see that it *is* still the case that $u_\varepsilon(t) \to 0$ at fixed t for any exponent $\alpha < 1$, but this is a consequence of more subtle stochastic cancellations.

It turns out that the natural time and length scales at which one does see non-trivial behaviour are given by

$$T_\varepsilon = \left(\frac{d}{2} - \alpha\right)\log \varepsilon^{-1}, \qquad L_\varepsilon = \sqrt{T_\varepsilon}.$$

The main result of this article is that for $\sigma > 1$, $u(\sigma T_\varepsilon, x L_\varepsilon) \to \pm 1$ on sets $\Omega_\sigma^{\pm 1}$, which evolve under mean curvature flow. The initial data (at time $\sigma = 1$) for that flow is given by the nodal domains of a centred Gaussian field $\{\Psi_1(x) : x \in \mathbf{R}^d\}$ with Gaussian correlation function, also known as the Bargmann–Fock field:

$$\mathbf{E}[\Psi_\sigma(x)\Psi_\sigma(y)] = \sigma^{-\frac{d}{2}} \exp\left(-\frac{|x-y|^2}{8\sigma}\right), \qquad \sigma > 0. \tag{1.3}$$

In what follows we will write Ψ short for Ψ_1. This Gaussian field emerges from the linearisation near zero of (1.2), within a time interval of order 1 (in the original variables) around $t_\star(\varepsilon)$ with

$$t_\star(\varepsilon) \overset{\text{def}}{=} T_\varepsilon + \tau_\star(\varepsilon), \quad \tau_\star(\varepsilon) \overset{\text{def}}{=} \frac{d}{4} \log\log \varepsilon^{-1} + \mathfrak{c}, \quad \mathfrak{c} \overset{\text{def}}{=} \frac{d}{4} \log\left(4\pi(d-2\alpha)\right), \tag{1.4}$$

where $\mathfrak{c}$ is chosen to make certain expressions look nicer. In fact, roughly up to time $t_\star(\varepsilon)$ the dynamics of u_ε is governed by the linearisation of (1.2) around the origin, but at that time the solution becomes of order 1 and starts to follow the flow $\Phi \colon \mathbf{R} \times \mathbf{R} \to [-1, 1]$ of the nonlinear ODE

$$\partial_t \Phi = \Phi - \Phi^3, \qquad \lim_{t \to -\infty} e^{-t}\Phi(t, u) = u, \tag{1.5}$$

which is given by

$$\Phi(t, u) = \frac{u}{\sqrt{e^{-2t} + u^2}}.$$

We will show that a good approximation of $u_\varepsilon(t_\star(\varepsilon), x L_\varepsilon)$ is given by $\Phi(0, \Psi(x))$.

Then, after an additional time of order T_ε, u_ε is finally governed by the large scale behaviour of the Allen–Cahn equation with random initial condition $\Phi(0, \Psi)$, namely by the mean curvature flow evolution of the set $\{\Psi = 0\}$. To capture the different time scales we introduce the process

$$U_\varepsilon(\sigma, x) \overset{\text{def}}{=} (e^{(1-\sigma)T_\varepsilon} \vee 1) u_\varepsilon(\sigma T_\varepsilon + \tau_\star(\varepsilon), x L_\varepsilon), \tag{1.6}$$

and we summarise our results in the theorem below.

Theorem 1.1 *The process U_ε converges in law as follows for $\varepsilon \to 0$*

$$U_\varepsilon(\sigma, x) \to \Psi_\sigma(x), \qquad\qquad \sigma \in (0, 1), x \in \mathbf{R}^d \ ,$$
$$U_\varepsilon(1 + t/T_\varepsilon, x) \to \Phi(t, \Psi(x)), \qquad\qquad t \in \mathbf{R}, x \in \mathbf{R}^d \ .$$

The convergences hold for fixed t or σ and locally uniformly over the variable x.

Denote furthermore by Ω_σ^+, Ω_σ^- respectively the mean curvature flow evolutions of the sets $\Omega_1^+ = \{x : \Psi(x) > 0\}$, $\Omega_1^- = \{x : \Psi(x) < 0\}$. Then, there exist a coupling

between $\{U_\varepsilon\}_{\varepsilon \in (0,1)}$ and Ψ such that, for every $\sigma > 1$, one has $\lim_{\varepsilon \to 0} U_\varepsilon(\sigma, x) = \pm 1$ locally uniformly for $x \in \Omega_\sigma^\pm$, in probability.

Remark 1.2 Although level set solutions to mean curvature flow with continuous initial datum are unique, they might exhibit fattening, meaning that the evolution of the zero level set $\Gamma_\sigma = \mathbf{R}^d \backslash (\Omega_\sigma^+ \cup \Omega_\sigma^-)$ might have a non-empty interior for some $\sigma > 1$. In dimension $d = 2$ this does not happen, as the initial interface is composed of disjoint smooth curves which evolve under classical mean curvature flow, c.f. Corollary 5.7. In dimension $d \geq 3$ this picture is less clear, and in addition the initial nodal set was recently proven to contain unbounded connected components [5].

1.1 Structure of the paper

The rest of this article is divided as follows. In Sect. 2, we reformulate Theorem 1.1 more precisely and we provide its proof, taking for granted the more technical results given in subsequent sections. In Sect. 3 we study a Wild expansion of the solution to (1.2) which allows us to take advantage of stochastic cancellations to control the solutions on a relatively short initial time interval of length smaller than $(1-\alpha) \log \varepsilon^{-1}$. In Sect. 4 we then show how the Bargmann–Fock field arises from this expansion and we track solutions up to timescales of order $t_\star$. In Sect. 5, we finally conclude by showing that our estimates are sufficiently tight to allows us to patch into existing results on the convergence to mean curvature flow on large scales.

1.2 Notations

For a set $\mathcal{X}$ and two functions $f, g \colon \mathcal{X} \to \mathbf{R}$, we write $f \lesssim g$ if there exists a constant $c > 0$ such that $f(x) \leq cg(x)$ for all $x \in \mathcal{X}$. If $\mathcal{X}$ is a locally compact metric space we define $C_{\mathrm{loc}}(\mathcal{X})$ as the space of continuous and locally bounded functions, equipped with the topology of uniform convergence on compact sets. We will write P_t, P_t^1 for the following semigroups on $\mathbf{R}^d$:

$$P_t \varphi = e^{t\Delta} \varphi, \qquad P_t^1 \varphi = e^{t(\Delta+1)} \varphi, \quad t \geq 0.$$

We denote with $\mathrm{sgn} \colon \mathbf{R} \to \{-1, 0, 1\}$ the sign function, with the convention that $\mathrm{sgn}(0) = 0$.

2 Main results

In what follows we consider $(t, x) \mapsto u_\varepsilon(t, x)$ as a process on $\mathbf{R} \times \mathbf{R}^d$ by defining $u_\varepsilon(t, x) = u_\varepsilon(0, x)$ for $t < 0$. We then show the first part of Theorem 1.1, namely

Theorem 2.1 *Let u_ε solve (1.2) and Ψ_σ, Φ, U_ε be given respectively by (1.3), (1.5) and (1.6). Then*

1. *For any $\sigma \in (0, 1)$ the process $\{U_\varepsilon(\sigma, x) : x \in \mathbf{R}^d\}$ converges locally uniformly in law to $\{\Psi_\sigma(x) : x \in \mathbf{R}^d\}$.*

2. *The process $\{u_\varepsilon(t_\star(\varepsilon) + t, xL_\varepsilon) : (t, x) \in \mathbf{R} \times \mathbf{R}^d\}$ converges locally uniformly in law to $\{\Phi(t, \Psi(x)) : (t, x) \in \mathbf{R} \times \mathbf{R}^d\}$.*

Proof The first statement follows, for $\sigma \in \left(0, \frac{1-\alpha}{d/2-\alpha}\right)$, from similar calculations as in Lemma 4.4 in combination with Proposition 3.15. For $\sigma \in \left[\frac{1-\alpha}{d/2-\alpha}, 1\right)$ it follows again by similar calculations as in Lemma 4.4, this time in combination with Lemma 4.3. Let us pass to the second statement. For some $\overline{\alpha} \in (\alpha, 1)$ consider

$$t_1(\varepsilon) = (\overline{\alpha} - \alpha) \log \varepsilon^{-1}, \qquad t_2(\varepsilon) = T_\varepsilon - \frac{1}{2} \log \log \varepsilon^{-1}.$$

It is shown in Proposition 4.6 that the limit as $\varepsilon \to 0$ of the process $(t, x) \mapsto u_\varepsilon(t_\star + t, L_\varepsilon x)$ is identical, in probability, to that of $(t, x) \mapsto w_\varepsilon^N(t_\star + t, L_\varepsilon x)$ for some fixed N sufficiently large. From the definition of w_ε^N, we have

$$w_\varepsilon^N(t, x) = \frac{P_{t-t_1}^1 u_\varepsilon^N(t_1, x)}{(1 + (1 - e^{-2(t-t_2)})(P_{t-t_1}^1 u_\varepsilon^N(t_1, x))^2)^{1/2}},$$

where u_ε^N is the Wild expansion truncated at level N given by (3.5). Applying Lemma 4.4, we see that the process $(t, x) \mapsto P_{t_\star + t - t_1}^1 u_\varepsilon^N(t_1, L_\varepsilon x)$ converges to $(t, x) \mapsto e^t \Psi(x)$ in $C_{\text{loc}}(\mathbf{R} \times \mathbf{R}^d)$ in distribution. Since furthermore $t_\star \gg t_1$, it follows that the process $(t, x) \mapsto w_\varepsilon^N(t_\star + t, L_\varepsilon x)$ converges in $C_{\text{loc}}(\mathbf{R} \times \mathbf{R}^d)$ in distribution to the process

$$(t, x) \mapsto \frac{e^t \Psi(x)}{(1 + e^{2t} \Psi(x)^2)^{1/2}} = \Phi(t, \Psi(x)),$$

thus concluding the proof. $\qquad\square$

We now turn to the proof of the second part of Theorem 1.1. This is trickier to formulate due to the potential fattening of level sets already alluded to in Remark 1.2. In particular, let us denote by $(\sigma, x) \mapsto v(f; \sigma, x) \in \{-1, 0, 1\}$ the sign of the level set solution to mean curvature flow associated to the initial interface $\{f = 0\}$ in the sense of Definition 5.5, with the difference that the initial condition is given at time $\sigma = 1$ in the present scale. We will then write Γ for the random interface $\Gamma = \{(\sigma, x) \in [1, \infty) \times \mathbf{R}^d : v(\Psi; \sigma, x) = 0\} \subseteq \mathbf{R}^{d+1}$ and $\Gamma_\sigma = \{x \in \mathbf{R}^d : (\sigma, x) \in \Gamma\}$. In addition, in order to define locally uniform convergence in the complement of Γ, let us introduce for any $\delta \in (0, 1)$ the random sets

$$K_\delta = \{z = (\sigma, x) \in (1, \infty) \times \mathbf{R}^d : |z| \le \delta^{-1}, \ \sigma > 1 + \delta, \ d(z, \Gamma) \ge \delta\}$$
$$K_\delta^1 = \{x \in \mathbf{R}^d : |x| \le \delta^{-1}, \ d(x, \Gamma_1) \le \delta\}. \tag{2.1}$$

Here $d(p, X)$ is the usual point-to-set distance. Furthermore, we can define the (random) norms

$$\|f\|_{K_\delta} = \sup_{(\sigma, x) \in K_\delta} |f(\sigma, x)|$$

and analogously for K_δ^1. With these definitions at hand, the next result describes the formation of the initial interface, which appears if we wait an additional time of order $\log \log \varepsilon^{-1}$. Hence we define, for $\kappa > 0$:

$$t_\star^\kappa(\varepsilon) \overset{\text{def}}{=} t_\star(\varepsilon) + \kappa \log \log \varepsilon^{-1}. \tag{2.2}$$

Proposition 2.2 *Consider any sequence of times* $\{t(\varepsilon)\}_{\varepsilon \in (0,1)} \subseteq (0, \infty)$ *such that for some* $\kappa > 0$

$$\liminf_{\varepsilon \to 0}\{t(\varepsilon) - t_\star^\kappa(\varepsilon)\} \geq 0, \qquad \limsup_{\varepsilon \to 0} \frac{t(\varepsilon) - t_\star(\varepsilon)}{T_\varepsilon} = 0.$$

Then there exists a coupling between u_ε *and* Ψ *such that for any* $\delta, \zeta \in (0, 1)$

$$\lim_{\varepsilon \to 0} \mathbf{P}(\|u_\varepsilon(t(\varepsilon), \cdot L_\varepsilon) - \text{sgn}(\Psi(\cdot))\|_{K_\delta^1} > \zeta) = 0.$$

Proof Up to taking subsequences, it suffices to prove the result in either of the following two cases:

$$\liminf_{\varepsilon \to 0}\{t(\varepsilon) - t_\star^{\frac{1}{2}}(\varepsilon)\} > 0, \qquad \text{or} \qquad \limsup_{\varepsilon \to 0}\{t(\varepsilon) - t_\star^{\frac{1}{2}}(\varepsilon)\} \leq 0.$$

In the first case, the result is a consequence of Proposition 5.4 while in the second case it follows from Lemma 5.3. In both cases the choice of the coupling is provided by Lemma 5.2. $\qquad \square$

In the case $t(\varepsilon) = t_\star^\kappa(\varepsilon)$ for some $\kappa \in (0, \frac{1}{4})$ one can also obtain the above result as a consequence of Proposition 4.6. Finally, we show that if we wait for an additional time of order T_ε, the interface moves under mean curvature flow. This result is a consequence of Proposition 5.6.

Theorem 2.3 *Consider* $U_\varepsilon(\sigma, x)$ *as in* (1.6) *for* $\sigma > 1$, $x \in \mathbf{R}^d$. *There exists a coupling between* u_ε *and* Ψ *such that for any* $\delta, \zeta \in (0, 1)$

$$\lim_{\varepsilon \to 0} \mathbf{P}\big(\|U_\varepsilon(\cdot) - v(\Psi; \cdot)\|_{K_\delta} > \zeta\big) = 0.$$

3 Wild expansion

The next two sections are devoted to tracking u_ε for times up to and around $t_\star(\varepsilon)$. In order to complete this study we divide the interval $[0, t_\star(\varepsilon)]$ into different regions. Let $\bar{\alpha} \in (\alpha, 1)$ be fixed and define

$$t_1(\varepsilon) \overset{\text{def}}{=} (\bar{\alpha} - \alpha) \log \varepsilon^{-1},$$
$$t_2(\varepsilon) \overset{\text{def}}{=} \left(\frac{d}{2} - \alpha\right) \log \varepsilon^{-1} - \frac{1}{2} \log \log \varepsilon^{-1}. \tag{3.1}$$

We observe that for ε small one has

$$0 \ll t_1(\varepsilon) \ll t_2(\varepsilon) \ll t_\star(\varepsilon).$$

The specific forms of these times are chosen such that in complement to the moment estimates, the error terms in various estimates which appear below remain small when $\varepsilon \to 0$. The constant $-1/2$ which appears in $t_2(\varepsilon)$ can be replaced by any negative number, and in spatial dimensions 3 or higher can be replaced by 0. The separation of time scales is due to that fact that the linear and nonlinear terms in (1.2) have different effects on the solution over each period.

During the time period $[0, t_1(\varepsilon)]$, the Laplacian dominates and (1.2) can be treated as a perturbed heat equation. One expects that a truncated Wild expansion u_ε^N (where N is the level of the truncation) provides a good approximation for u_ε during this initial time period, see Proposition 3.15 at the end of this section. During the time period $[t_1(\varepsilon), t_2(\varepsilon)]$, the linear term increases the size of the solution from $\mathcal{O}(\varepsilon^\theta)$ for any $0 < \theta < \frac{d}{2} - \overline{\alpha}$ to almost a size of order 1. During this period, the estimate (3.27) in Proposition 3.15 does not reflect the actual size of the solution u_ε. However, the leading order term $X_\varepsilon^\bullet$ in u_ε^N (which is the solution to the linearisation near 0 of (1.2)) at time $t_2(\varepsilon)$ is of order $(\log \varepsilon^{-1})^{-d/4}$. Hence the solution u_ε still remains small, and we expect that the non-linear term u_ε^3 in (1.2) is negligible: see Lemma 4.3. Eventually one starts to see the classical dynamic of the Allen–Cahn equation during $[t_2(\varepsilon), t_\star(\varepsilon)]$, as explained in Proposition 4.6.

To establish the aforementioned results, we will frequently make use of the following formulation of the maximum principle.

Lemma 3.1 *Let $(a, b) \subset \mathbf{R}$ be a finite interval and let G, R be measurable functions on $(a, b) \times \mathbf{R}^d$ such that G is non-negative. Suppose that v is a function in $C^2([a, b] \times \mathbf{R}^d)$ satisfying*

$$(\partial_t - \Delta - 1)v = -Gv + R \quad on \quad (a, b) \times \mathbf{R}^d.$$
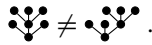
Then,

$$|v(t, x)| \le P_{t-a}^1 |v|(a, \cdot) + \int_a^t P_{t-s}^1 |R|(s) \, ds \quad \forall (t, x) \in [a, b] \times \mathbf{R}^d.$$

3.1 Moment estimates

Let $\mathcal{T}_3$ be the set of rooted ternary ordered trees defined inductively by postulating that $\bullet \in \mathcal{T}_3$ and, for any $\tau_i \in \mathcal{T}_3$, one has $\tau = [\tau_1, \tau_2, \tau_3] \in \mathcal{T}_3$. In general, for a finite collection of trees $\tau_1, \dots, \tau_n$ we denote by $\tau = [\tau_1, \dots, \tau_n]$ the tree obtained by connecting the roots of the τ_i with a new node, which in turn becomes the root of the tree τ:

$$[\tau_1, \dots, \tau_n] = \overset{\tau_1 \tau_2 \; \cdots \; \tau_n}{\bigvee}.$$

These trees are *ordered* in the sense that $[\tau_1, \tau_2, \tau_3] \neq [\tau_1, \tau_3, \tau_2]$ unless $\tau_2 = \tau_3$, and similarly for all other permutations. For example we distinguish the following two trees:

$$\raisebox{-0.5em}{\includegraphics[height=2em]{tree1}} \neq \raisebox{-0.5em}{\includegraphics[height=2em]{tree2}} .$$

For $\tau \in \mathcal{T}_3$, we write $i(\tau)$ for the number of inner nodes of $[\tau]$ (*i.e.* nodes which are not leaves, but including the root) and $\ell(\tau)$ be the number of leaves of $[\tau]$. Then ℓ and i are related by

$$\ell(\tau) = 2i(\tau) + 1. \tag{3.2}$$

For each positive integer n, $\pi_n = (2n - 1)!!$ is the number of pairings of $2n$ objects so that each object is paired with exactly one other object. For $\tau = [\tau_1, \tau_2, \tau_3] \in \mathcal{T}_3$, we define X_ε^τ inductively by solving

$$\partial_t X_\varepsilon^\bullet = (\Delta + 1)X_\varepsilon^\bullet, \qquad\qquad X_\varepsilon^\bullet(0, \cdot) = \eta_\varepsilon(\cdot), \tag{3.3}$$
$$\partial_t X_\varepsilon^\tau = (\Delta + 1)X_\varepsilon^\tau - X_\varepsilon^{\tau_1} X_\varepsilon^{\tau_2} X_\varepsilon^{\tau_3}, \qquad\qquad X_\varepsilon^\tau(0, \cdot) = 0. \tag{3.4}$$

For $N \geq 1$, we write $\mathcal{T}_3^N \subset \mathcal{T}_3$ for the trees with $i(\tau) \leq N$ and we define

$$u_\varepsilon^N \overset{\text{def}}{=} \sum_{\tau \in \mathcal{T}_3^N} X_\varepsilon^\tau, \tag{3.5}$$

which is a truncated Wild expansion of u_ε. Then u_ε^N satisfies the following equation

$$(\partial_t - \Delta - 1)u_\varepsilon^N = -(u_\varepsilon^N)^3 + R_\varepsilon^N, \quad u_\varepsilon^N(0, \cdot) = u_\varepsilon(0, \cdot), \tag{3.6}$$

where

$$R_\varepsilon^N = \sum_{\tau_1, \tau_2, \tau_3 \in \mathcal{T}_3^N : [\tau_1, \tau_2, \tau_3] \notin \mathcal{T}_3^N} X_\varepsilon^{\tau_1} X_\varepsilon^{\tau_2} X_\varepsilon^{\tau_3}. \tag{3.7}$$

Indeed, using (3.4) and (3.3), we have

$$(\partial_t - \Delta - 1)u_\varepsilon^N = - \sum_{\tau = [\tau_1, \tau_2, \tau_3] \in \mathcal{T}_3^N \setminus \{\bullet\}} X_\varepsilon^{\tau_1} X_\varepsilon^{\tau_2} X_\varepsilon^{\tau_3}$$
$$= -\left(\sum_{\tau \in \mathcal{T}_3^N} X_\varepsilon^\tau \right)^3 + R_\varepsilon^N = -(u_\varepsilon^N)^3 + R_\varepsilon^N.$$

Let us now introduce some graphical notations to represent conveniently X_ε^τ and the integrals that will be associated to them. In what follows the negative heat kernel

$-P^1(t - s, x - y)$ (here the negative sign appears to keep track of the sign of the polynomial term in (3.4)) will be represented by a directed edge

$$-K_\frown((t, x), (s, y)) = -P^1(t - s, x - y) = \quad \begin{array}{c} \bullet\, (s, y) \\ \Big/ \\ \bullet\, (t, x) \end{array} \quad .$$

Each endpoint of an edge represents a space-time variable and the kernel is evaluated at their difference. Three different nodes $\circ$, $\bullet$ and $\bullet$ can be attached to an end of an edge, which represent respectively a fixed space-time variable, integration with respect to Lebesgue measure $ds\,dy$, and integration with respect to the (random) measure $-\delta(s)\eta_\varepsilon(y)\,ds\,dy$. Here, δ is the Dirac mass at 0 and the minus sign is again just a matter of convention. For example, we have

$$\begin{array}{c} \bullet \\ \Big| \\ \bullet\,(t, x) \end{array} = \int_{\mathbf{R}^d} P^1(t, x - y)\eta_\varepsilon(y)\,dy = X^\bullet_\varepsilon(t, x),$$

as well as $X^{\bullet\bullet\triangledown}_\varepsilon = \begin{array}{c}\triangledown\\ \bullet\end{array}$.

Since all the edges follow the natural direction toward the root, we may drop the arrows altogether and just write $X^\bullet_\varepsilon = \begin{array}{c}\bullet\\\bullet\end{array}$ and $X^{\bullet\bullet\triangledown}_\varepsilon = \begin{array}{c}\triangledown\\\bullet\end{array}$. For a general tree $\tau \in \mathcal{T}_3$, X^τ_ε is represented by the tree $[\tau]$ with its root coloured red and leaves coloured green.

Given a tree $\tau \in \mathcal{T}_3$, we would like to estimate the moment $\mathbf{E}[X^\tau_\varepsilon(t, x)]^2$. The contracting kernel (see (1.1))

$$K_\frown((s, y), (\bar{s}, \bar{y})) = \varepsilon^{d-2\alpha}\delta(s)\delta(\bar{s})\int_{\mathbf{R}^d} \varphi^\varepsilon_y(z)\varphi^\varepsilon_{\bar{y}}(z)\,dz, \qquad (3.8)$$

which appears when one contracts two noise nodes (*i.e.* two green nodes), is represented as a green edge. Note that since these kernels are symmetric, green edges are naturally undirected. Moreover, since we are only interested in upper bound and up to replacing φ^ε with $|\varphi^\varepsilon|$, we may assume that

$$\varphi^\varepsilon(x) \geq 0, \qquad \forall x \in \mathbf{R}^d.$$

Then for each ε, we define the *positive* kernel

$$p^\varepsilon_t(x) = \int_{\mathbf{R}^d} p_t(x - y)\varphi^\varepsilon_y(0)\,dy, \qquad p_t(x) = \frac{e^{-\frac{|x|^2}{4t}}}{(4\pi t)^{\frac{d}{2}}}.$$

From Young's inequality, we see that

$$\|p^\varepsilon_t\|_{L^2(\mathbf{R}^d)} \leq (\|p_t\|_{L^2(\mathbf{R}^d)}\|\varphi^\varepsilon_\cdot\|_{L^1(\mathbf{R}^d)}) \wedge (\|p_t\|_{L^1(\mathbf{R}^d)}\|\varphi^\varepsilon_\cdot\|_{L^2(\mathbf{R}^d)}) \leq c(t \vee \varepsilon^2)^{-\frac{d}{4}},$$

for some constant $c > 0$. Another elementary observation is the bound

$$
\begin{aligned}
\text{[graph]} \quad &= e^{t_1+t_2} \varepsilon^{d-2\alpha} \int_{\mathbf{R}^d} p_{t_1}^\varepsilon(x_1 - z) p_{t_2}^\varepsilon(x_2 - z) \, \mathrm{d}z \\
&\le (c\varepsilon^{\frac{d}{2}-\alpha})^2 e^{t_1+t_2} (t_1 \vee \varepsilon^2)^{-\frac{d}{4}} (t_2 \vee \varepsilon^2)^{-\frac{d}{4}}.
\end{aligned} \tag{3.9}
$$

Here we did not use any green nodes because the distinction between green and black nodes is captured by the difference between the kernels $K_\frown$ and $K_\frown$, and as per assumption every black node indicates integration over both space and time variables.

We begin with an estimate for $\mathbf{E}\overset{2}{\overset{\bullet}{\bullet}}(t, x)$. The second moment is obtained by contracting two leaves,

$$
\mathbf{E}[\overset{2}{\overset{\bullet}{\bullet}}(t, x)] = \text{[graph]}_{(t,x)} \le (ce^t \varepsilon^{\frac{d}{2}-\alpha})^2 (t \vee \varepsilon^2)^{-\frac{d}{2}}.
$$

For a general tree τ, the second moment $\mathbf{E}|X_\varepsilon^\tau(t, x)|^2$ is obtained by summing over all pairwise contractions between the leaves of $[\tau, \bar{\tau}]$, where $\bar{\tau}$ is an identical copy of the tree τ. For instance,

$$
\mathbf{E}\text{[graph]}^2 = 6 \text{[graph]} + 9 \text{[graph]}. \tag{3.10}
$$

Let us take a closer look at each term. In particular, we can extend our graphical notation in order to obtain efficient estimates for such trees.

More tree spaces. First, we observe that each of the graphs above comes with a specific structure: it consists of a tree with an even number of leaves, together with a pairing among the leaves. In addition the root is coloured red and every node that is not a leaf is the root of up to three planted trees in $[\mathcal{T}_3]$. This suggest the introduction of the following space of paired forests $\mathcal{F}$ (later on it will be convenient to work with forests and not only trees).

Let $\mathcal{T}$ be the set of all finite rooted trees and for $\{\tau_i\}_{i=1}^k \subseteq \mathcal{T}$ define $\ell(\tau_1 \sqcup \cdots \sqcup \tau_k) = \ell(\tau_1) + \cdots \ell(\tau_k)$ the total number of leaves of a forest. Then define $\mathcal{F}$ as

$$
\mathcal{F} = \Big\{ (\tau_1 \sqcup \cdots \sqcup \tau_k, \gamma) : k \in \mathbf{N}, \ \tau_i \in \mathcal{T},
$$
$$
\text{with } \ell(\tau_1 \sqcup \cdots \sqcup \tau_k) \in 2\mathbf{N} \text{ and } \gamma \in \mathcal{P}_\ell(\tau_1, \ldots, \tau_k) \Big\}.
$$

where $\mathcal{P}_\ell$ is the set of possible pairings among the union of all the leaves of the trees $\tau_1, \ldots, \tau_k$.

We naturally think of elements of $\mathcal{F}$ as coloured graphs $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, with the pairing γ giving rise to green edges. Observe that the forest structure of $\tau_1 \sqcup \cdots \sqcup \tau_k$ induces a partial order on the set $\mathcal{V}$ of all vertices by defining $v > v'$ if $v' \ne v$ lies on the unique (directed) path joining v to its relative root. As we mentioned, we will be

especially interested in elements of $\mathcal{F}$ where each tree τ_i is an element of $[\mathcal{T}_3]$, with the twist that up to two such trees can be planted on the same root:

$$\text{Each } \tau_i \text{ consists of either 1 or 2 trees in } [\mathcal{T}_3] \text{ glued at their roots.} \qquad (3.11)$$

We then set

$$\mathcal{F}_3 = \Big\{ (\tau_1 \sqcup \cdots \sqcup \tau_k, \gamma) \in \mathcal{F} \text{ such that (3.11) holds true}\}.$$

Finally, we want to introduce a colouring for $\mathcal{G} \in \mathcal{F}$. We already mentioned that γ is represented by green edges among the leaves and that all the roots are coloured red, but we will also introduce yellow vertices.

Colourings. To simplify the upcoming constructions, let us now write $c(e) \in \{\frown, \curvearrowright\}$ and $c(v) \in \{\bullet, \circ, \bullet\}$ for the colour of edges and vertices of some graph $\mathcal{G} \in \mathcal{F}$ and let K_c denote the kernel associated to each edge colour. As we will see later on, colours will appear only in certain locations and will be valuated in specific ways. To clarify these points we compile the following tables for the nodes

Colour	Where?	Valuation
🔴	Roots	Suprema over spatial variables
○	Inner nodes of paths	Disappear under spatial integration
●	Leaves or untouched inner nodes	Integrate

and define kernels

$$K_{\frown} = (3.8), \qquad K_{\curvearrowright}((t, x), (s, y)) = \begin{cases} P^1(t - s, x - y), & \text{if } t > s, \\ 0 & \text{else.} \end{cases}$$

Now, if $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, we write $\mathcal{V}_c$ for the set of vertices of colour c. Given $z_{\mathcal{V}} \in (\mathbf{R}^{d+1})^{\mathcal{V}}$ and $A \subset \mathcal{V}$, we write $z_A \in (\mathbf{R}^{d+1})^A$ for the corresponding projection, as well as $t_A \in \mathbf{R}^A$ and $x_A \in (\mathbf{R}^d)^A$ for its temporal and spatial components. In the particular case when $A = \mathcal{V}_c$ for some colour c we will sometimes simply write z_c instead of $z_{\mathcal{V}_c}$. Finally, given an oriented edge $e \in \mathcal{E}$, we write $e = (e_-, e_+)$.

Structure of the estimates. With these notations at hand, for any subgraph $\mathcal{G}' = (\mathcal{V}', \mathcal{E}') \subseteq \mathcal{G} \in \mathcal{F}$, we set

$$K_{\mathcal{G}'}(z) = \prod_{e \in \mathcal{E}'} K_{c(e)}(z_{e_+} - z_{e_-}), \qquad z \in (\mathbf{R}^{d+1})^{\mathcal{V}'}.$$

We can then evaluate the graph at its red roots by

$$\mathcal{G}(z_{\bullet}) = \int_{(\mathbf{R}^{d+1})^{\mathcal{I}}} K_{\mathcal{G}}(z) \, dz_{\mathcal{I}},$$

where $\mathcal{I} = \mathcal{V} \setminus \mathcal{V}_\bullet$. Now, the crux of our bounds is to pick certain spatial variables and use estimates of the following two kinds, for some $F, G \geq 0$

$$\int_0^t \int_{\mathbf{R}^d} F(x, s) G(x, s) \, dx \, ds \leq \int_0^t \sup_x F(x, s) \int_{\mathbf{R}^d} G(x, s) \, dx \, ds, \qquad (3.12)$$

$$\sup_{x \in \mathbf{R}^d} F(x) G(x) \leq \left(\sup_x F(x) \right) \left(\sup_x G(x) \right). \qquad (3.13)$$

In our setting F, G will be two components of $K_{\mathcal{G}}$, linked to two subgraphs $\mathcal{G}^{(1)}, \mathcal{G}^{(2)} \subseteq \mathcal{G}$, which, appropriately glued together, form $\mathcal{G}$.

Glueing. In pictures, we describe such a glueing as follows:

$$(3.14)$$
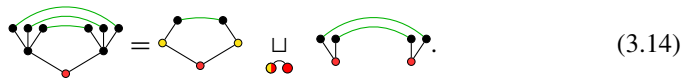
When glueing red nodes onto yellow nodes we will be in a setting in which we use (3.12). As the notation suggests, we also glue red nodes together, and in this case we will use the estimate (3.13). Glueing red nodes is necessary, say in our running example, to decompose

The glueing procedure we just informally suggested is formalised like this.

Definition 3.2 Let $\mathcal{G}^{(1)}, \mathcal{G}^{(2)} \in \mathcal{F}$ and consider $A^{(1)} \subseteq \mathcal{V}_\circ^{(1)} \cup \mathcal{V}_\bullet^{(1)}, A^{(2)} \subseteq \mathcal{V}_\bullet^{(2)}$. Given a bijection $\mathfrak{oo}: A^{(1)} \to A^{(2)}$, we define

$$\mathcal{G} = (\mathcal{V}, \mathcal{E}) = \mathcal{G}^{(1)} \underset{\mathfrak{oo}}{\sqcup} \mathcal{G}^{(2)} = (\mathcal{G}^{(1)} \sqcup \mathcal{G}^{(2)})/\mathfrak{oo},$$

the latter being the equivalence relation generated by $v \sim \mathfrak{oo}(v)$ for all $v \in A^{(1)}$. In this way we also obtain two emebddings $\mathfrak{e}^{(1)}, \mathfrak{e}^{(2)}$ with $\mathfrak{e}^{(i)}: \mathcal{G}^{(i)} \hookrightarrow \mathcal{G}^{(1)} \underset{\mathfrak{oo}}{\sqcup} \mathcal{G}^{(2)}$, which we can combine to a map (which is no longer injective)

$$\mathfrak{e}: \mathcal{G}^{(1)} \sqcup \mathcal{G}^{(2)} \to \mathcal{G}^{(1)} \underset{\mathfrak{oo}}{\sqcup} \mathcal{G}^{(2)}.$$

Finally, the colour assigned to an identified node $v \in \mathcal{V}$, i.e. some v belonging to $\mathfrak{e}A^{(1)}$, is

$$c(v) = \bullet, \quad \text{if } v \in \mathfrak{e}\mathcal{V}_\circ^{(1)}, \qquad c(v) = \bullet, \quad \text{if } v \in \mathfrak{e}\mathcal{V}_\bullet^{(1)}. \qquad (3.15)$$

Here and in what follows we write $\mathfrak{e}A$ short for $\mathfrak{e}(A)$. In general, via this procedure the new graph $\mathcal{G}$ may not belong to the space of forests $\mathcal{F}$. We therefore introduce the following notion of compatibility between the colours and the structure of $\mathcal{F}$.

Definition 3.3 A graph $\mathcal{G} \in \mathcal{F}$ has *compatible* colouring if for any $v \in \mathcal{V}$

$$c(v) = \bullet \qquad \Leftrightarrow v \text{ is a root,}$$
$$c(v) = \circ \qquad \Rightarrow v \text{ is neither a leaf nor a root, i.e. an inner node.}$$

Remark 3.4 Note that the rooted forest structure of $\mathcal{G}$ is uniquely determined by its colouring of edges (black or green) and vertices (black, red, or yellow).

From here on we always work with *compatible* graphs. In particular, we have the following result.

Lemma 3.5 *Given compatible graphs $\mathcal{G}^{(1)}, \mathcal{G}^{(2)} \in \mathcal{F}$ and $\bullet\!\!\bullet$ as in Definition 3.2, $\mathcal{G}^{(1)} \underset{\bullet\!\bullet}{\sqcup} \mathcal{G}^{(2)}$ is a compatible element of $\mathcal{F}$, with the same pairing as $\mathcal{G}^{(1)} \sqcup \mathcal{G}^{(2)}$.*

Proof Since $\mathcal{V}_{\bullet}^{(2)}$ consists of roots and since these are always glued onto either roots or inner nodes by Definition 3.3, $\mathcal{G}^{(1)} \underset{\bullet\!\bullet}{\sqcup} \mathcal{G}^{(2)}$ again has a natural forest structure. The rule (3.15) furthermore guarantees that the compatibility property is preserved. (Note also that no new yellow nodes are created.) □

Valuation. Now we define a valuation $|\mathcal{G}|$ of a compatible graph $\mathcal{G} \in \mathcal{F}$ as follows:

- Consider the kernel $K_{\mathcal{G}}$ defined by the edges of $\mathcal{G}$ and integrate it over all the black variables (i.e. variables associated to black vertices), with all remaining variables fixed.
- Second, integrate over all the spatial components of the yellow variables.
- Third, take the supremum over the temporal components of the yellow variables.
- Finally, take the supremum over the spatial components of red variables.

Clearly, the result of such a valuation will depend only on the red temporal variables $|\mathcal{G}| = |\mathcal{G}|(t_{\bullet})$. In formulae

$$|\mathcal{G}|(t_{\bullet}) = \sup_{x_{\bullet}, t_{\circ}} \int K_{\mathcal{G}}(z) \, \mathrm{d}z_{\bullet} \, \mathrm{d}x_{\circ}.$$

Remark 3.6 This definition is consistent with our evaluation of the graph at a given point in the sense that $\mathcal{G}(z_{\bullet}) \leq |\mathcal{G}|(t_{\bullet})$ if all nodes of $\mathcal{G}$ apart from the red roots are coloured black.

Now, our aim will be to combine the valuation with the glueing of two graphs and obtain an estimate of the sort $|\mathcal{G}^{(1)} \underset{\bullet\!\bullet}{\sqcup} \mathcal{G}^{(2)}| \leq |\mathcal{G}^{(1)}| \cdot |\mathcal{G}^{(2)}|$. Of course, in this form the estimate cannot be true, since the right-hand side has more free (red) variables than the left-hand side; these additional free variables will be integrated over. In addition, in order to obtain a useful bound we have to also take care of the time supremum over $t_{\circ}$: here it will be natural to assume that all yellow nodes are covered by the glueing procedure, and we call such a glueing *onto*. We capture this in the following notation.

Definition 3.7 Consider two compatible graphs $\mathcal{G}^{(1)}, \mathcal{G}^{(2)} \in \mathcal{F}$ and $A^{(1)} \subseteq \mathcal{V}_{\circ}^{(1)} \cup \mathcal{V}_{\bullet}^{(1)}$, $A^{(2)} \subseteq \mathcal{V}_{\bullet}^{(2)}$, together with a bijection $\text{🔴🔴}: A^{(2)} \to A^{(1)}$. If

$$\mathcal{V}_{\circ}^{(1)} \subseteq A^{(1)}, \qquad \mathcal{V}_{\circ}^{(2)} = \emptyset,$$

then we say that 🔴🔴 defines an **onto** glueing of $\mathcal{G}^{(2)}$ onto $\mathcal{G}^{(1)}$. In particular, for $\mathcal{G} = \mathcal{G}^{(1)} \underset{\text{🔴🔴}}{\sqcup} \mathcal{G}^{(2)} = (\mathcal{V}, \mathcal{E})$ it holds that $\mathcal{V}_{\circ} = \emptyset$.

Now the key result for our bounds is the following lemma. Here we introduce the temporal integration bound $\mathfrak{t} = \max_{v \in \mathcal{V}_{\bullet}} t_v$.

Lemma 3.8 *Consider two compatible $\mathcal{G}^{(1)}, \mathcal{G}^{(2)} \in \mathcal{F}$ and an onto glueing 🔴🔴. Then for $\mathcal{G} = \mathcal{G}^{(1)} \underset{\text{🔴🔴}}{\sqcup} \mathcal{G}^{(2)}$ we have*

$$|\mathcal{G}|(t_{\mathcal{V}_{\bullet}}) \leq |\mathcal{G}^{(1)}|(t_{\mathfrak{e}\mathcal{V}_{\bullet}^{(1)}}) \cdot \int_{[0,\mathfrak{t}]^{\mathfrak{e}\mathcal{V}_{\circ}^{(1)}}} |\mathcal{G}^{(2)}|(t_{\mathfrak{e}\mathcal{V}_{\bullet}^{(2)}}) \, dt_{\mathfrak{e}\mathcal{V}_{\circ}^{(1)}}.$$

Remark 3.9 We observe that the temporal variables on which the valuation depends are inherited from the glued graph $\mathcal{G}$, thus determining a pairing among variables of $\mathcal{G}^{(1)}$ and $\mathcal{G}^{(2)}$. This is captured by the map $\mathfrak{e}$. We note in particular that $\mathfrak{e}\mathcal{V}_{\circ}^{(1)} \subseteq \mathfrak{e}\mathcal{V}_{\bullet}^{(2)}$ since the glueing is onto, as well as $\mathfrak{e}\mathcal{V}_{\bullet}^{(1)} \subseteq \mathcal{V}_{\bullet}$. In fact, we can write $\mathfrak{e}\mathcal{V}_{\bullet}^{(2)}$ as the disjoint union

$$\mathfrak{e}\mathcal{V}_{\bullet}^{(2)} = \left(\mathfrak{e}\mathcal{V}_{\circ}^{(1)} \cap \mathfrak{e}\mathcal{V}_{\bullet}^{(2)} \right) \sqcup \left(\mathcal{V}_{\bullet} \cap \mathfrak{e}\mathcal{V}_{\bullet}^{(2)} \right).$$

Remark 3.10 A special case is given by $\mathcal{V}_{\circ}^{(1)} = \mathcal{V}_{\circ}^{(2)} = \emptyset$ in which case one has $\mathcal{V}_{\bullet} = \mathcal{V}_{\bullet}^{(1)} \sqcup \mathcal{V}_{\bullet}^{(1)}$ and $|\mathcal{G}| \leq |\mathcal{G}^{(1)}| \cdot |\mathcal{G}^{(2)}|$.

Proof Since by assumption 🔴🔴 is onto, we have $\mathcal{V}_{\circ} = \emptyset$ and thus

$$|\mathcal{G}|(t_{\mathcal{V}_{\bullet}}) = \sup_{x_{\mathcal{V}_{\bullet}}} \int K_{\mathcal{G}}(z) \, dz_{\mathcal{V}_{\bullet}}.$$

Note first that we can write

$$\mathcal{V}_{\bullet} = \mathfrak{e}\mathcal{V}_{\bullet}^{(1)} \sqcup \mathfrak{e}\mathcal{V}_{\bullet}^{(2)} \sqcup \mathfrak{e}\mathcal{V}_{\circ}^{(1)},$$

by the assumptions on the colouring of $\mathcal{G}^{(1)}, \mathcal{G}^{(2)}$ (in particular, recall that $\mathfrak{e}\mathcal{V}_{\bullet}^{(1)} \subseteq \mathcal{V}_{\bullet}$ because we colour nodes black after glueing). As a consequence, we have

$$\int K_{\mathcal{G}}(z) \, dz_{\mathcal{V}_{\bullet}} = \int \left(\int K_{\mathcal{G}^{(1)}}(z) \, dz_{\mathfrak{e}\mathcal{V}_{\bullet}^{(1)}} \right) \cdot \left(\int K_{\mathcal{G}^{(2)}}(z) \, dz_{\mathfrak{e}\mathcal{V}_{\bullet}^{(2)}} \right) dz_{\mathfrak{e}\mathcal{V}_{\circ}^{(1)}},$$

where we used that $\mathfrak{e}\mathcal{V}_\bullet^{(1)}$ and $\mathfrak{e}\mathcal{V}_\bullet^{(2)}$ do not intersect. It then follows from (3.12) that

$$
\int K_{\mathcal{G}}(z)\,\mathrm{d}z_{\mathcal{V}_\bullet}
$$
$$
= \int_{\mathbf{R}^{\mathfrak{e}\mathcal{V}_\circ^{(1)}}} \int_{(\mathbf{R}^d)^{\mathfrak{e}\mathcal{V}_\circ^{(1)}}} \left(\int K_{\mathcal{G}^{(1)}}(z)\,\mathrm{d}z_{\mathfrak{e}\mathcal{V}_\bullet^{(1)}} \right) \cdot \left(\int K_{\mathcal{G}^{(2)}}(z)\,\mathrm{d}z_{\mathfrak{e}\mathcal{V}_\bullet^{(2)}} \right) \mathrm{d}x_{\mathfrak{e}\mathcal{V}_\circ^{(1)}}\,\mathrm{d}t_{\mathfrak{e}\mathcal{V}_\circ^{(1)}}
$$
$$
\le \int_{\mathbf{R}^{\mathfrak{e}\mathcal{V}_\circ^{(1)}}} \left(\int K_{\mathcal{G}^{(1)}}(z)\,\mathrm{d}z_{\mathfrak{e}\mathcal{V}_\bullet^{(1)}}\,\mathrm{d}x_{\mathfrak{e}\mathcal{V}_\circ^{(1)}} \right) \cdot \left(\sup_{x_{\mathfrak{e}\mathcal{V}_\circ^{(1)}}} \int K_{\mathcal{G}^{(2)}}(z)\,\mathrm{d}z_{\mathfrak{e}\mathcal{V}_\bullet^{(2)}} \right) \mathrm{d}t_{\mathfrak{e}\mathcal{V}_\circ^{(1)}}.
$$

Next, observe that $\mathfrak{e}\mathcal{V}_\circ^{(1)} \subseteq \mathfrak{e}\mathcal{V}_\circ^{(2)}$, and that from the definition of the kernel the time variables are ordered so that $t_v \in [0, t]$ for any $v \in \mathcal{V}$. In particular the previous estimate is bounded by

$$
\int K_{\mathcal{G}}(z)\,\mathrm{d}z_{\mathcal{V}_\bullet}
$$
$$
\le \int_{[0,t]^{\mathfrak{e}\mathcal{V}_\circ^{(1)}}} \left(\int K_{\mathcal{G}^{(1)}}(z)\,\mathrm{d}z_{\mathfrak{e}\mathcal{V}_\bullet^{(1)}}\,\mathrm{d}x_{\mathfrak{e}\mathcal{V}_\circ^{(1)}} \right) \cdot \left(\sup_{x_{\mathfrak{e}\mathcal{V}_\circ^{(2)}}} \int K_{\mathcal{G}^{(2)}}(z)\,\mathrm{d}z_{\mathfrak{e}\mathcal{V}_\bullet^{(2)}} \right) \mathrm{d}t_{\mathfrak{e}\mathcal{V}_\circ^{(1)}}
$$
$$
\le |\mathcal{G}^{(1)}|(t_{\mathfrak{e}\mathcal{V}_\circ^{(1)}}) \int_{[0,t]^{\mathfrak{e}\mathcal{V}_\circ^{(1)}}} |\mathcal{G}^{(2)}|(t_{\mathfrak{e}\mathcal{V}_\circ^{(2)}})\,\mathrm{d}t_{\mathfrak{e}\mathcal{V}_\circ^{(1)}},
$$

as required. □

Running example. Finally, we can use Lemma 3.8 to conclude the estimates for (3.10). It is important to remark that in all cases of our interest the supremum over $t_{\mathcal{V}_\circ^{(1)}}$ in the valuation of $\mathcal{G}^{(1)}$ will be superfluous, by an application of the semigroup property. This can be swiftly seen by proceeding in the calculation for our running example. Since no confusion can occur, we now write t instead of $\mathfrak{t}$, observing that the latter does indeed coincide in our case of interest with the temporal variable of the only red node present in the original graph. By (3.14) and Lemma 3.8, we obtain

Regarding the first factor, it follows from the semigroup property and the fact that we integrate over all yellow spatial variables that

$$
\quad (3.16)
$$

for some $c > 0$. Hence, using Remark 3.10, we have

$$\left| \begin{array}{c} \text{(graph)} \end{array} \right| \leq \mathfrak{s}_\varepsilon(t) \int_{[0,t]^2} \left| \begin{array}{c} \text{(graph)} \end{array} \right| (t_\bullet) \cdot \left| \begin{array}{c} \text{(graph)} \end{array} \right| (t_\bullet) \, dt_\bullet$$

$$\leq \mathfrak{s}_\varepsilon(t) \int_{[0,t]^2} \mathfrak{s}_\varepsilon(t_{\bullet,1}) \mathfrak{s}_\varepsilon(t_{\bullet,2}) \, dt_\bullet$$

$$\leq (t \vee \varepsilon^2)^{-\frac{d}{2}} (ce^t \varepsilon^{\frac{d}{2}-\alpha})^6 \left(\int_0^t (s \vee \varepsilon^2)^{-\frac{d}{2}} \, ds \right)^2 \stackrel{\text{def}}{=} B_\varepsilon(t) .$$

The second graph on the right-hand side of (3.10) can be bounded by $B_\varepsilon(t)$ in an analogous way. We recall that every smaller tree is evaluated at the temporal variable it has inherited from the original large tree (so the last two trees have the same contribution, but they are evaluated at different time variables). We omit writing the dependence on such variables for convenience. We conclude that

$$\mathbf{E} \begin{array}{c} \text{(graph)} \end{array}^2 (t, x) \leq 15 B_\varepsilon(t). \tag{3.17}$$

Here, the number $15 = 6 + 9$ is the total number of ways of pairing 6 green nodes. In addition, we observe that $B_\varepsilon(t)$ can be rewritten, by a change of variables, as

$$B_\varepsilon(t) = (t \vee \varepsilon^2)^{-\frac{d}{2}} (ce^t \varepsilon^{\frac{d}{2}-\alpha})^2 \left(ce^t \varepsilon^{1-\alpha} \left(\int_0^{t\varepsilon^{-2}} (s \vee 1)^{-\frac{d}{2}} \, ds \right)^{\frac{1}{2}} \right)^4 ,$$

where $4 = 2(\ell(\tau) - 1)$, for the trees τ that we took under consideration. We now claim that an analogous bound holds for every tree τ. More precisely, we have the following result.

Proposition 3.11 *For every τ in $\mathcal{T}_3$ and $q \in [2, \infty)$,*

$$\|X_\varepsilon^\tau(t, x)\|_{L^q(\Omega)} \lesssim (t \vee \varepsilon^2)^{-\frac{d}{4}} e^t \varepsilon^{\frac{d}{2}-\alpha} \left(e^t \varepsilon^{1-\alpha} \Gamma_{d/2}^{1/2}(t\varepsilon^{-2}) \right)^{\ell(\tau)-1} \tag{3.18}$$

and

$$\|\nabla X_\varepsilon^\tau(t, x)\|_{L^q(\Omega)} \lesssim (t \vee \varepsilon^2)^{-\frac{d}{4}-\frac{1}{2}} e^t \varepsilon^{\frac{d}{2}-\alpha} \left(e^t \varepsilon^{1-\alpha} \Gamma_{d/2}^{1/2}(t\varepsilon^{-2}) \right)^{\ell(\tau)-1} , \tag{3.19}$$

where for each $a > 0$, $\Gamma_a(t) = \int_0^t (s \vee 1)^{-a} \, ds$.

Remark 3.12 We observe that for $t \geq \varepsilon^2$ we have

$$1 \leq \Gamma_{\frac{d}{2}}(t\varepsilon^{-2}) \leq 1 + \begin{cases} \log(t\varepsilon^{-2}) & \text{if } d = 2, \\ (1 - \frac{2}{d})^{-1} & \text{if } d > 2, \end{cases}$$

which leads to the time scale $t_1(\varepsilon)$ (in particular to the constraint $\bar{\alpha} < 1$), since the estimates in Proposition 3.11 improve as $\ell(\tau)$ increases only for times $t \leq t_1(\varepsilon)$.

Proof By hypercontractivity, it suffices to show (3.18) for $q = 2$ since each X_ε^τ belongs to a finite Wiener chaos. Let τ be a tree in $\mathcal{T}_3$ and $\bar\tau$ be an identical copy of τ. Let γ be a way of pairing the leaves of $[\tau, \bar\tau]$ such that each leaf is paired with exactly one other leaf. The coloured graph $[\tau, \bar\tau]_\gamma$ is obtained by connecting the leaves of $[\tau, \bar\tau]$ with green edges according to the pairing rule γ and connecting the two roots to a red node, so that

$$\mathbf{E}|X_\varepsilon^\tau(t, x)|^2 = \sum_\gamma [\tau, \bar\tau]_\gamma(z_\bullet) \le \sum_\gamma |[\tau, \bar\tau]_\gamma|(t_\bullet).$$

As usual, the value associated to the single red node is fixed to (t, x) and since no confusion can occur we write t instead of $t_\bullet$. Hence, the estimate (3.18) for $q = 2$ is obtained once we show that for any pairing rule γ,

$$|[\tau, \bar\tau]_\gamma|(t) \le (t \vee \varepsilon^2)^{-\frac{d}{2}} (ce^t \varepsilon^{\frac{d}{2}-\alpha})^{2\ell(\tau)} \left(\int_0^t (s \vee \varepsilon^2)^{-\frac{d}{2}} \, ds \right)^{\ell(\tau)-1}, \quad (3.20)$$

where $|\cdot|$ is the valuation introduced above. Let us write $\mathcal{G} = (\mathcal{V}, \mathcal{E}) \in \mathcal{F}$ for the compatible tree $[\tau, \bar\tau]_\gamma$.

To prove (3.20) we introduce an algorithm that allows us to iterate the type of bounds we described previously. We start by constructing a succession $\{p_i\}_{i=0}^{N-1}$ of self-avoiding paths $p_i = (\mathcal{V}_{p_i}, \mathcal{E}_{p_i}) \subseteq (\mathcal{V}, \mathcal{E})$ that cover the entire tree $\mathcal{G}$.

By self-avoiding path we mean a path with endpoints $p_-, p_+ \in \mathcal{V}$ (with possibly $p_- = p_+$) such that all points, apart from possibly p_-, p_+ appear exactly once in an edge (in particular, all edges are distinct). In addition we ask that the endpoints are minimal points with respect to the ordering inherited by the forest structure of $\mathcal{G}$.

The paths $\{p_i\}_{i=0}^{N-1}$ will cover the entire tree $\mathcal{G}$ in the sense that we will iteratively define compatible graphs $\mathcal{G}^{(i)} \in \mathcal{F}_3$ for $i \in \{0, \ldots, N\}$ by onto glueings, satisfying

$$\mathcal{G}^{(N)} = \emptyset, \quad \mathcal{G}^{(i-1)} = p_{i-1} \underset{\bullet\bullet}{\sqcup} \mathcal{G}^{(i)}, \quad \mathcal{G}^{(0)} = \mathcal{G}.$$

In particular it is useful to note that we obtain the following chain of embeddings

$$\mathcal{G}^{(N)} \xrightarrow{\mathfrak{e}_{N-1}} \mathcal{G}^{(N-1)} \xrightarrow{\mathfrak{e}_{N-2}} \cdots\cdots \xrightarrow{\mathfrak{e}_1} \mathcal{G}^{(1)} \xrightarrow{\mathfrak{e}_0} \mathcal{G}^{(0)}$$
$$\mathfrak{e}_{N-1}\uparrow \qquad\qquad \mathfrak{e}_i\uparrow \qquad\quad \mathfrak{e}_1\uparrow \qquad \mathfrak{e}_0\uparrow \quad,$$
$$p_{N-1} \qquad\qquad p_i \qquad\quad p_1 \qquad p_0$$

where each $\mathfrak{e}_i$ is the map of Definition 3.2. Thus we obtain an overall map

$$\mathfrak{e}: p_0 \sqcup \cdots \sqcup p_{N-1} \to \mathcal{G}, \quad (3.21)$$

defined by composing $\mathfrak{e} = \mathfrak{e}_0 \circ \cdots \circ \mathfrak{e}_i$ on p_i. The way the paths p_i are chosen is described by the following procedure, starting from the graph $\mathcal{G}^{(0)} = \mathcal{G}$ introduced above.

1. **Construction of the path.** Assume that for some $i \in \mathbf{N} \cup \{0\}$ we are given a compatible graph $\mathcal{G}^{(i)} = (\mathcal{V}^{(i)}, \mathcal{E}^{(i)}) \in \mathcal{F}_3$, with $c(v) \in \{\bullet, \circ\}$ for all $v \in \mathcal{V}^{(i)}$ (note that this is the case for $\mathcal{G}^{(0)}$, since we have one red root and no node is coloured yellow). We define p_i by choosing as a starting point any $v \in \mathcal{V}^{(i)}$ such that $c(v) = \bullet$, namely any root of $\mathcal{G}^{(i)}$.

   By assumption, since $\mathcal{G}^{(i)}$ is an element of $\mathcal{F}_3$, we can follow a path up through any one of the two maximal trees rooted at v and let us call the sub-tree we choose $\tau(v)$. Then, because of oddity, we can choose to arrive at a leaf such that the pairing γ leads to another leaf, belonging to some $\overline{\tau}(v) \neq \tau(v)$. We then continue the path by descending the maximal subtree this leaf belongs to, until we reach its red root (this may be again v, since v can be the root of up to two trees). We colour all the nodes of p_i yellow, apart from its endpoints and leaves, which are left red and black respectively.

2. **Removal.** We then define the compatible graph $\mathcal{G}^{(i+1)} = (\mathcal{V}^{(i+1)}, \mathcal{E}^{(i)} \setminus \mathcal{E}_{p_i}) \in \mathcal{F}$ with the colourings inherited by $\mathcal{G}^{(i)}$. Here $\mathcal{V}^{(i+1)}$ is obtained from $\mathcal{V}^{(i)}$ by removing all singletons (meaning the two leafs connected by the only green edge p_i has crossed and possibly the endpoints of p_i, if any of them was the root of just one tree, or p_i was a cycle). We finally colour $c(v) = \bullet$ for $v \in \mathcal{V}^{(i+1)} \cap \mathcal{V}_{p_i}$. In particular (3.11) still holds true for $\mathcal{G}^{(i+1)}$, so that $\mathcal{G}^{(i+1)}$ lies in $\mathcal{F}_3$. Moreover, we obtain that $\mathcal{G}^{(i)}$ is the *onto* glueing of $\mathcal{G}^{(i+1)}$ onto p_i

$$\mathcal{G}^{(i)} = p_i \underset{\bullet\bullet}{\sqcup} \mathcal{G}^{(i+1)},$$

   with the map $\bullet\bullet$ defined by inverting the removal we just described.

3. **Conclusion.** We can now repeat steps 1 and 2 for N times until $\mathcal{G}^{(N)} = \emptyset$ (if not, there must still be a red root available and we can still follow the previous algorithm). Since at every step we are removing exactly one green edge, we see that this algorithm terminates in $N = \ell(\tau)$ steps.

We are left with a collection of N paths $\{p_i\}_{i=0}^{\ell(\tau)-1}$ such that

1. $\mathcal{E}$ is the disjoint union $\mathcal{E} = \bigsqcup_{i=1}^{\ell(\tau)-1} \mathcal{E}_{p_i}$.
2. Every $v \in \mathcal{V}$ which is not a leaf appears in exactly thee paths, twice as a root (coloured red) and once as an inner node (coloured yellow). Instead, every leaf appears in exactly one path, coloured black.

We observe that by the semigroup property, every path p_i yields a contribution

$$|p_i|(t_\bullet) \leq \begin{cases} \sqrt{\mathfrak{s}_\varepsilon(t_{p_{i,+}})\mathfrak{s}_\varepsilon(t_{p_{i,-}})} & \text{if } p_{i,+} \neq p_{i,-}, \\ \mathfrak{s}_\varepsilon(t_{p_i}) & \text{if } p_{i,+} = p_{i,-} = p_i, \end{cases} \tag{3.22}$$

where $\mathfrak{s}_\varepsilon$ is defined in (3.16). To conclude, we may now use Lemma 3.8 to obtain

$$\mathcal{G}(t, x) \leq \mathfrak{s}_\varepsilon(t) \int_{[0,t]^{\varepsilon \mathcal{V}_{p_0,\bullet}}} |\mathcal{G}^{(1)}|(t_{\varepsilon \mathcal{V}_\bullet^{(1)}}) \, \mathrm{d}t_{\varepsilon \mathcal{V}_{p_0,\bullet}},$$

where the map $\mathfrak{e}$ is defined in (3.21). And similarly, for $i \in \{1, \ldots, N-1\}$, using the definition of the map $\mathfrak{e}$, we find

$$|\mathcal{G}^{(i)}|(t_{\mathfrak{e}\mathcal{V}^{(i)}_{\bullet}}) \leq |p_i|(t_{\mathfrak{e}\mathcal{V}_{p_i,\circ}}) \int_{[0,t]^{\mathfrak{e}\mathcal{V}_{p_i,\circ}}} |\mathcal{G}^{(i+1)}|(t_{\mathfrak{e}\mathcal{V}^{(i+1)}_{\bullet}}) \, \mathrm{d}t_{\mathfrak{e}\mathcal{V}_{p_i,\circ}},$$

so that overall

$$\mathcal{G}(t,x) \leq \mathfrak{s}_\varepsilon(t) \int_{[0,t]^{\overline{\mathcal{I}}}} \prod_{i=1}^{\ell(\tau)-1} |p_i|(t_{\mathfrak{e}\mathcal{V}_{p_i,\bullet}}) \, \mathrm{d}t_{\overline{\mathcal{I}}},$$

where $\overline{\mathcal{I}} = \sqcup_{i=0}^{N-1} \mathfrak{e}\mathcal{V}_{p_i,\circ} \subseteq \mathcal{V}$ is the set of inner nodes that are not leaves, in view of the considerations above. Since in addition by our observations every node in $\overline{\mathcal{I}}$ appears exactly twice as a root of some path (either a cycle or two different paths), by (3.22) we finally obtain

$$\mathcal{G}(t,x) \leq \mathfrak{s}_\varepsilon(t) \int_{[0,t]^{\overline{\mathcal{I}}}} \mathfrak{s}_\varepsilon(t_{\overline{\mathcal{I}}}) \, \mathrm{d}t_{\overline{\mathcal{I}}} \leq \mathfrak{s}_\varepsilon(t) \left(\int_0^t \mathfrak{s}_\varepsilon(s) \, \mathrm{d}s \right)^{|\overline{\mathcal{I}}|}$$

$$\leq (t \vee \varepsilon^2)^{-\frac{d}{2}} (c e^t \varepsilon^{\frac{d}{2}-\alpha})^{2\ell(\tau)} \left(\int_0^t (s \vee \varepsilon^2)^{-\frac{d}{2}} \, \mathrm{d}s \right)^{\ell(\tau)-1},$$

which is (3.20). Here we have used that $|\overline{\mathcal{I}}| = 2i(\tau)$ since we have two copies of the same tree, and $2i(\tau) = \ell(\tau) - 1$.

As for (3.19), we can follow the same argument as above, to write:

$$\mathbf{E}|\nabla X_\varepsilon^\tau(t,x)| = \sum_\tau [\nabla \tau, \nabla \overline{\tau}]_\gamma(t,x),$$

where $\nabla \tau$ indicates the same convolution integral associated to the tree τ as above, only with the kernel $-e^t \nabla p_t(x)$ used in the edge connecting to the root. Graphically, if τ is built starting from the trees τ_1, τ_2, τ_3 we have
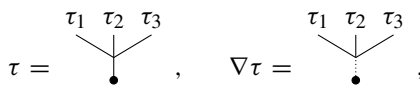
where the dotted line indicates convolution against the kernel $-e^t \nabla p_t(x)$. In particular, we can decompose the graph $[\nabla \tau, \nabla \overline{\tau}]_\gamma$ into the same set of paths $\{p_i\}_{i=0}^{N-1}$ as for $[\tau, \overline{\tau}]_\gamma$. The only difference is that now the path p_0 contains two dotted lines coming into the root: all other paths remain unchanged. By the semigroup property the contribution of the path p_0 is the same as that of

(t,x)

Hence, using the bound (for some $c > 0$)

$$\|\nabla p_t^\varepsilon\|_{L^2(\mathbf{R}^d)} \le (\|\nabla p_t\|_{L^2(\mathbf{R}^d)}\|\varphi^\varepsilon\|_{L^1(\mathbf{R}^d)}) \wedge (\|p_t\|_{L^1(\mathbf{R}^d)}\|\nabla \varphi^\varepsilon\|_{L^2(\mathbf{R}^d)})$$
$$\le c(t \vee \varepsilon^2)^{-\frac{d}{4}-\frac{1}{2}},$$

we find that the contribution of p_0 is bounded from above by

$$(c\varepsilon^{\frac{d}{2}-\alpha})^2 e^{2t}(t \vee \varepsilon^2)^{-\frac{d}{2}-1},$$

which together with the previous calculations yields (3.19). $\qquad\square$

We can extend the previous result to longer timescales as follows.

Proposition 3.13 *For every $q \in [2, \infty)$, τ in $\mathcal{T}_3$ and every $t > t_1(\varepsilon)$, the following estimates hold true*

$$\|P^1_{t-t_1} X^\tau_\varepsilon(t_1, x)\|_{L^q(\Omega)} \lesssim (t \vee \varepsilon^2)^{-\frac{d}{4}} e^t \varepsilon^{\frac{d}{2}-\alpha}\left(e^{t_1}\varepsilon^{1-\alpha}\Gamma^{1/2}_{d/2}(t_1\varepsilon^{-2})\right)^{\ell(\tau)-1},$$
$$(3.23)$$

$$\|\nabla P^1_{t-t_1} X^\tau_\varepsilon(t_1, x)\|_{L^q(\Omega)} \lesssim (t \vee \varepsilon^2)^{-\frac{d}{4}-\frac{1}{2}} e^t \varepsilon^{\frac{d}{2}-\alpha}\left(e^{t_1}\varepsilon^{1-\alpha}\Gamma^{1/2}_{d/2}(t_1\varepsilon^{-2})\right)^{\ell(\tau)-1}.$$
$$(3.24)$$

Proof By hypercontractivity it suffices to prove the result for $q = 2$. We can change the graphical notation we introduced previously by adding blue inner nodes • associated to a time $t_1 > 0$ to indicate integration against the measure $-\delta_{t_1}(s) \, ds \, dy$: once more, the minus sign is a matter of convention. It is natural though since our lines represent integration against $-P^1$ as a consequence of the minus sign appearing in (3.4). With this notation we have

$$P^1_{t-t_1} X^\bullet_\varepsilon(t_1, x) = \underset{(t, x)}{\bullet}\, {\scriptstyle t_1}\,, \quad P^1_{t-t_1} X^\tau_\varepsilon(t_1, x) = \overset{\tau_1 \ \tau_2 \ \tau_3}{\underset{(t, x)}{\bullet}}\,{\scriptstyle t_1}\,,$$

for $\tau = [\tau_1, \tau_2, \tau_3]$. For convenience we denote by $P^1_{t-t_1}\tau_{t_1}$ the trees obtained in this manner. Now we can follow the proof of Proposition 3.11 to write

$$\mathbf{E}|P^1_{t-t_1} X^\tau_\varepsilon(t_1, x)|^2 = \sum_\gamma [P^1_{t-t_1}\tau_{t_1}, P^1_{t-t_1}\overline{\tau}_{t_1}]_\gamma(x),$$

where $\overline{\tau}$ is an identical copy of τ. The possible paring rules γ for the tree $P^1_{t-t_1}\tau_{t_1}$ (with an identical copy of itself) are the same as for the tree τ (since we only added a node to the edge entering the root), and also the set of paths $\{p_i\}_{i=0}^{N-1}$ in which the tree is decomposed remains unchanged, up to adding the two nodes corresponding to

t_1 to the path p_0. By the semigroup property, the contribution of the path p_0 is then given by

$$(ce^t \varepsilon^{\frac{d}{2}-\alpha})^2 (t \vee \varepsilon^2)^{-\frac{d}{2}},$$

while the contribution of all other paths remains unchanged, with the additional constraint that all temporal variables (apart from the red root) are constrained to the interval $[0, t_1]$.

Hence, following the calculations of the proof of Proposition 3.11 we can bound

$$[P^1_{t-t_1}\tau_{t_1}, P^1_{t-t_1}\overline{\tau}]_\gamma \leq (ce^t \varepsilon^{\frac{d}{2}-\alpha})^2 (t \vee \varepsilon^2)^{-\frac{d}{2}} \left(\int_0^{t_1} \mathfrak{s}_\varepsilon(s) \, ds \right)^{2i(\tau)}$$

$$\lesssim (e^t \varepsilon^{\frac{d}{2}-\alpha})^2 (t \vee \varepsilon^2)^{-\frac{d}{2}} \left(e^{t_1} \varepsilon^{1-\alpha} \Gamma^{1/2}_{d/2}(t_1 \varepsilon^{-2}) \right)^{\ell(\tau)-1}.$$

This completes the proof of (3.23). For (3.24) one can follow the same arguments in combination with the estimates in the proof of Proposition 3.11. □

Next we can compare the Wild expansion and the original solution via a comparison principle.

Lemma 3.14 *For every* $\varepsilon \in (0,1)$, $N \geq 1$ *and* $(t,x) \in \mathbf{R}_+ \times \mathbf{R}^d$, *let* u^N_ε *and* R^N *be as in* (3.6), (3.7). *Then*

$$|u_\varepsilon - u^N_\varepsilon|(t,x) \leq P^1 * |R^N_\varepsilon|(t,x). \tag{3.25}$$

Proof From (1.2) and (3.6), it follows that the remainder $v_\varepsilon = u_\varepsilon - u^N_\varepsilon$ satisfies the following equation

$$(\partial_t - \Delta - 1)v_\varepsilon = -\left(v_\varepsilon^2 + 3u^N_\varepsilon v_\varepsilon + 3(u^N_\varepsilon)^2 \right) v_\varepsilon - R^N_\varepsilon, \quad v_\varepsilon(0,\cdot) = 0. \tag{3.26}$$

The estimate (3.25) is a direct consequence of Lemma 3.1 and the fact that $v_\varepsilon^2 + 3u^N_\varepsilon v_\varepsilon + 3(u^N_\varepsilon)^2$ is non-negative. □

Eventually, if we combine all the results we have obtained so far, we obtain an estimate on the distance of the Wild expansion from the original solution until times of order $(1-\alpha)\log \varepsilon^{-1}$.

Proposition 3.15 *For every* $t \leq (1-\alpha)\log \varepsilon^{-1}$, $q \in [2, \infty)$ *and* $N \in \mathbf{N}$

$$\|u_\varepsilon(t,x) - u^N_\varepsilon(t,x)\|_{L^q(\Omega)} \lesssim \varepsilon^{2-3\alpha} e^{3t} \left(e^t \varepsilon^{1-\alpha} \Gamma^{3/2}_{d/2}(t\varepsilon^{-2}) \right)^{N-1} \tag{3.27}$$

and

$$\|u_\varepsilon(t,x)\|_{L^q(\Omega)} \lesssim (t \vee \varepsilon^2)^{-\frac{d}{4}} e^t \varepsilon^{\frac{d}{2}-\alpha} \Gamma^{N/2}_{d/2}(t\varepsilon^{-2})$$

$$+ \varepsilon^{2-3\alpha} e^{3t} \left(e^t \varepsilon^{1-\alpha} \Gamma^{3/2}_{d/2}(t\varepsilon^{-2}) \right)^{N-1}. \tag{3.28}$$

In addition, for every $c > 1$, uniformly over $t_1(\varepsilon) < t < c \log \varepsilon^{-1}$

$$\|P^1_{t-t_1(\varepsilon)} u_\varepsilon(t_1(\varepsilon), y)\|_{L^q(\Omega)} \lesssim t^{-\frac{d}{4}} e^t \varepsilon^{\frac{d}{2}-\alpha}. \tag{3.29}$$

Proof From the definition of R^N_ε and the moment estimate in Proposition 3.11, we see that for every $q \geq 2$,

$$\|R^N_\varepsilon(t, x)\|_{L^q(\Omega)} \lesssim (t \vee \varepsilon^2)^{-3\frac{d}{4}} (e^t \varepsilon^{\frac{d}{2}-\alpha})^3 \sum \left(e^t \varepsilon^{1-\alpha} \Gamma^{1/2}_{d/2}(t\varepsilon^{-2}) \right)^{\ell(\tau_1)+\ell(\tau_2)+\ell(\tau_3)-3},$$

where the sum is taken over all trees τ_1, τ_2, τ_3 in $\mathcal{T}^N_3$ such that $\tau = [\tau_1, \tau_2, \tau_3] \notin \mathcal{T}^N_3$. Since $\ell(\tau) = \ell(\tau_1) + \ell(\tau_2) + \ell(\tau_3) \in [N+2, 3N]$ and t satisfies $e^t \varepsilon^{1-\alpha} \leq 1$, each term in the above sum is at most $(e^t \varepsilon^{1-\alpha})^{N-1} \Gamma_{d/2}(t\varepsilon^{-2})^{\frac{3}{2}(N-1)}$. This yields

$$\|R^N_\varepsilon(t, x)\|_{L^q(\Omega)} \lesssim (t \vee \varepsilon^2)^{-3\frac{d}{4}} (e^t \varepsilon^{\frac{d}{2}-\alpha})^3 (e^t \varepsilon^{1-\alpha})^{N-1} \Gamma_{d/2}(t\varepsilon^{-2})^{\frac{3}{2}(N-1)},$$

which, when combined with Lemma 3.14 gives

$$\|u_\varepsilon(t, x) - u^N_\varepsilon(t, x)\|_{L^q(\Omega)} \lesssim (e^t \varepsilon^{\frac{d}{2}-\alpha})^3 (e^t \varepsilon^{1-\alpha} \Gamma^{3/2}_{d/2})^{N-1} \int_0^t (s \vee \varepsilon^2)^{-\frac{3d}{4}} \, ds.$$

To derive (3.27), we observe that $\int_0^t (s \vee \varepsilon^2)^{-\frac{3d}{4}} \, ds \lesssim \varepsilon^{2-\frac{3d}{2}}$ for $d \geq 2$. The estimate (3.28) is a direct consequence of (3.27) and Proposition 3.11.

Let us now show (3.29). By the triangle inequality, we have

$$\|P^1_{t-t_1} u_\varepsilon(t_1)\|_{L^q(\Omega)} \leq P^1_{t-t_1} \|(u_\varepsilon - u^N_\varepsilon)(t_1)\|_{L^q(\Omega)} + \|P^1_{t-t_1} u^N_\varepsilon(t_1)\|_{L^q(\Omega)}.$$

The first term is estimated using (3.27),

$$P^1_{t-t_1} \|(u_\varepsilon - u^N_\varepsilon)(t_1)\|_{L^q(\Omega)}(x) \lesssim e^{t-t_1} \varepsilon^{2-3\alpha} e^{3t_1} (e^{t_1} \varepsilon^{1-\alpha} \Gamma^{3/2}_{d/2}(t_1 \varepsilon^{-2}))^{N-1}.$$

Writing the second term as $\sum_{\tau \in \mathcal{T}^N_3} P^1_{t-t_1} X^\tau_\varepsilon(t_1, x)$ and applying Proposition 3.13, we obtain

$$\|P^1_{t-t_1} u^N_\varepsilon(t_1, x)\|_{L^q} \lesssim (t \vee \varepsilon^2)^{-\frac{d}{4}} e^t \varepsilon^{\frac{d}{2}-\alpha}.$$

In the previous estimate, we have used the fact that $e^{t_1} \varepsilon^{1-\alpha} \Gamma^{1/2}_{d/2}(t_1 \varepsilon^{-2}) \leq 1$ for sufficiently small ε. Combining these estimates yields

$$\|P^1_{t-t_1} u_\varepsilon(t_1, y)\|_q \lesssim t^{-\frac{d}{4}} e^t \varepsilon^{\frac{d}{2}-\alpha} + e^t \varepsilon^{2-2\bar\alpha-\alpha} \left(\varepsilon^{1-\bar\alpha} \Gamma^{3/2}_{d/2}(t_1 \varepsilon^{-2}) \right)^{N-1}.$$

for every $t \geq t_1$. By choosing N sufficiently large, the previous inequality implies (3.29). □

4 Front formation

4.1 Tightness criteria

In this subsection we recall a tightness criterion that is useful in the upcoming discussion. Observe that a subset $\mathcal{F}$ of $C_{loc}(\mathbf{R}^d)$ is compact if and only if for every compact set $K \subset \mathbf{R}^d$, the projection $\mathcal{F}_K$ of $\mathcal{F}$ onto $C(K)$ is compact. In particular the following is a consequence of the Arzelà–Ascoli theorem and of the Morrey–Sobolev inequality.

Lemma 4.1 *Let $\{X_\varepsilon\}_{\varepsilon \in (0,1)}$ be a family of stochastic processes with sample paths in $C_{loc}(\mathbf{R}^d)$. Suppose that for every integer $n \geq 1$, there exists $q > d$ such that*

$$\sup_\varepsilon \sup_{|z| \leq n} \mathbf{E}[|X_\varepsilon(z)|^q + |\nabla X_\varepsilon(z)|^q] < \infty. \tag{4.1}$$

Then the family of the laws of X_ε on $C_{loc}(\mathbf{R}^d)$ is tight. In addition, if for every $n \geq 1$, there exists a $q > d$ such that

$$\limsup_{\varepsilon \downarrow 0} \sup_{|z| \leq n} \mathbf{E}[|X_\varepsilon(z)|^q + |\nabla X_\varepsilon(z)|^q] = 0,$$

then X_ε converges to 0 in probability with respect to the topology of $C_{loc}(\mathbf{R}^d)$.

In particular, we deduce the following criterion.

Lemma 4.2 *Let $\{g_\varepsilon\}_{\varepsilon \in (0,1)}$ be a family of random fields on $\mathbf{R}^d$. Let $t(\varepsilon)$ and $L(\varepsilon)$ be positive numbers such that $\lim_{\varepsilon \to 0} t(\varepsilon) = \infty$. Assume that there exists a $q > d + 1$ such that*

$$\sup_\varepsilon \left(1 + \frac{L(\varepsilon)}{\sqrt{t(\varepsilon)}}\right) e^{t(\varepsilon)} \sup_{y \in \mathbf{R}^d} \|g_\varepsilon(y)\|_{L^q(\Omega)} < \infty.$$

Then the family $\{P^1_{t(\varepsilon)+t} g_\varepsilon(L(\varepsilon)x) : (t, x) \in \mathbf{R} \times \mathbf{R}^d\}_\varepsilon$ is tight on $C_{loc}(\mathbf{R} \times \mathbf{R}^d)$. In addition, if

$$\limsup_{\varepsilon \to 0} \left(1 + \frac{L(\varepsilon)}{\sqrt{t(\varepsilon)}}\right) e^{t(\varepsilon)} \sup_{y \in \mathbf{R}^d} \|g_\varepsilon(y)\|_{L^q(\Omega)} = 0.$$

then $\{P^1_{t(\varepsilon)+t} g_\varepsilon(L(\varepsilon)x) : (t, x) \in \mathbf{R} \times \mathbf{R}^d\}$ converges in probability to 0 in $C_{loc}(\mathbf{R} \times \mathbf{R}^d)$.

Proof This is a consequence of the following estimates:

$$\|P^1_{t(\varepsilon)+t} g_\varepsilon(L(\varepsilon)x)\|_q \leq e^{t(\varepsilon)+t} \sup_y \|g_\varepsilon(y)\|_q,$$

$$\|\nabla P^1_{t(\varepsilon)+t} g_\varepsilon(L(\varepsilon)x)\|_q \lesssim L(\varepsilon)(t(\varepsilon)+t)^{-1/2} e^{t(\varepsilon)+t} \sup_y \|g_\varepsilon(y)\|_q,$$

$$\|\partial_t P^1_{t(\varepsilon)+t} g_\varepsilon(L(\varepsilon)x)\|_q \lesssim (1 + (t(\varepsilon)+t)^{-1}) e^{t(\varepsilon)+t} \sup_y \|g_\varepsilon(y)\|_q,$$

in conjunction with Lemma 4.1. $\qquad\square$

4.2 Weak convergence

Now recall $t_2(\varepsilon)$ from (3.1) and define

$$t_2^\kappa(\varepsilon) \stackrel{\text{def}}{=} \left(\frac{d}{2} - \alpha\right) \log \varepsilon^{-1} - \left(\kappa + \frac{1}{2}\right) \log \log \varepsilon^{-1}.$$

Then we can use the tightness criteria above to describe the behaviour of u_ε up to times of order $t_2^\kappa(\varepsilon)$.

Lemma 4.3 *For every $N > d/(2 - 2\bar{\alpha}) - 1$ and every $\kappa \geq 0$ the process*

$$(t, x) \mapsto \left(P^1_{t_\star^\kappa(\varepsilon) + t - t_2^\kappa(\varepsilon)} |u_\varepsilon(t_2^\kappa(\varepsilon), \cdot) - P^1_{t_2^\kappa(\varepsilon) - t_1(\varepsilon)} u_\varepsilon^N(t_1(\varepsilon), \cdot)|\right)(x L_\varepsilon),$$

converges to 0 in probability on $C_{loc}(\mathbf{R} \times \mathbf{R}^d)$.

Proof For each $t \geq t_1$, we define

$$S_\varepsilon^N(t, x) = u_\varepsilon(t, x) - P^1_{t - t_1} u_\varepsilon^N(t_1, x).$$

Since $h_\varepsilon(t, x) \stackrel{\text{def}}{=} P^1_{t - t_1} u_\varepsilon^N(t_1, x)$ solves $(\partial_t - \Delta - 1)h = 0$ with initial condition (at time t_1) $h(t_1, x) = u_\varepsilon^N(t_1, x)$, the difference $S_\varepsilon^N = u_\varepsilon - h_\varepsilon$ satisfies

$$(\partial_t - \Delta - 1)S_\varepsilon^N = -\left((S_\varepsilon^N)^2 + 3(S_\varepsilon^N)h_\varepsilon + 3h_\varepsilon^2\right) S_\varepsilon^N - h_\varepsilon^3 \quad \text{for} \quad t > t_1$$

with initial condition $S_\varepsilon^N(t_1, x) = u_\varepsilon(t_1, x) - u_\varepsilon^N(t_1, x)$. An application of Lemma 3.1 gives

$$|S_\varepsilon^N(t_2^\kappa, x)| \leq \int_{t_1}^{t_2^\kappa} \int_{\mathbf{R}^d} p^1_{t_2^\kappa - s}(x - y)|P^1_{s - t_1} u_\varepsilon^N(t_1, y)|^3 \, dy \, ds + P^1_{t_2^\kappa - t_1} |u_\varepsilon - u_\varepsilon^N|(t_1, x).$$

Let $q > d + 1$ be fixed. Applying Proposition 3.15, we have

$$\sup_y \|S_\varepsilon^N(t_2^\kappa, y)\|_q \lesssim \int_{t_1}^{t_2^\kappa} e^{t_2^\kappa - s} \left(e^s s^{-\frac{d}{4}} \varepsilon^{\frac{d}{2} - \alpha}\right)^3 \, ds + e^{t_2^\kappa - t_1} \varepsilon^{2 - 3\bar{\alpha}} \left(\varepsilon^{1 - \bar{\alpha}} \Gamma_{d/2}^{3/2}(t_1 \varepsilon^{-2})\right)^{N-1}$$

and hence

$$e^{t_\star^\kappa - t_2^\kappa} \sup_y \|S_\varepsilon^N(t_2^\kappa, y)\|_q \lesssim e^{t_\star^\kappa - t_2^\kappa} \left(e^{t_2^\kappa} \varepsilon^{\frac{d}{2} - \alpha} t_1^{-\frac{d}{4}}\right)^3 (t_2^\kappa - t_1)$$

$$+ e^{t_\star^\kappa - t_1} \varepsilon^{2 - 3\bar{\alpha}} \left(\varepsilon^{1 - \bar{\alpha}} \Gamma_{d/2}^{3/2}(t_1 \varepsilon^{-2})\right)^{N-1}.$$

From the definitions it is evident that

$$e^{t_\star^\kappa - t_2^\kappa} \left(e^{t_2^\kappa} \varepsilon^{\frac{d}{2} - \alpha} t_1^{-\frac{d}{4}} \right)^3 (t_2^\kappa - t_1) \lesssim \left(\log \varepsilon^{-1} \right)^{-\frac{d}{2} - \kappa}.$$

We observe that at this point we used that $t_2(\varepsilon)$ contains an additional negative term $-\frac{1}{2} \log \log \varepsilon^{-1}$, at least in dimension $d = 2$, to obtain an upper bound that vanishes as $\varepsilon \to 0$. As for the second term, we have

$$e^{t_\star^\kappa - t_1} \varepsilon^{2 - 3\tilde{\alpha}} \left(\varepsilon^{1 - \tilde{\alpha}} \Gamma_{d/2}^{3/2} (t_1 \varepsilon^{-2}) \right)^{N-1} \lesssim \varepsilon^{(1-\tilde{\alpha})(N+1) - \frac{d}{2}} \left(\log \varepsilon^{-1} \right)^{-\kappa} \Gamma_{d/2}^{3(N-1)/2} (t_1 \varepsilon^{-2}),$$

which also vanishes as $\varepsilon \to 0$ by our choice of N. It follows that

$$\limsup_{\varepsilon \to 0} \left(1 + \frac{L_\varepsilon}{\sqrt{t_\star^\kappa - t_2^\kappa}} \right) e^{t_\star^\kappa - t_2^\kappa} \sup_y \| S_\varepsilon^N (t_2^\kappa, y) \|_q = 0.$$

Applying the second part of Lemma 4.2, the above estimates imply that the process $(t, x) \mapsto P^1_{t_\star^\kappa - t_2^\kappa + t} S_\varepsilon^N (t_2^\kappa, L_\varepsilon x)$ converges to 0 in probability in $C_{\mathrm{loc}}(\mathbf{R} \times \mathbf{R}^d)$. □

Next, we observe that the truncated Wild expansion converges to the Bargmann–Fock field around time $t_\star(\varepsilon)$.

Lemma 4.4 *For every $\kappa \geq 0$, as $\varepsilon \to 0$, we have:*

(i) *The process $\{e^{-\kappa \log \log \varepsilon^{-1}} P^1_{t_\star^\kappa - t_1 + t} X_\varepsilon^\bullet(t_1, L_\varepsilon x) : (t, x) \in \mathbf{R} \times \mathbf{R}^d\}$ converges in law in $C_{loc}(\mathbf{R} \times \mathbf{R}^d)$ to $\{e^t \Psi(x) : (t, x) \in \mathbf{R} \times \mathbf{R}^d\}$ with Ψ as in (1.3).*
(ii) *For each $\tau \in \mathcal{T}_3 \setminus \{\bullet\}$, the process $\{e^{-\kappa \log \log \varepsilon^{-1}} P^1_{t_\star^\kappa - t_1 + t} X_\varepsilon^\tau(t_1, L_\varepsilon x) : (t, x) \in \mathbf{R} \times \mathbf{R}^d\}$ converges in probability to 0 in $C_{loc}(\mathbf{R} \times \mathbf{R}^d)$.*

Proof We first show that the family $\{e^{-\kappa \log \log \varepsilon^{-1}} P^1_{t_\star^\kappa - t_1 + t} X_\varepsilon^\tau(t_1, L_\varepsilon x) : (t, x) \in \mathbf{R} \times \mathbf{R}^d\}_\varepsilon$ is tight in $C_{\mathrm{loc}}(\mathbf{R} \times \mathbf{R}^d)$ for every $\tau \in \mathcal{T}_3$. Hence, let us fix $q > d + 1$. From Proposition 3.15, we have

$$e^{t_\star^\kappa - t_1 + t} e^{-\kappa \log \log \varepsilon^{-1}} \sup_y \| X_\varepsilon^\tau(t_1, y) \|_{L^q} \lesssim e^{t_\star + t} \varepsilon^{\frac{d}{2} - \alpha} t_1^{-\frac{d}{4}} = \mathcal{O}(1) \quad \text{as} \quad \varepsilon \downarrow 0.$$

In addition, we observe that

$$\frac{L_\varepsilon}{\sqrt{t_\star^\kappa - t_1}} = \mathcal{O}(1) \quad \text{as} \quad \varepsilon \downarrow 0,$$

so that an application of Lemma 4.2 implies tightness for the sequence.

Let us now prove the first point. By construction, the process

$$(t, x) \mapsto e^{-\kappa \log \log \varepsilon^{-1}} P^1_{t_\star - t_1 + t} X_\varepsilon^\bullet(t_1, L_\varepsilon x) = e^{-\kappa \log \log \varepsilon^{-1}} X_\varepsilon^\bullet(t_\star^\kappa + t, L_\varepsilon x)$$

is a centered Gaussian random field with covariance

$$\mathbf{E}[e^{-\kappa \log \log \varepsilon^{-1}} X_\varepsilon^\bullet(t_\star^\kappa + t, L_\varepsilon x) e^{-\kappa \log \log \varepsilon^{-1}} X_\varepsilon^\bullet(t_\star^\kappa + t, L_\varepsilon y)]$$
$$= e^{2t} e^{2t_\star} \varepsilon^{d-2\alpha} p_{t_\star^\kappa + t}^\varepsilon * p_{t_\star^\kappa + t}^\varepsilon (L_\varepsilon(x - y)).$$

Now, it is straightforward to verify that

$$\lim_{\varepsilon \to 0} (t_\star^\kappa)^{-\frac{d}{2}} e^{2t_\star} \varepsilon^{d-2\alpha} = \left(\frac{d}{2} - \alpha\right)^{-\frac{d}{2}} e^{2\mathfrak{c}}$$

with $\mathfrak{c}$ as in (1.4) and

$$\lim_{\varepsilon \to 0} (t_\star^\kappa)^{\frac{d}{2}} p_{t_\star^\kappa + t}^\varepsilon * p_{t_\star^\kappa + t}^\varepsilon (L_\varepsilon(x - y)) = (8\pi)^{-\frac{d}{2}} \exp\left(-\frac{|x - y|^2}{8}\right),$$

so that (i) is verified.

As for the second point, let $\tau \in \mathcal{T}_3 \setminus \{\bullet\}$ and q be fixed such that $q > d + 1$. From Proposition 3.13, we have

$$e^{t_\star^\kappa - t_1} e^{-\kappa \log \log \varepsilon^{-1}} \sup_y \|X_\varepsilon^\tau(t_1, y)\|_q \lesssim \left(\varepsilon^{1-\bar\alpha} \Gamma_{d/2}^{1/2}(t_1 \varepsilon^{-2})\right)^{\ell(\tau)-1}.$$

Since $\bar\alpha < 1$ and $\ell(\tau) > 1$, the right-hand side above vanishes as $\varepsilon \to 0$. By Lemma 4.2, this implies (ii). $\qquad\square$

Remark 4.5 It is possible to show that the process $\{P_{t_\star - t_1 + t}^1 u_\varepsilon(t_1, L_\varepsilon x) : (t, x) \in \mathbf{R} \times \mathbf{R}^d\}$ converges in law to $\{e^t \Psi(x) : (t, x) \in \mathbf{R} \times \mathbf{R}^d\}$ in $C_{\mathrm{loc}}(\mathbf{R} \times \mathbf{R}^d)$. However, this fact is not needed in what follows, so its proof is omitted.

In the upcoming result we will work with the flow $\overline{\Phi}(t, u)$ given by

$$\overline{\Phi}(t, u) = \frac{e^t u}{\left(1 + (e^{2t} - 1)u^2\right)^{1/2}}, \qquad t \in \mathbf{R}, \quad u \in \mathbf{R}, \tag{4.2}$$

which solves (1.5) with initial condition $\overline{\Phi}(0, u) = u$ (the feature that distinguishes Φ from $\overline{\Phi}$ is the initial condition). With this definition we have for $t > 0$

$$\partial_u \overline{\Phi}(t, u) = e^t (1 + u^2(e^{2t} - 1))^{-\frac{3}{2}},$$
$$\frac{\partial_u^2 \overline{\Phi}(t, u)}{\partial_u \overline{\Phi}(t, u)} = 3u \frac{e^{2t} - 1}{1 + u^2(e^{2t} - 1)} \quad \text{and} \quad |\partial_u^2 \overline{\Phi}(t, u)| \lesssim e^{2t}. \tag{4.3}$$

To obtain the last bound we observe that

$$|\partial_u^2 \overline{\Phi}(t, u)| = \frac{3|u|e^t(e^{2t}-1)}{(1+u^2(e^{2t}-1))^{\frac{5}{2}}}.$$

Then for $|u| \leq e^{-t}$ we have $|\partial_u^2 \overline{\Phi}(t, u)| \lesssim e^{2t} - 1 \lesssim e^{2t}$, since in the denominator $1 + u^2(e^{2t} - 1) \geq 1$. Instead for $|u| > e^{-t}$ the leading term in the denominator is $u^2 e^{2t}$, so we find

$$|\partial_u^2 \Phi(t, u)| = \frac{3|u|e^t(e^{2t} - 1)}{(1 + u^2(e^{2t} - 1))^{\frac{5}{2}}} \leq \frac{3|u|e^t(e^{2t} - 1)}{u^5 e^{5t}} \lesssim \frac{1}{|u|^4} \frac{e^{3t}}{e^{5t}} \lesssim e^{2t}.$$

Next, let us define for every $t > t_2^\kappa$

$$w_\varepsilon^{N, \kappa}(t, x) = \overline{\Phi}(t - t_2^\kappa, P_{t - t_2^\kappa} P_{t_2^\kappa - t_1}^1 u_\varepsilon^N(t_1, x))$$

$$= \overline{\Phi}(t - t_2^\kappa, e^{-(t - t_2^\kappa)} P_{t - t_1}^1 u_\varepsilon^N(t_1, x)).$$

When $\kappa = 0$, we simply write w_ε^N for $w_\varepsilon^{N,0}$. Now we can prove the main result of this section, which states that $w_\varepsilon^{N,\kappa}$ is a good approximation of u_ε in a time window of order one about time $t_\star^\kappa(\varepsilon)$, and at spatial scales of order L_ε.

Proposition 4.6 *For every $\kappa \in [0, \frac{1}{4})$, the process*

$$(t, x) \mapsto |u_\varepsilon(t_\star^\kappa(\varepsilon) + t, xL_\varepsilon) - w_\varepsilon^{N, \kappa}(t_\star^\kappa(\varepsilon) + t, xL_\varepsilon)|,$$

converges to 0 in probability in $C_{loc}(\mathbf{R} \times \mathbf{R}^d)$ as $\varepsilon \to 0$.

Proof From its definition, we see that $w_\varepsilon^{N,\kappa}$ satisfies for $t > t_2^\kappa(\varepsilon)$

$$(\partial_t - \Delta - 1)w_\varepsilon^{N,\kappa} = -(w_\varepsilon^{N,\kappa})^3 + f_\varepsilon^N, \quad w_\varepsilon^{N,\kappa}(t_2^\kappa, \cdot) = P_{t_2^\kappa - t_1}^1 u_\varepsilon^N(t_1, \cdot),$$

where

$$f_\varepsilon^N(t, x) = -\partial_u^2 \overline{\Phi}\left(t - t_2^\kappa, e^{-(t - t_2^\kappa)} P_{t - t_1}^1 u_\varepsilon^N(t_1, x)\right) e^{-2(t - t_2^\kappa)} |\nabla P_{t - t_1}^1 u_\varepsilon^N(t_1, x)|^2.$$

It follows that the difference $v_\varepsilon^N = w_\varepsilon^{N,\kappa} - u_\varepsilon$ satisfies the equation

$$(\partial_t - \Delta - 1)v_\varepsilon^N = -((v_\varepsilon^N)^2 + 3u_\varepsilon v_\varepsilon^N + 3u_\varepsilon^2)v_\varepsilon^N + f_\varepsilon^N \quad \forall t > t_2^\kappa$$

with initial condition at t_2^κ given by $v_\varepsilon^N(t_2^\kappa, \cdot) = P_{t_2^\kappa - t_1}^1 u_\varepsilon^N(t_1, \cdot) - u_\varepsilon(t_2^\kappa, \cdot)$. Applying Lemma 3.1, we see that

$$|u_\varepsilon - w_\varepsilon^{N, \kappa}|(t, x) \leq P_{t - t_2^\kappa}^1 |P_{t_2^\kappa - t_1}^1 u_\varepsilon^N(t_1, \cdot) - u_\varepsilon(t_2^\kappa, \cdot)|(x)$$

$$+ \int_{t_2^\kappa}^t \int_{\mathbf{R}^d} p_{t-s}^1(x - y)|f_\varepsilon^N(s, y)| \, \mathrm{d}y \, \mathrm{d}s$$

for every $(t, x) \in [t_2^\kappa, \infty) \times \mathbf{R}^d$. By Lemma 4.3, the process

$$(t, x) \mapsto \left(P_{t_\star^\kappa + t - t_2^\kappa}^1 |P_{t_2^\kappa - t_1}^1 u^N(t_1, \cdot) - u_\varepsilon(t_2^\kappa, \cdot)|\right)(L_\varepsilon x)$$

converges to 0 in probability in $C_{\text{loc}}(\mathbf{R} \times \mathbf{R}^d)$. We also note that in view of (4.3), $|f_\varepsilon^N(t, x)| \lesssim |\nabla P_{t-t_1}^1 u_\varepsilon^N(t_1, x)|^2$. Hence it remains to show that the process

$$H_\varepsilon(t, x) = \int_{t_2^\kappa}^{t_\star^\kappa + t} \int_{\mathbf{R}^d} p_{t_\star^\kappa + t - s}^1(L_\varepsilon x - y)|\nabla P_{s-t_1}^1 u_\varepsilon^N(t_1, y)|^2 \, \mathrm{d}y \, \mathrm{d}s$$

converges to 0 in probability in $C_{\text{loc}}(\mathbf{R} \times \mathbf{R}^d)$.

For every $z_1 = (\sigma_1, x_1)$, $z_1 = (\sigma_2, x_2)$ in $\mathbf{R} \times \mathbf{R}^d$, define the distance $d(z_1, z_2) = |\sigma_1 - \sigma_2|^{\frac{1}{2}} + |z_1 - z_2|$. Let K be a compact set in $\mathbf{R}^{d+1}$ and $q = d + 2$. It suffices to show that

$$\lim_{\varepsilon \to 0} \sup_{z \in K} \mathbf{E}|H_\varepsilon(z)|^q + \mathbf{E} \int_{K \times K} \frac{|H_\varepsilon(z_1) - H_\varepsilon(z_2)|^q}{d(z_1, z_2)^q} \, \mathrm{d}z_1 \, \mathrm{d}z_2 = 0. \tag{4.4}$$

In fact, by the Morrey–Sobolev inequality, the above estimate implies that $\sup_{z \in K} |H_\varepsilon(z)|$ has vanishing q-th moment as $\varepsilon \to 0$, which in turn, implies the convergence of H_ε to 0 in $C_{\text{loc}}(\mathbf{R} \times \mathbf{R}^d)$ in probability. From Proposition 3.13, we have

$$\|\nabla P_{s-t_1}^1 u_\varepsilon^N(t_1, y)\|_q \lesssim s^{-\frac{d}{4} - \frac{1}{2}} e^s \varepsilon^{\frac{d}{2} - \alpha},$$

which implies that

$$e^{t_\star^\kappa - s} \sup_{y \in \mathbf{R}^d} \|\nabla P_{s-t_1}^1 u_\varepsilon^N(t_1, y)\|_q^2 \lesssim \left((t_2^\kappa)^{-\frac{d}{4} - \frac{1}{2}} e^{t_\star^\kappa} \varepsilon^{\frac{d}{2} - \alpha} \right)^2$$

$$\lesssim \left(\log \varepsilon^{-1} \right)^{2\kappa - 1} \left(\frac{\log \varepsilon^{-1}}{t_2^\kappa} \right)^{\frac{d}{2} + 1}, \tag{4.5}$$

for every $s \in (t_2^\kappa, t_\star^\kappa + \text{dist}(0, K))$. It follows that for every $z = (\sigma, x)$ in K

$$\|H_\varepsilon(\sigma, x)\|_q \lesssim \int_{t_2^\kappa}^{t_\star^\kappa + \sigma} e^{t_\star^\kappa - s} \sup_{y \in \mathbf{R}^d} \|\nabla P_{s-t_1}^1 u_\varepsilon^N(t_1, y)\|_{2q}^2 \, \mathrm{d}s$$

$$\lesssim \left(\log \varepsilon^{-1} \right)^{2\kappa - 1} \left(\frac{\log \varepsilon^{-1}}{t_2^\kappa} \right)^{\frac{d}{2} + 1} \left(t_\star^\kappa - t_2^\kappa + \text{dist}(0, K) \right).$$

From the definitions of t_2 and $t_\star$, the right-hand side above vanishes as $\varepsilon \to 0$.

For every $z_1 = (\sigma_1, x_1)$, $z_2 = (\sigma_2, x_2) \in K$, we now estimate the increment $H_\varepsilon(z_2) - H_\varepsilon(z_1)$. The increment in the spatial variables $\|H_\varepsilon(\sigma_1, x_1) - H_\varepsilon(\sigma_1, x_2)\|_q$ can be estimated uniformly by a constant multiple of

$$|x_1 - x_2| \int_{t_2}^{t_\star + \sigma_1} \int_{\mathbf{R}^d} L_\varepsilon |\nabla p_{t_\star + \sigma_1 - s}^1(y)| \, \mathrm{d}y \left(\sup_y \|\nabla P_{s-t_1}^1 u_\varepsilon^N(t_1, y)\|_{2q}^2 \right) \mathrm{d}s.$$

Taking into account the fact that $L_\varepsilon \|\nabla p^1_{t_\star+\sigma_1-s}\|_{L^1(\mathbf{R}^d)} \lesssim L_\varepsilon(t_\star + \sigma_1 - s)^{-\frac{1}{2}} e^{t_\star-s}$ uniformly for every $z_1 \in K$, this gives the estimate

$$\|H_\varepsilon(\sigma_1, x_1) - H_\varepsilon(\sigma_1, x_2)\|_q$$
$$\lesssim |x_1 - x_2| \int_{t_2}^{t_\star+\sigma_1} L_\varepsilon(t_\star + \sigma_1 - s)^{-\frac{1}{2}} \sup_y e^{t_\star-s} \|\nabla P^1_{s-t_1} u^N_\varepsilon(t_1, y)\|^2_{2q} \, ds \, .$$

Simplifying the integration on the right-hand side, using the estimate (4.5), we obtain

$$\|H_\varepsilon(\sigma_1, x_1) - H_\varepsilon(\sigma_1, x_2)\|_q$$
$$\lesssim |x_1 - x_2| \left(\log \varepsilon^{-1}\right)^{2\kappa-\frac{1}{2}} \left(\frac{\log \varepsilon^{-1}}{t_2^\kappa}\right)^{\frac{d}{2}+1} \left(t_\star^\kappa - t_2^\kappa + \sigma\right)^{\frac{1}{2}} \, .$$

To estimate the increment in the time variables, we assume without lost of generality that $\sigma_1 < \sigma_2$ and write

$$H(\sigma_2, x_2) - H(\sigma_1, x_2)$$
$$= \int_{t_\star^\kappa+\sigma_1}^{t_\star^\kappa+\sigma_2} \int_{\mathbf{R}^d} p^1_{t_\star^\kappa+\sigma_2-s}(L_\varepsilon x - y)|\nabla P^1_{s-t_1} u^N_\varepsilon(t_1, y)|^2 \, dy \, ds$$
$$+ \int_{t_2^\kappa}^{t_\star^\kappa+\sigma_1} \int_{\mathbf{R}^d} \left(p^1_{t_\star^\kappa+\sigma_2-s} - p^1_{t_\star^\kappa+\sigma_1-s}\right)(L_\varepsilon x - y)|\nabla P^1_{s-t_1} u^N_\varepsilon(t_1, y)|^2 \, dy \, ds$$
$$=: I_1 + I_2 \, .$$

The first term is estimated easily

$$\|I_1\|_q \lesssim \int_{t_\star^\kappa+\sigma_1}^{t_\star^\kappa+\sigma_2} e^{t_\star^\kappa+\sigma_2-s} \sup_y \|\nabla P^1_{s-t_1} u^N_\varepsilon(t_1, y)\|^2_{2q} \, ds$$
$$\lesssim |\sigma_2 - \sigma_1| \left(\log \varepsilon^{-1}\right)^{2\kappa-1} \left(\frac{\log \varepsilon^{-1}}{t_2^\kappa}\right)^{\frac{d}{2}+1} \, .$$

For the second term, we use the elementary estimates

$$|p^1_{t_\star^\kappa+\sigma_2-s}(y) - p^1_{t_\star^\kappa+\sigma_1-s}(y)| \le \int_{\sigma_1}^{\sigma_2} |\partial_\sigma p^1_{t_\star^\kappa+\sigma-s}(y)| \, d\sigma$$

and

$$\|\partial_\sigma p^1_{t_\star^\kappa+\sigma-s}\|_{L^1(\mathbf{R}^d)} \lesssim e^{t_\star^\kappa+\sigma-s} \left(1 + (t_\star^\kappa + \sigma - s)^{-1}\right)$$

to obtain

$$\|I_2\|_q \lesssim \int_{t_2^\kappa}^{t_\star^\kappa+\sigma_1} \int_{\sigma_1}^{\sigma_2} \left(1 + (t_\star^\kappa + \sigma - s)^{-1}\right) \, d\sigma \, e^{t_\star^\kappa+\sigma_2-s} \sup_y \|\nabla P^1_{s-t_1} u^N_\varepsilon(t_1, y)\|^2_{2q} \, ds \, .$$

Using the estimate (4.5), it is straightforward to verify that

$$\|I_2\|_q \lesssim |\sigma_2 - \sigma_1|^{\frac{1}{2}} \left(\log \varepsilon^{-1}\right)^{2\kappa-1} \left(\frac{\log \varepsilon^{-1}}{t_2^{\kappa}}\right)^{\frac{d}{2}+1} (t_{\star}^{\kappa} - t_2^{\kappa} + \text{dist}(0, K))^{\frac{1}{2}}.$$

Combining these estimates yields

$$\mathbf{E} \int_{K \times K} \frac{|H_\varepsilon(z_1) - H_\varepsilon(z_2)|^q}{d(z_1, z_2)^q} \, dz_1 \, dz_2$$

$$\lesssim \left(\log \varepsilon^{-1}\right)^{2\kappa-\frac{1}{2}} \left(\frac{\log \varepsilon^{-1}}{t_2^{\kappa}}\right)^{\frac{d}{2}+1} \left(t_{\star}^{\kappa} - t_2^{\kappa} + \text{dist}(0, K)\right)^{\frac{1}{2}}.$$

Since $\kappa < \frac{1}{4}$, the right-hand side above vanishes as $\varepsilon \to 0$, which implies (4.4) and completes the proof. $\quad\square$

5 Front propagation

In this section we prove Proposition 2.2 regarding the formation of the initial front and Theorem 2.3 regarding its evolution via mean curvature flow. We start by recalling an *a-priori* bound on solutions to the Allen–Cahn equation.

Lemma 5.1 *For every $u_0 \in C_{\text{loc}}(\mathbf{R}^d)$ satisfying $\sup_{x \in \mathbf{R}^d} |u_0(x)e^{-|x|}| < \infty$, let u be the solution to the Allen–Cahn equation (1.2) with initial condition u_0. Then*

$$|u(t, x)| \leq \frac{e^t}{\sqrt{e^{2t} - 1}}, \qquad \forall t > 0, \ x \in \mathbf{R}^d.$$

Proof We observe that with the definition of $\overline{\Phi}$ as in (4.2) we have $\Xi(t) = e^t/\sqrt{e^{2t} - 1} = \lim_{u \to \infty} \overline{\Phi}(t, u)$, so that $\Xi(t)$ is a space-independent solution to the Allen–Cahn equation on $(0, \infty)$. By comparison, since $\lim_{t \to 0} \Xi(t) = \infty$, and approximating u_0 with uniformly bounded initial conditions, we obtain $u(t) \leq \Xi(t)$ for all $t > 0$. Similarly one derives the upper bound. The growth condition on u_0 is used to justify the approximation procedure, as well as the well-posedness of the heat flow started in u_0. $\quad\square$

Since our results are concerned with convergence in law of the process u_ε, the choice of underlying probability space is not relevant. In the next lemma we build a probability space on which our family of processes converges in probability. We note that the Skorokhod representation theorem is usually stated for discrete families of random variables, rather than continuous ones hence, for completeness, we include a proof of our statement.

Lemma 5.2 *There exists a probability space $(\Omega, \mathcal{F}, \mathbf{P})$ supporting a sequence of processes $\{u_\varepsilon(t, x) : (t, x) \in [t_\star(\varepsilon), \infty) \times \mathbf{R}^d\}$ with the same law as the solutions to (1.2), and a Gaussian process $\{\Psi(x) : x \in \mathbf{R}^d\}$ satisfying (1.3), such that*

$\{u_\varepsilon(t_\star(\varepsilon), xL_\varepsilon) \; : \; x \in \mathbf{R}^d\}$ *converges, as* $\varepsilon \to 0$, *to* $\{\Phi(0, \Psi(x)) \; : \; x \in \mathbf{R}^d\}$ *in probability in* $C_{\mathrm{loc}}(\mathbf{R}^d)$.

Proof Let us consider a probability space $(\Omega, \mathcal{F}, \mathbf{P})$ supporting a white noise η on $\mathbf{R}^d$. Then define, for $\varepsilon \in (0, 1)$, the random fields $x \mapsto X_\varepsilon^\bullet(t_\star(\varepsilon), xL_\varepsilon)$ by

$$X_\varepsilon^\bullet(t_\star(\varepsilon), xL_\varepsilon) = K_\varepsilon * \varphi^\varepsilon * \eta,$$

$$K_\varepsilon(x) = e^{t_\star(\varepsilon) + \mathfrak{c}} \varepsilon^{\frac{d}{2} - \alpha} L_\varepsilon^{\frac{d}{2}} (4\pi t_\star(\varepsilon))^{-\frac{d}{2}} \exp\left\{ -\frac{|x|^2 L_\varepsilon^2}{4 t_\star(\varepsilon)} \right\}.$$

Here φ^ε is as in (1.1), $\mathfrak{c}$ as in (1.4) and we observe that in so far $x \mapsto X_\varepsilon^\bullet(t_\star(\varepsilon), xL_\varepsilon)$ is a time-independent Gaussian process constructed to have the same law as the solution to (3.3) at time $t_\star(\varepsilon)$. From the definition of $t_\star(\varepsilon)$ we obtain the convergence of $K_\varepsilon(x) \to K(x)$, with

$$K(x) = \frac{(8\pi)^{\frac{d}{4}}}{(4\pi)^{\frac{d}{2}}} \exp\left(-\frac{|x|^2}{4} \right),$$

from which we obtain that $X_\varepsilon^\bullet(t_\star(\varepsilon), xL_\varepsilon) \to \Psi(x)$ uniformly over x, almost surely (with Ψ having the required covariance structure), as $\varepsilon \to 0$. In addition, starting from the process $X_\varepsilon^\bullet(t_\star(\varepsilon), \cdot)$ we can construct a sequence of noises η_ε with the same law as (but not identical to) the initial conditions $\varepsilon^{\frac{d}{2} - \alpha} \eta * \varphi^\varepsilon$ appearing in (1.2), and such that $X_\varepsilon^\bullet(t_\star(\varepsilon), x) = P_{t_\star(\varepsilon)}^1 \eta_\varepsilon$.

Now we can follow step by step the proof of Theorem 2.1, to find that if we consider u_ε the solution to (1.2) with the initial condition η_ε we just constructed, then $u_\varepsilon(t_\star(\varepsilon), xL_\varepsilon) \to \Phi(0, \Psi(x))$ in probability in $C_{\mathrm{loc}}(\mathbf{R}^d)$ as $\varepsilon \to 0$. □

The next lemma establishes the formation of the fronts by time $t_\star^\kappa(\varepsilon)$ for some $\kappa \in (0, \frac{1}{2}]$. We write $C_b^k(\mathbf{R}^d; \mathbf{R})$ for the space of k times differentiable functions with all derivatives continuous and bounded. Recall further that

$$t_\star^\kappa(\varepsilon) = t_\star(\varepsilon) + 2\kappa \log L_\varepsilon - 2\kappa \log(d/2 - \alpha).$$

We then define the random nodal set

$$\Gamma_1 = \{x \in \mathbf{R}^d \; : \; \Psi(x) = 0\},$$

and recall the definition of K_δ^1 from (2.1). The proof of the following lemma and of the subsequent proposition follow roughly the approach of [4, Theorem 4.1].

Lemma 5.3 *Consider* $(\Omega, \mathcal{F}, \mathbf{P})$ *as in Lemma 5.2. For any* $0 < \kappa \leq \frac{1}{2}$ *and any sequence* $\{t(\varepsilon)\}_{\varepsilon \in (0,1)}$ *with* $t(\varepsilon) \geq t_\star(\varepsilon)$ *such that*

$$\limsup_{\varepsilon \to 0} \left(t(\varepsilon) - t_\star^{\frac{1}{2}}(\varepsilon) \right) \leq 0 \leq \liminf_{\varepsilon \to 0} \left(t(\varepsilon) - t_\star^\kappa(\varepsilon) \right),$$

it holds that for all $\delta, \zeta \in (0, 1)$

$$\lim_{\varepsilon \to 0} \mathbf{P}\left(\|u_\varepsilon(t(\varepsilon), \cdot L_\varepsilon) - \operatorname{sgn}(\Psi(\cdot))\|_{K_\delta^1} > \zeta\right) = 0.$$

Proof Suppose that along a subsequence $\{\varepsilon_n\}_{n \in \mathbf{N}}$, with $\varepsilon_n \in (0, 1)$, $\lim_{n \to \infty} \varepsilon_n = 0$, it holds that for some $\delta, \zeta, \zeta' \in (0, 1)$

$$\lim_{n \to \infty} \mathbf{P}(\|u_{\varepsilon_n}(t(\varepsilon_n), \cdot L_{\varepsilon_n}) - \operatorname{sgn}(\Psi(\cdot))\|_{K_\delta^1} > \zeta) \geq \zeta'.$$

By our choice of probability space, up to further refining the subsequence, we can assume that $u_{\varepsilon_n}(t_\star(\varepsilon_n), \cdot L_{\varepsilon_n}) \to \Phi(0, \Psi(\cdot))$ almost surely in $C_{\mathrm{loc}}(\mathbf{R}^d)$. We will then show that our assumption is absurd, by proving that almost surely

$$\lim_{n \to \infty} \|u_{\varepsilon_n}(t(\varepsilon_n), \cdot L_{\varepsilon_n}) - \operatorname{sgn}(\Psi(\cdot))\|_{K_\delta^1} = 0.$$

In particular, it will suffice to show that for any $x_0 \in \Gamma_1^c$ there exist (random) $\lambda(x_0), \varrho(x_0), \varepsilon(x_0) > 0$ such that for all $\varepsilon_n \in (0, \varepsilon(x_0))$:

$$1 - \frac{\lambda}{L_{\varepsilon_n}^{2\kappa}} \leq \operatorname{sgn}(\Psi(x)) \cdot u_{\varepsilon_n}(t(\varepsilon_n), x L_{\varepsilon_n}) \leq 1 + \lambda \varepsilon_n^{\frac{d}{2} - \alpha}, \quad \forall x \in B_\delta(x_0). \quad (5.1)$$

For the sake of clarity, let us refrain from writing the subindex n and fix x_0 such that $\Psi(x_0) > 0$ (the opposite case follows analogously). For the upper bound we use Lemma 5.1 to find for some $\lambda > 0$

$$|u_\varepsilon(t(\varepsilon), x L_\varepsilon)| \leq \frac{e^{t(\varepsilon)}}{\sqrt{e^{2t(\varepsilon)} - 1}} \leq 1 + \lambda e^{t(\varepsilon)} \leq 1 + \lambda \varepsilon^{\frac{d}{2} - \alpha}.$$

To establish the lower bound, consider for any constant $K > 0$ and any $\psi \in C_b^2$ the following function (here $\overline{\Phi}$ is as in (4.2)),

$$v_\varepsilon(t, x) = \overline{\Phi}\left(t, \psi(x) - \frac{K}{L_\varepsilon} t\right).$$

We see that, since $\partial_u \overline{\Phi} \geq 0$

$$\partial_t v_\varepsilon = L_\varepsilon^{-2} \Delta v_\varepsilon + v_\varepsilon(1 - v_\varepsilon^2) - K L_\varepsilon^{-1} \partial_u \overline{\Phi} - L_\varepsilon^{-2} \partial_u \overline{\Phi} \Delta \psi - L_\varepsilon^{-2} \partial_u^2 \overline{\Phi} |\nabla \psi|^2$$

$$\leq L_\varepsilon^{-2} \Delta v_\varepsilon + v_\varepsilon(1 - v_\varepsilon^2) - \frac{\partial_u \overline{\Phi}}{L_\varepsilon^2}\left(K L_\varepsilon - |\Delta \psi| - \frac{|\partial_u^2 \overline{\Phi}|}{|\partial_u \overline{\Phi}|} |\nabla \psi|^2\right).$$

Now we observe that for $t > 0$, similarly to (4.3) distinguishing the cases $|u| \leq e^{-t}$ and $|u| \geq e^{-t}$, we can bound

$$\left|\frac{\partial_u^2 \overline{\Phi}(t, u)}{\partial_u \overline{\Phi}(t, u)}\right| \lesssim \left|u \frac{e^{2t} - 1}{1 + u^2(e^{2t} - 1)}\right| \lesssim e^t.$$

Hence there exists a $K(\psi) > 0$ such that v_ε is a subsolution to $\partial_t u = L_\varepsilon^{-2}\Delta u + u(1 - u^2)$ on the time interval $[0, t(\varepsilon) - t_\star(\varepsilon)]$ with initial condition ψ. Here we use that $\limsup_{\varepsilon \to 0}\{t(\varepsilon) - t_\star^{\frac{1}{2}}(\varepsilon)\} \leq 0$, so that for a constant $C > 0$ independent of ε we have $\exp(t(\varepsilon) - t_\star(\varepsilon)) \leq CL_\varepsilon$. In particular, by our assumptions and by the upper bound of (5.1), we can choose $\psi \in C_b^2$ so that $\psi(x) > 0$ for x in a closed ball $\overline{B}_\varrho(x_0)$ about x_0 and such that for some $\varepsilon(\psi) > 0$

$$u_\varepsilon(t_\star(\varepsilon), xL_\varepsilon) \geq \overline{\Phi}(0, \psi(x)) = \psi(x), \qquad \forall x \in \mathbf{R}^d, \ \varepsilon \in (0, \varepsilon(\psi)).$$

Now, using that $u = 1$ is an exponentially stable fixed point for $\overline{\Phi}$, we have that for every $u > 0$ there exists a $\lambda(u)$ that can be chosen locally uniformly over u, such that

$$\overline{\Phi}(t(\varepsilon) - t_\star(\varepsilon), u) \geq \overline{\Phi}(2\kappa \log L_\varepsilon - C, u) \geq 1 - \frac{\lambda}{L_\varepsilon^{2\kappa}}$$

for all ε sufficiently small, and for $C > 0$ such that $t(\varepsilon) \geq t_\star^\kappa(\varepsilon) - C$ for all ε. Then by comparison, using that $\frac{t(\varepsilon) - t_\star(\varepsilon)}{L_\varepsilon} \lesssim \frac{\log L_\varepsilon}{L_\varepsilon} \to 0$, for ε sufficiently small:

$$u_\varepsilon(t(\varepsilon), xL_\varepsilon) \geq \inf_{y \in B_\varrho(x_0)} v_\varepsilon(t(\varepsilon) - t_\star(\varepsilon), y) \geq 1 - \frac{\lambda}{L_\varepsilon^{2\kappa}}, \qquad \forall x \in B_\varrho(x_0).$$

This completes the proof of (5.1) and of the lemma. $\qquad\square$

The following proposition treats slightly longer time scales. Recall the definition of K_δ^1 given in (2.1), for $\delta \in (0, 1)$.

Proposition 5.4 *Consider $(\Omega, \mathcal{F}, \mathbf{P})$ as in Lemma 5.2 and fix any sequence $\{t(\varepsilon)\}_{\varepsilon \in (0,1)}$, with $t(\varepsilon) \geq t_\star(\varepsilon)$, and such that*

$$\liminf_{\varepsilon \to 0}\left(t(\varepsilon) - t_\star^{\frac{1}{2}}(\varepsilon)\right) > 0, \qquad \lim_{\varepsilon \to 0}\frac{t(\varepsilon) - t_\star(\varepsilon)}{T_\varepsilon} = 0.$$

Then for any $\delta, \zeta \in (0, 1)$

$$\lim_{\varepsilon \to 0} \mathbf{P}\left(\|u_\varepsilon(t(\varepsilon), \cdot L_\varepsilon) - \text{sgn}(\Psi(\cdot))\|_{K_\delta^1} > \zeta\right) = 0.$$

Proof As in the proof of the previous lemma, it suffices to prove that for any subsequence $\{\varepsilon_n\}_{n \in \mathbf{N}}$ with $\varepsilon_n \in (0, 1)$, $\lim_{n \to \infty} \varepsilon_n = 0$ for which almost surely

$$u_\varepsilon(t_\star(\varepsilon), \cdot L_\varepsilon) \to \Phi(0, \Psi(\cdot)) \quad \text{in } C_{\text{loc}}(\mathbf{R}^d),$$

it holds that for all $\delta \in (0, 1)$

$$\lim_{n \to \infty} \|u_{\varepsilon_n}(t(\varepsilon_n), \cdot L_{\varepsilon_n}) - \text{sgn}(\Psi(\cdot))\|_{K_\delta^1} = 0.$$

Hence we will work with a fixed realization of all random variables and for the sake of clarity we will refrain from writing the subindex n. In addition, since the case $\Psi < 0$ is identical to the case $\Psi > 0$, let us choose an $x_0 \in \mathbf{R}^d$ such that $\Psi(x_0) > 0$ and define $s(\varepsilon) = t(\varepsilon) - t_\star^{\frac{1}{2}}(\varepsilon)$. We can also assume that $s(\varepsilon) \geq 0$ for all ε. Our aim is to prove the convergence $\lim_{\varepsilon \to 0} u_\varepsilon(t_\star^{\frac{1}{2}}(\varepsilon) + s(\varepsilon), x L_\varepsilon) = 1$ holds true for all x in a ball $B_\varrho(x_0)$ about x_0 of radius $\varrho > 0$.

By Lemma 5.3 we already know that there exist $\overline{\lambda}(x_0), \varrho(x_0) > 0$ and $\varepsilon(x_0) \in (0, 1)$ such that:

$$u_\varepsilon(t_\star^{\frac{1}{2}}(\varepsilon), x L_\varepsilon) \geq 1 - \frac{\overline{\lambda}}{L_\varepsilon}, \qquad \forall \varepsilon \in (0, \varepsilon(x_0)), \quad x \in B_\varrho(x_0),$$

$$u_\varepsilon(t_\star^{\frac{1}{2}}(\varepsilon), x L_\varepsilon) \geq -1 - \overline{\lambda} \varepsilon^{\frac{d}{2} - \alpha}, \qquad \forall x \in \mathbf{R}^d. \tag{5.2}$$

Here the second bound is a consequence of Lemma 5.1 (in fact in the second statement $\overline{\lambda}$ can be chosen deterministic and independent of x_0). Our aim is to show that this front does not move after an additional time $s(\varepsilon)$. Let us define $\widetilde{u}_\varepsilon(\sigma, x) = u_\varepsilon(t_\star^{\frac{1}{2}}(\varepsilon) + \sigma s(\varepsilon), x L_\varepsilon)$, which solves

$$\partial_\sigma \widetilde{u}_\varepsilon = \frac{s(\varepsilon)}{L_\varepsilon^2} \Delta \widetilde{u}_\varepsilon + s(\varepsilon) \widetilde{u}_\varepsilon(1 - \widetilde{u}_\varepsilon^2), \qquad \widetilde{u}_\varepsilon(0, \cdot) = \widetilde{u}_{\varepsilon,0}(\cdot), \tag{5.3}$$

with an initial condition $\widetilde{u}_{\varepsilon,0}(x) = u_\varepsilon(t_\star^{\frac{1}{2}}(\varepsilon), x L_\varepsilon)$. Our purpose is to construct an explicit subsolution $\underline{u}_\varepsilon$ to (5.3) with initial condition "close" to $1_{B_{\frac{\varrho}{2}}(x_0)}$, such that $\lim_{\varepsilon \to 0} \underline{u}_\varepsilon(1, x) = 1$ for all x in a neighbourhood of x_0. Our ansatz is that close to the interface the subsolution is of the following form, for $\zeta(\varepsilon) = \frac{s(\varepsilon)}{L_\varepsilon^2}$:

$$\mathsf{q}\Big(L_\varepsilon\, d(\zeta(\varepsilon)\, \sigma, x) + L_\varepsilon \mathcal{O}(\zeta(\varepsilon))\Big) - \mathcal{O}(L_\varepsilon^{-1}).$$

Here $d(\sigma, x)$ is the signed distance function associated to the mean curvature flow evolution at time $\sigma \geq 0$ of the ball $B_{\frac{\varrho}{2}}(x_0)$ with the sign convention $d(0, x) > 0$ if $x \in B_{\frac{\varrho}{2}}(x_0)$, and $d(0, x) \leq 0$ if $x \in B_{\frac{\varrho}{2}}^c(x_0)$ and $\mathsf{q}(u) = \tanh(u)$ is the traveling wave solution to the Allen–Cahn equation:

$$\ddot{\mathsf{q}} + \mathsf{q}(1 - \mathsf{q}^2) = 0, \qquad \lim_{x \to \pm\infty} \mathsf{q} = \pm 1.$$

Our first step is to construct precisely the subsolution near the interface. Define $\overline{d}(\sigma, x) = e^{-\mu\sigma} d(\sigma, x)$, for some $\mu > 0$ that will be chosen later on. We observe that there are $\sigma' > 0$, $\varrho' \in (0, \varrho/2)$ such that $d(\sigma, x)$ is smooth in the set

$$Q_\varrho = [0, \sigma'] \times \{x : |x - x_0| \in [\varrho/2 - \varrho', \varrho/2 + \varrho']\}$$

and (see for example [6, Equation (6.4), p. 663]) there exists a constant $C > 0$, depending on $\varrho, \varrho', \sigma'$, such that

$$\partial_\sigma d - \Delta d \leq Cd \quad \text{on} \quad Q_\varrho.$$

In particular, fixing $\mu \geq C$ we have $(\partial_\sigma - \Delta)\overline{d} \leq 0$ on Q_ϱ. Then consider, for some $K_1 > 0$

$$w_\varepsilon(\sigma, x) = \mathfrak{q}\left(L_\varepsilon\, \overline{d}(\zeta(\varepsilon)\sigma, x) - K_1\frac{s(\varepsilon)}{L_\varepsilon}\sigma\right) - \frac{\lambda}{L_\varepsilon}.$$

We claim that for K_1 sufficiently large w_ε is a subsolution to

$$\partial_\sigma w_\varepsilon - \zeta(\varepsilon)\Delta w_\varepsilon - s(\varepsilon)w_\varepsilon(1 - w_\varepsilon^2) \leq 0, \quad \text{on} \quad Q_\varrho^\varepsilon, \tag{5.4}$$

where

$$Q_\varrho^\varepsilon = [0, \zeta(\varepsilon)^{-1}\sigma'] \times \{x : |x - x_0| \in [\varrho/2 - \varrho', \varrho/2 + \varrho']\}.$$

In fact, since $\dot{\mathfrak{q}} \geq 0$, we can compute

$$\partial_\sigma w_\varepsilon = L_\varepsilon\dot{\mathfrak{q}}\,\zeta(\varepsilon)\,\partial_\sigma\overline{d} - K_1\frac{s(\varepsilon)}{L_\varepsilon}\dot{\mathfrak{q}}$$

$$= \zeta(\varepsilon)\Delta w_\varepsilon + L_\varepsilon\zeta(\varepsilon)\,\dot{\mathfrak{q}}(\partial_\sigma - \Delta)\overline{d} - \zeta(\varepsilon)L_\varepsilon^2\ddot{\mathfrak{q}}|\nabla d|^2 - K_1\frac{s(\varepsilon)}{L_\varepsilon}\dot{\mathfrak{q}}$$

$$\leq \zeta(\varepsilon)\Delta w_\varepsilon - s(\varepsilon)\ddot{\mathfrak{q}} - K_1\frac{s(\varepsilon)}{L_\varepsilon}\dot{\mathfrak{q}}$$

where we used that $|\nabla d|^2 = 1$ in a neighborhood of the boundary of $B_{\frac{\varrho}{2}}$. Now we use the definition of $\mathfrak{q}$ to rewrite the last term as

$$\zeta(\varepsilon)\Delta w_\varepsilon + s(\varepsilon)\mathfrak{q}(1 - \mathfrak{q}^2) - K_1\tfrac{s(\varepsilon)}{L_\varepsilon}\dot{\mathfrak{q}}.$$

At this point we would like to replace $\mathfrak{q}$ with $w_\varepsilon = \mathfrak{q} - \frac{\lambda}{L_\varepsilon}$ in the nonlinearity. We observe that since $u \mapsto u(1 - u^2)$ is decreasing near $u = 1$ and $u = -1$ there exists a $\gamma \in (0, 1)$ such that if $|\mathfrak{q}| \in (\gamma, 1)$ and ε is sufficiently small, then $\mathfrak{q}(1 - \mathfrak{q}^2) \leq w_\varepsilon(1 - w_\varepsilon^2)$. On the other hand, on the set $|\mathfrak{q}| \leq \gamma$ there exists a constant $c(\gamma) > 0$ such that $\dot{\mathfrak{q}} \geq c(\gamma) > 0$. Hence in this last case:

$$\zeta(\varepsilon)\Delta w_\varepsilon + s(\varepsilon)\mathfrak{q}(1 - \mathfrak{q}^2) - K_1\frac{s(\varepsilon)}{L_\varepsilon}\dot{\mathfrak{q}}$$

$$\leq \zeta(\varepsilon)\Delta w_\varepsilon + s(\varepsilon)w_\varepsilon(1 - w_\varepsilon^2)$$

$$+ s(\varepsilon)\left\{\frac{\lambda}{L_\varepsilon} - 3\frac{\lambda}{L_\varepsilon}w_\varepsilon^2 - \frac{c(\gamma)K_1}{L_\varepsilon} + \mathcal{O}\left(\frac{1}{L_\varepsilon^2}\right)\right\}1_{\{|\mathfrak{q}|\leq\gamma\}}$$

$$\leq \zeta(\varepsilon)\Delta w_\varepsilon + s(\varepsilon)w_\varepsilon(1 - w_\varepsilon^2),$$

where the last inequality holds for all ε sufficiently small, and provided K_1 is chosen large enough. Hence we have proven (5.4).

The next step is to extend this subsolution w_ε to all $x \in \mathbf{R}^d$ (at the moment it is defined only for $|x - x_0| \in [\varrho/2 - \varrho', \varrho/2 + \varrho']$). Here we follow two different arguments in the interior and the exterior of the ball $B_{\frac{\varrho}{2}}$. Let us start with the exterior. We observe that for any fixed $\varrho'' \in (0, \varrho')$ it holds that for all ε sufficiently small and some $\lambda' > 0$:

$$w_\varepsilon(\sigma, x) \leq -1 - \frac{\lambda'}{L_\varepsilon},$$
$$\forall (\sigma, x) \in [0, \zeta(\varepsilon)^{-1}\sigma'] \times \{x : |x - x_0| \in [\varrho/2 + \varrho'', \varrho/2 + \varrho']\}.$$

Here we use that asymptotically, for $x \to -\infty$, we have $\mathsf{q}(x) \leq -1 + 2e^{-2x}$, so that by definition, for some $c(x_0) > 0$

$$\sup_{|x - x_0| \in [\varrho/2 + \varrho'', \varrho/2 + \varrho']} w_\varepsilon(\sigma, x) \leq \mathsf{q}\left(-L_\varepsilon \left(c(x_0) + \mathcal{O}(\zeta(\varepsilon))\right)\right) - \frac{\lambda}{L_\varepsilon}$$
$$\leq -1 + 2\exp\left(-\left(c(x_0) + \mathcal{O}(\zeta_\varepsilon)\right) L_\varepsilon\right) - \frac{\lambda}{L_\varepsilon} \leq -1 - \frac{\lambda'}{L_\varepsilon},$$

where the last inequality holds for ε sufficiently small. Now consider $\bar{\lambda}$ as in (5.2) and $\overline{\Phi}$ as in (4.2). Then let S^ε be the set

$$S^\varepsilon = \{x : |x - x_0| \in [\varrho/2 + \varrho'', \varrho/2 + \varrho']\}.$$

Since $\sigma \mapsto \overline{\Phi}(s(\varepsilon)\sigma, -1 - \bar{\lambda}\varepsilon^{\frac{d}{2} - \alpha}) \stackrel{\text{def}}{=} \overline{\Phi}_\varepsilon(\sigma)$ is a spatially homogeneous solution to (5.3) we find that for small ε,

$$\underline{w}_\varepsilon(\sigma, x) = \begin{cases} \max\{w_\varepsilon(\sigma, x), \overline{\Phi}_\varepsilon(\sigma)\}, & \text{if } |x - x_0| \in [\frac{\varrho}{2} - \varrho', \frac{\varrho}{2} + \varrho'], \\ \overline{\Phi}_\varepsilon(\sigma) & \text{else.} \end{cases}$$

is a subsolution to (5.3) in the viscosity sense (cf. the proof of Lemma 3.1) on the set

$$[0, \zeta(\varepsilon)^{-1}\sigma'] \times \{x : |x - x_0| \geq \varrho/2 - \varrho'\}.$$

Here we use that asymptotically $\lambda' L_\varepsilon^{-1} \gg \varepsilon^{\frac{d}{2} - \alpha}$, so that $\underline{w}_\varepsilon(\sigma, x) = \overline{\Phi}_\varepsilon(\sigma)$ for all $x \in S^\varepsilon$.

Finally, we want to extend the subsolution to the interior of the ball $B_{\frac{\varrho}{2}}(x_0)$. To complete this extension we consider a convex combination between $\underline{w}_\varepsilon$ and the constant $1 - \frac{\lambda}{L_\varepsilon}$. Let us fix a decreasing smooth function $\Upsilon : \mathbf{R} \to [0, 1]$ such that $\Upsilon(x) = 1$ if

$x \leq \left(\frac{\varrho - \varrho'}{2}\right)^2$, and $\Upsilon(x) = 0$ if $x \geq \left(\frac{\varrho - \varrho'/2}{2}\right)^2$. Then define, for some constant $K_2 > 0$

$$\underline{u}_\varepsilon(\sigma, x) = \left(1 - \Upsilon\left(|x - x_0|^2 + K_2\zeta(\varepsilon)\sigma\right)\right)\underline{w}_\varepsilon(\sigma, x)$$
$$+ \Upsilon(|x - x_0|^2 + K_2\zeta(\varepsilon)\sigma)\left(1 - \frac{\lambda}{L_\varepsilon}\right).$$

Note that by considerations on the support of Υ and the domain of definition of $\underline{w}_\varepsilon$, $\underline{u}_\varepsilon$ is well defined. We claim that if K_2 is sufficiently large the function $\underline{u}_\varepsilon$ is a viscosity subsolution to (5.3) on $[0, \zeta(\varepsilon)^{-1}\sigma'] \times \mathbf{R}^d$. In fact, by our previous calculations we find

$$(\partial_\sigma - \zeta(\varepsilon)\Delta)\underline{u}_\varepsilon \leq s(\varepsilon)\,(1 - \Upsilon)\,\underline{w}_\varepsilon(1 - \underline{w}_\varepsilon^2)$$
$$+ \zeta(\varepsilon)\Big(\underline{w}_\varepsilon - \Big(1 - \frac{\lambda}{L_\varepsilon}\Big)\Big)\Big[2\dot{\Upsilon} + 4\ddot{\Upsilon}|x - x_0|^2 - K_2\dot{\Upsilon}\Big]$$
$$\leq s(\varepsilon)\,(1 - \Upsilon)\,\underline{w}_\varepsilon(1 - \underline{w}_\varepsilon^2)$$
$$+ 4\,\zeta(\varepsilon)\,\ddot{\Upsilon}\,|x - x_0|^2(\underline{w}_\varepsilon - (1 - \lambda/L_\varepsilon)),$$

assuming $K_2 \geq 2$, and using that $\underline{w}_\varepsilon \leq 1 - \frac{\lambda}{L_\varepsilon}$ and $\dot{\Upsilon} \leq 0$, as Υ is decreasing. Now on the set $\{\Upsilon > 0\}$, we have $\underline{w}_\varepsilon \geq 0$, provided ε is sufficiently small. So using the concavity of $[0, 1] \ni u \mapsto u(1 - u^2)$:

$$(1 - \Upsilon)\underline{w}_\varepsilon(1 - \underline{w}_\varepsilon^2) \leq \underline{u}_\varepsilon(1 - \underline{u}_\varepsilon^2) - \Upsilon(1 - \lambda/L_\varepsilon)(1 - (1 - \lambda/L_\varepsilon)^2).$$

Furthermore, we can find a constant $\nu \in (0, 1)$ such that

$$\ddot{\Upsilon} \geq 0 \quad \text{if} \quad \Upsilon \leq \nu.$$

Hence we see that on the set $\{\Upsilon \leq \nu\}$

$$(\partial_\sigma - \zeta(\varepsilon)\Delta)\underline{u}_\varepsilon \leq s(\varepsilon)\,(1 - \Upsilon)\,\underline{w}_\varepsilon(1 - \underline{w}_\varepsilon^2)$$
$$\leq s(\varepsilon)\underline{u}_\varepsilon(1 - \underline{u}_\varepsilon^2).$$

On the other hand, on the set $\{\Upsilon > \nu\}$ we have for some $C > 0$:

$$(\partial_\sigma - \zeta(\varepsilon)\Delta)\underline{u}_\varepsilon \leq s(\varepsilon)\underline{u}_\varepsilon(1 - \underline{u}_\varepsilon^2) - \nu s(\varepsilon)(1 - \lambda/L_\varepsilon)(1 - (1 - \lambda/L_\varepsilon)^2) + C\,\zeta(\varepsilon)\,|\ddot{\Upsilon}|$$
$$\leq s(\varepsilon)\underline{u}_\varepsilon(1 - \underline{u}_\varepsilon^2) - 2\nu\frac{\lambda s(\varepsilon)}{L_\varepsilon} + \mathcal{O}(\zeta(\varepsilon))$$
$$\leq s(\varepsilon)\underline{u}_\varepsilon(1 - \underline{u}_\varepsilon^2),$$

where the last inequality holds for ε sufficiently small.

To conclude, for ε sufficiently small, we have constructed a subsolution to (5.3) such that, by (5.2) and up to choosing λ sufficiently large, the initial condition satisfies

$$\tilde{u}_{\varepsilon,0}(\cdot) \geq \underline{u}_\varepsilon(0, \cdot).$$

By comparison, since $\lim_{\varepsilon \to 0} \underline{u}_\varepsilon(1, x) = 1$, for all x in a neighborhood of x_0 and through the upper bound of Lemma 5.1 our proof is complete. □

The next result establishes convergence to level set solutions of mean curvature flow. Recall that we have defined

$$U_\varepsilon(\sigma, x) = u_\varepsilon(\sigma T_\varepsilon + \tau_\star(\varepsilon), xL_\varepsilon) = u_\varepsilon(t_\star(\varepsilon) + (\sigma - 1)T_\varepsilon, xL_\varepsilon).$$

In these variables the initial condition for the mean curvature flow appears at time $\sigma = 1$. The level set formulation of mean curvature flow is then given by viscosity solutions to the following equation for $(t, x) \in [1, \infty) \times \mathbf{R}^d$

$$\partial_t w = \Delta w - \frac{(\nabla w)^{\otimes 2} : \nabla^2 w}{|\nabla w|^2}, \qquad w(1, x) = w_1(x). \qquad (5.5)$$

Here $(\nabla w)^{\otimes 2} : \nabla^2 w = \sum_{i,j=1}^d \partial_i w \partial_j w \partial_{ij} w$. If w_1 is uniformly continuous on $\mathbf{R}^d$ there exists a unique viscosity solution to the above equation, see e.g. [8, Theorem 1.8]. Furthermore, we will be only interested in the evolution of the sets $\{w > 0\}$, $\{w < 0\}$ and $\{w = 0\}$, which motivates the following definition.

Definition 5.5 For any $f \in C_{\mathrm{loc}}(\mathbf{R}^d; \mathbf{R})$ we define $v(f; \cdot, \cdot) \colon [1, \infty) \times \mathbf{R}^d \to \{-1, 0, 1\}$ by

$$v(f; \sigma, x) = \mathrm{sgn}\left(w(\sigma, x)\right), \qquad (\sigma, x) \in [1, \infty) \times \mathbf{R}^d,$$

with w the viscosity solution to (5.5) with an arbitrary initial condition $w_1 \in C_b^2(\mathbf{R}^d; \mathbf{R})$ satisfying:

$$\{w_1 > 0\} = \{f > 0\}, \quad \{w_1 < 0\} = \{f < 0\}, \quad \{w_1 = 0\} = \{f = 0\}.$$

The function $v(f; \cdot, \cdot)$ does not depend on the particular choice of w_1 by [6, Theorem 5.1].

We recall the definition of the sets K_δ as in (2.1) for $\delta \in (0, 1)$:

$$K_\delta = \{z = (\sigma, x) \in (1, \infty) \times \mathbf{R}^d \ : \ |z| \le \delta^{-1}, \ \sigma > 1 + \delta, \ d(z, \Gamma) \ge \delta\}.$$

Now we can state our concluding result.

Proposition 5.6 *Consider* $(\Omega, \mathcal{F}, \mathbf{P})$ *as in Lemma 5.2 and let* $v(\Psi; \cdot, \cdot) \colon [1, \infty) \times \mathbf{R}^d \to \{-1, 0, 1\}$ *be as in Definition 5.5. Then for any* $\delta, \zeta \in (0, 1)$

$$\lim_{\varepsilon \to 0} \mathbf{P}\left(\|U_\varepsilon(\cdot) - v(\Psi; \cdot)\|_{K_\delta} \ge \zeta\right) = 0.$$

Proof In analogy to the previous results, up to considering suitable subsequences we can assume that $U_\varepsilon(1, \cdot) \to \overline{\Phi}(0, \Psi(\cdot))$ almost surely in $C_{\mathrm{loc}}(\mathbf{R}^d)$ as $\varepsilon \to 0$. Then it suffices to prove that almost surely, for any $\delta \in (0, 1)$

$$\lim_{\varepsilon \to 0} \|U_\varepsilon(\cdot) - v(\Psi; \cdot)\|_{K_\delta} = 0.$$

In the setting just introduced, our aim is to construct suitable super- and sub-solutions to $u_\varepsilon(t_\star(\varepsilon) + (\sigma - 1)T_\varepsilon, xL_\varepsilon)$ with initial conditions that are independent of ε and constant outside of a compact set. After establishing convergence to mean curvature flow for these super- and sub-solutions we will use comparison to obtain the convergence of the original sequence. For convenience we will restrict ourselves to the construction of subsolutions, which will guarantee convergence on the set $\{v(\Psi; \cdot, \cdot) > 0\}$. The construction of supersolutions is analogous.

Consider a smooth function $\varphi \colon \mathbf{R}^d \to [-1, 0]$ such that $\varphi(x) = 0$, for all $|x| \leq \frac{1}{2}$ and $\varphi(x) = -1$, for all $|x| \geq 1$. Then for any $R \geq 1$ define $\varphi_R(x) = \varphi(x/R)$ and consider

$$\underline{u}_R = \varphi_R + (1 + \varphi_R)\frac{\Psi}{\sqrt{1 + \Psi^2}} - R^{-1}.$$

In particular by our assumptions for any $R \geq 1$ there exists a $\varepsilon(R) \in (0, 1)$ such that

$$\underline{u}_R(x) \leq U_\varepsilon(1, x), \quad \forall x \in \mathbf{R}^d, \quad \varepsilon \in (0, \varepsilon(R)). \tag{5.6}$$

Here we use the convergence $U_\varepsilon(1, \cdot) \to \overline{\Phi}(0, \Psi(\cdot))$ in $C_{\mathrm{loc}}(\mathbf{R}^d)$ together with the a-priori bound from Lemma 5.1, which guarantees that $U_\varepsilon(\sigma, \cdot) \geq -1 - R^{-1}$ for $\sigma \geq 1$ and ε sufficiently small. Moreover, locally uniformly over $x \in \mathbf{R}^d$: $\lim_{R \to \infty} \underline{u}_R(x) = \overline{\Phi}(0, \Psi(x))$. Now, let $\underline{u}_{R,\varepsilon}(\sigma, x)$ be the solution to the rescaled Allen–Cahn equation

$$\partial_t \underline{u}_{R,\varepsilon} = \Delta \underline{u}_{R,\varepsilon} + T_\varepsilon \underline{u}_{R,\varepsilon}(1 - \underline{u}_{R,\varepsilon}^2), \qquad \underline{u}_{R,\varepsilon}(1, x) = \underline{u}_R(x). \tag{5.7}$$

Then by [4, Theorem 4.1] $\lim_{\varepsilon \to 0} \underline{u}_{R,\varepsilon} = \mathrm{sgn}(\underline{w}_R)$, in $C_{\mathrm{loc}}((1, \infty) \times \mathbf{R}^d \backslash \{\underline{w}_R = 0\})$, where $\underline{w}_R$ is the unique level set solution to (5.5) with initial condition $\underline{u}_R$. By comparison, using the ordering (5.6) and the fact that U_ε also solves (5.7) with a different initial condition, we obtain that for any $R \geq 1$

$$\lim_{\varepsilon \to 0} U_\varepsilon(\sigma, x) = 1 \qquad \text{locally uniformly over} \qquad (\sigma, x) \in \underline{\mathcal{P}}^R,$$

where $\underline{\mathcal{P}}^R = \{\underline{w}_R(\cdot, \cdot) > 0\} \subseteq (1, \infty) \times \mathbf{R}^d$.

Now we would like to pass similarly to the limit $R \to \infty$, but we have to take care of the fact that the limiting initial condition $\lim_{R \to \infty} \underline{u}_R = \Psi/\sqrt{1 + \Psi^2}$ may not be uniformly continuous on $\mathbf{R}^d$, which complicates the construction of solutions. Of course, this problem is not significant, since we can rescale the initial condition arbitrarily by multiplying with a positive function without modifying the mean curvature flow evolution.

Hence consider a partition of the unity $\{\varphi_k\}_{k\in\mathbf{N}}$ of $\mathbf{R}^d$. Namely, for every $k \in \mathbf{N}$, let $\varphi_k \colon \mathbf{R}^d \to [0, \infty)$ be a smooth function with compact support, such that $\sum_{k\in\mathbf{N}} \varphi_k(x) = 1$. We will assume that the partition is locally finite, in the sense that there exists an $M \in \mathbf{N}$ such that for every $k_1, \ldots, k_M \in \mathbf{N}$ with $k_i \neq k_j$ for all $i \neq j$ we have $\cap_{k=1}^M \operatorname{supp}(\varphi_k) = \emptyset$. In addition we assume that $\sup_{k\in\mathbf{N}} \|\varphi_k\|_{C^2} < \infty$. Then let

$$a_k = 1 + \sup_{x\in\operatorname{supp}(\varphi_k)} \{|\Psi(x)| + |\nabla\Psi(x)| + |\nabla^2\Psi(x)|\}$$

and define

$$g(x) = \sum_{k\in\mathbf{N}} \frac{1}{a_k}\varphi_k(x)\sqrt{1 + \Psi^2(x)}.$$

Now we can consider the solution $\underline{v}_R$ to (5.5) with initial condition $\underline{v}_{R,1}(x) = g(x) \cdot \underline{u}_R(1, x)$, so that from the definition of g we obtain:

$$\sup_{R\geq 1} \|\underline{v}_{R,1}\|_{C_b^2(\mathbf{R}^d)} < \infty. \tag{5.8}$$

By [6, Theorem 5.1] the evolution of the interface does not depend on the particular choice of the initial condition, as long as they share the same initial interface. In particular, since g is strictly positive we find that

$$\underline{\mathcal{P}}^R = \{\underline{v}_R(\cdot, \cdot) > 0\}.$$

The bound (5.8) now guarantees a uniform bound on the solutions (this is the consequence of the comparison principle in [6, Theorem 4.1]) for any $\alpha \in (0, 1)$:

$$\sup_{R\geq 1} \|\underline{v}_R\|_{C_b^\alpha((1,\infty)\times\mathbf{R}^d)} < \infty.$$

Since in addition

$$\lim_{R\to\infty} \underline{v}_{R,1} = \sum_{k\in\mathbf{N}} \varphi_k \frac{\Psi}{a_k} \overset{\text{def}}{=} w_1,$$

with $w_1 \in C_b^2(\mathbf{R}^d; \mathbf{R})$, using compactness as well as stability of viscosity solutions we find that $\lim_{R\to\infty} \underline{v}_R(\cdot, \cdot) = w(\cdot, \cdot)$ in $C([1, \infty) \times \mathbf{R}^d; \mathbf{R})$, where the latter is the unique viscosity solution to (5.5) with initial condition w_1. This completes the proof. □

The following corollary relates level set solutions to mean curvature flow to classical solutions in the case $d = 2$. Consider the unit torus $\mathbf{T} = \mathbf{R}/\mathbf{Z}$. A continuous (resp. smooth) closed curve is any continuous (resp. smooth) map $\gamma \colon \mathbf{T} \to \mathbf{R}^2$. We say that the curve is non self-intersecting if the map γ is injective.

Corollary 5.7 *If $d = 2$, the set $\Gamma_1 = \{\Psi = 0\}$ consists of a countable collection of disjoint smooth, closed, non self-intersecting curves: $\Gamma_1 = \bigcup_{i \in \mathbf{N}} \gamma_i(\mathbf{T})$. Then, for any $\sigma \geq 1$ the set $\Gamma_\sigma = \{v(\Psi; \sigma, \cdot) = 0\}$, with $v(\Psi)$ as in Proposition 5.6, is the disjoint union $\Gamma_\sigma = \bigcup_{i \in \mathbf{N}} \gamma_{i,\sigma}(\mathbf{T})$ of continuous curves $\gamma_{i,\sigma}$ which are the mean curvature flow evolution of the curves γ_i, in the sense of [7].*

Proof The fact that Γ_1 is the disjoint union of smooth closed curves is proven in Lemma 5.8. Each of these curves evolves according to the level set notion of mean curvature flow, which for concreteness we denote by $\mathcal{M}_\sigma \gamma_i(\mathbf{T})$, for $i \in \mathbf{N}$ and $\sigma \geq 1$. We observe that [6, Theorem 7.3] implies that the sets $\mathcal{M}_\sigma \gamma_i(\mathbf{T})$ are pairwise disjoint over $i \in \mathbf{N}$, for any $\sigma > 1$. In addition the level set mean curvature flow evolution of a smooth closed curve coincides with the classical evolution by [6, Theorem 6.1], as long as the latter is defined. Since by Lemma 5.8 our curves are non self-intersecting, [7] proves that the classical evolution is defined for all times, and our result follows. □

Lemma 5.8 *Let Ψ be as in (1.3), in dimension $d = 2$, and consider the random set $\Gamma_1 = \{x : \Psi(x) = 0\} \subseteq \mathbf{R}^2$. Then $\mathbf{P}$−almost surely the set Γ_1 is a countable union of smooth, non self-intersecting and disjoint closed curves: $\Gamma_1 = \bigcup_{i \in \mathbf{N}} \gamma_i(\mathbf{T})$, with $\gamma_i \colon \mathbf{T} \to \mathbf{R}^2$ smooth.*

Proof By [2, Corollary 1.4] there exists a null set $\mathfrak{N}_1$ such that for all $\omega \in \mathfrak{N}_1^c$ every connected component of $\Gamma_1(\omega)$, $\mathcal{P}_1(\omega)$ and $\mathcal{N}_1(\omega)$ is bounded.

Hence the proof is complete if we show that the boundary of every region is given by a smooth closed curve. Let us first show that there exists a null set $\mathfrak{N}_2$ such that for all $\omega \in \mathfrak{N}_2^c$ it holds that

$$|\nabla\Psi(\omega, x)| \neq 0, \qquad \forall x \in \Gamma_1(\omega). \tag{5.9}$$

This follows from Bulinskaya's lemma, see for example [1, Lemma 11.2.10], as long as we can prove the following nondegeneracy condition, namely that the map

$$h \colon \Omega \times \mathbf{R}^2 \to \mathbf{R}^3, \qquad h(\omega, x) = (\nabla\Psi(\omega, x), \Psi(\omega, x)),$$

satisfies that for any $x \in \mathbf{R}^2$ the probability measure

$$\mathbf{P}_x(A) = \mathbf{P}(h(x) \in A)$$

has a density $p_x(y)$ with respect to Lebesgue measure on $\mathbf{R}^3$ such that for some $C > 0$: $|p_x(y)| \leq C$, $\forall x \in \mathbf{R}^2$, $y \in \mathbf{R}^3$. In our case this condition is trivially satisfied as for every $x \in \mathbf{R}^2$, $h(x)$ is distributed as a Gaussian vector in $\mathbf{R}^3$ with diagonal covariance matrix (here $\lambda = \frac{d-2\alpha}{2}$ and $C(\lambda) > 0$ is a constant):

$$\mathbf{E}[\partial_i \Psi(x) \partial_j \Psi(x)] = \frac{1}{(4\pi\lambda)^d (2\lambda)^2} \int_{\mathbf{R}^2} e^{-2\frac{|y|^2}{4\lambda}} y_i y_j \, \mathrm{d}y = C(\lambda) 1_{\{i=j\}},$$

$$\mathbf{E}[\partial_i \Psi(x) \Psi(x)] = \frac{1}{(4\pi\lambda)^d 2\lambda} \int_{\mathbf{R}^2} e^{-2\frac{|y|^2}{4\lambda}} y_i \, \mathrm{d}y = 0.$$

It follows that (5.9) holds true. An application of the implicit function theorem allows us to deduce our result. □

References

1. Adler, R.J., Taylor, J.E.: Random Fields and Geometry. Springer Monographs in Mathematics. Springer, New York, xviii+448 (2007)
2. Beffara, V., Gayet, D.: Percolation of random nodal lines. Publ. Math. Inst. Hautes Études Sci. **126**, 131–176 (2017). https://doi.org/10.1007/s10240-017-0093-0. arxiv:1605.08605
3. Bray, A.: Theory of phase-ordering kinetics. Adv. Phys. **43**(3), 357–459 (1994). https://doi.org/10.1080/00018739400101505. arXiv:cond-mat/9501089
4. Barles, G., Souganidis, P.E.: A new approach to front propagation problems: theory and applications. Arch. Ration. Mech. Anal. **141**(3), 237–296 (1998). https://doi.org/10.1007/s002050050077
5. Duminil-Copin, H., Rivera, A., Rodriguez, P.-F., Vanneuville, H.: Existence of an unbounded nodal hypersurface for smooth Gaussian fields in dimension d ≥ 3, (2021). arXiv:2108.08008
6. Evans, L.C., Spruck, J.: Motion of level sets by mean curvature. I. J. Differ. Geom. **33**(3), 635–681 (1991)
7. Grayson, M.A.: The heat equation shrinks embedded plane curves to round points. J. Differ. Geom. **26**(2), 285–314 (1987)
8. Ishii, H., Souganidis, P.: Generalized motion of noncompact hypersurfaces with velocity having arbitrary growth on the curvature tensor. Tohoku Math. J. (2) **47**(2), 227–250 (1995). https://doi.org/10.2748/tmj/1178225593
9. Kohn, R.V., Otto, F.: Upper bounds on coarsening rates. Commun. Math. Phys. **229**(3), 375–395 (2002). https://doi.org/10.1007/s00220-002-0693-4
10. Ohta, T., Jasnow, D., Kawasaki, K.: Universal scaling in the motion of random interfaces. Phys. Rev. Lett. **49**, 1223–1226 (1982). https://doi.org/10.1103/PhysRevLett.49.1223