# GENERATION OF FINE TRANSITION LAYERS AND THEIR DYNAMICS FOR THE STOCHASTIC ALLEN–CAHN EQUATION

M. ALFARO, D. ANTONOPOULOU, G. KARALI, AND H. MATANO

Abstract. We study an $\varepsilon$-dependent stochastic Allen–Cahn equation with a mild random noise on a bounded domain in $\mathbb{R}^n$, $n \geq 2$. Here $\varepsilon$ is a small positive parameter that represents formally the thickness of the solution interface, while the mild noise $\xi^\varepsilon(t)$ is a smooth random function of $t$ of order $\mathcal{O}(\varepsilon^{-\gamma})$ with $0 < \gamma < 1/3$ that converges to white noise as $\varepsilon \to 0^+$. We consider initial data that are independent of $\varepsilon$ satisfying some non-degeneracy conditions, and prove that steep transition layers—or interfaces—develop within a very short time of order $\varepsilon^2|\ln \varepsilon|$, which we call the "generation of interface". Next we study the motion of those transition layers and derive a stochastic motion law for the sharp interface limit as $\varepsilon \to 0^+$, which is given in terms of the mean curvature flow with additive white noise. Furthermore, we prove that the thickness of the interface for $\varepsilon$ small is indeed of order $\mathcal{O}(\varepsilon)$ and that the solution profile near the interface remains close to that of a (squeezed) travelling wave; this means that the presence of the noise does not destroy the solution profile near the interface as long as the noise is spatially uniform. Our results on the motion of interface improve the earlier results of Funaki (1999) and Weber (2010) by considerably weakening the requirements for the initial data and establishing the robustness of the solution profile near the interface that has not been known before.

## 1. Introduction

We consider a stochastic Allen–Cahn equation with a Neumann boundary condition

$$
\begin{aligned}
&\partial_t u = \Delta u + \frac{1}{\varepsilon^2} f(u) + \frac{1}{\varepsilon}\xi^\varepsilon(t), \quad t > 0, \quad x \in \Omega, \\
&\frac{\partial u}{\partial \nu} = 0, \quad t > 0, \quad x \in \partial\Omega, \\
&u(x,0) = u_0(x), \quad x \in \Omega,
\end{aligned}
\tag{1.1}
$$

where $\Omega$ is a smooth open bounded domain in $\mathbb{R}^n$ ($n \geq 2$), $\nu$ is the outward unit normal vector to $\partial\Omega$, and $\varepsilon > 0$ is a small parameter. The nonlinearity $f$ is of the *bistable* type, and the perturbation term $\xi^\varepsilon(t)$ is what we call a mild noise which is a smooth but random function of $t$ that behaves like an irregular white noise in the limit as $\varepsilon \to 0$. As mentioned in [17], such an equation can be viewed as describing intermediate (mesoscopic) level phenomena between macroscopic and microscopic ones. In such a scale, an active noise appears as a correction term to the reaction-diffusion equation when fluctuations in the hydrodynamic limit is taken into account, see [32].

Our main goal is to make a detailed analysis of the *sharp interface limit* of the problem (1.1) as $\varepsilon \to 0$. In the deterministic case where the perturbation term $\xi^\varepsilon$ is replaced by non-random, uniformly bounded smooth functions, the sharp interface limit of (1.1) is well understood: it is known that the solution $u$ typically develops steep transition layers—or *interfaces*—of thickness $\mathcal{O}(\varepsilon)$ within a very short time, which we call the *generation of interface* (or one may call it the

1991 *Mathematics Subject Classification.* 35K55, 35B25, 60H30, 60H15.

*emergence of transition layers*). Furthermore, as $\varepsilon \to 0$, those layers converge to interfaces of thickness 0 whose law of motion is given by the curvature flow with some driving force (the *propagation of interface*). See [1, 9] and the references therein for details.

In the present problem, the perturbation term $\xi^\varepsilon(t)$ is random; furthermore it is no longer uniformly bounded as $\varepsilon \to 0$, since it converges to white noise in a certain sense. This makes the analysis harder than in the classical deterministic case. Nonetheless, as we shall see, it is possible to derive a number of results that are as optimal as those established for the classical deterministic problem.

Our first result concerns the *generation of interface*. More precisely, we consider solutions of (1.1) with $\varepsilon$-independent initial data, and show that steep transition layers of thickness $\mathcal{O}(\varepsilon)$ emerge within a very short time. This thickness estimate of order $\mathcal{O}(\varepsilon)$ is the same optimal estimate known for the classical deterministic problem. Next, we discuss the *propagation of interface* and show that the thickness of the layer remains of order $\mathcal{O}(\varepsilon)$ as time passes, and that in the sharp interface limit, as $\varepsilon \to 0$, the law of motion of the interface is given by

$$V = (n-1)\kappa + c\dot{W}_t, \tag{1.2}$$

where $V$ is the inward normal velocity, $\kappa$ denotes the mean curvature, $c$ is a positive constant and $\dot{W}_t$ is a white noise. Furthermore, we also show that the *profile of the solution* near the interface is well approximated by that of a "squeezed" travelling wave. This implies that the solution profile near the interface is quite robust and is not destroyed by the random noise, as long as the noise depends only on the time variable.

The singular limit of a stochastic Allen–Cahn equation of the form (1.1) was studied by Funaki in his pioneering work [17] for two space dimensions, and later by Weber [34] for general space dimensions $n \geq 2$. Both [17, 34] derive the motion law (1.2) but only for solutions whose initial data already have well-developed $\varepsilon$-dependent transition layers. Our results improve the work of [17, 34] in three notable aspects. First, as mentioned above, our paper studies the emergence of steep transition layers (the *generation of interface*) at the very initial stage of evolution, which is not discussed in [17, 34]. Secondly, our $\mathcal{O}(\varepsilon)$ estimate of the thickness of layers is optimal and therefore, is considerably better than the order $\mathcal{O}(\varepsilon^\alpha)$ $(0 < \alpha < 1)$ estimates presented in [17, 34]. Thirdly, we show the robustness of the solution profile around the interface in the presence of noise (*rigidity of profile*), a fact that has been totally unknown before.

Concerning results on the generation of interface, let us also mention the recent papers [22, 23]. In [22], the author considers the one-dimensional case with space-time white noise and studies both the generation and motion of the interface, thus improving the work [16], which did not consider the generation of interface. However, as we shall explain in subsection 1.2, the one-dimensional case is totally different from the multi-dimensional case as the curvature effect does not appear in the former. Therefore the problems treated in [16, 22] are different from the subject of the present paper. In [23], the author considers a multi-dimensional problem under a space-time noise that is smooth in $x$. However [23] deals with only the generation of interface, thus the motion of interface under such a noise remains unknown.

**Remark 1.1.** *We also refer to the theory of stochastic viscosity solutions of Lions and Souganidis which covers a large class of stochastic fully nonlinear partial differential equations with applications to phase transitions and propagation of fronts in the presence of noise; see, for example,* [24, 25, 26, 27] *and a more recent preprint* [28]*. In* [24]*, the authors introduce the notion of "stochastic viscosity solutions" for parabolic, possibly degenerate, second-order stochastic PDEs on* $\mathbb{R}^n$*. In particular* [25, subsection 2.3] *considers the specific case of the* $\varepsilon$*-dependent Allen-Cahn equation on*

*$\mathbb{R}^n$ with the same mild noise as in* [17]*, but, unlike* [17]*, for a general initial data $u_0$ not depending on $\varepsilon$ and derives the motion law for the sharp interface limit as $\varepsilon \to 0$. In* [28]*, the authors study the asymptotic behavior of the Allen-Cahn problem perturbed by a stochastic forcing; their result is global in time and does not require any regularity assumptions on the evolving front.*

*Notice, however, that the present paper focuses on the behavior of solutions of* (1.2) *on a relatively short time-interval as proposed in* [17] *and* [13]*, see Section 2. This allows us to prove much finer properties of the convergence of* (1.1) *to* (1.2)*, namely the optimal thickness estimate of the thin layers of solutions of* (1.1) *when $\varepsilon$ is very small (see* (3.3) *of Theorem 3.1), and the proof of the robustness of the layer profile (see Theorem 3.4), neither of which can hardly be obtained through the viscosity approach.*

1.1. **Assumptions.** Let us state our standing assumptions in the present paper. The nonlinearity is given by $f(u) := -W'(u)$, where $W(u)$ is a double-well potential with equal well-depth, taking its global minimum value at $u = a_\pm$. More specifically, we assume that

$$f \text{ is } C^2 \text{ and has exactly three zeros } a_- < a < a_+, \tag{1.3}$$

$$f'(a_\pm) < 0, \quad f'(a) > 0, \tag{1.4}$$

and

$$\int_{a_-}^{a_+} f(u)du = 0. \tag{1.5}$$

This last assumption (1.5) makes $f$ a *balanced* bistable nonlinearity. We shall use this assumption only in Section 5, where we study the propagation of interface. No such assumption is needed for the emergence of interface, which we discuss in Section 4.

Concerning the initial data, we assume that $u_0 \in C^2(\overline{\Omega})$, and define

$$C_0 := \|u_0\|_{C^0(\overline{\Omega})} + \|\nabla u_0\|_{C^0(\overline{\Omega})} + \|\Delta u_0\|_{C^0(\overline{\Omega})}. \tag{1.6}$$

The initial interface is defined by

$$\Gamma_0 := \{x \in \Omega : \ u_0(x) = a\}. \tag{1.7}$$

We assume that $\Gamma_0 \subset\subset \Omega$ is a $C^{2,\alpha}$ $(0 < \alpha < 1)$ hypersurface without boundary. Let $\Omega_0$ denote the region enclosed by $\Gamma_0$. Without loss of generality, we may assume that

$$u_0(x) < a \text{ for any } x \in \Omega_0 \text{ and } u_0(x) > a \text{ for any } x \in \Omega \setminus \overline{\Omega_0}. \tag{1.8}$$

We also assume that

$$\nabla u_0(x) \cdot n(x) \neq 0 \text{ for any } x \in \Gamma_0, \tag{1.9}$$

where $n = n(x)$ denotes the outward normal unit vector at $x \in \Gamma_0 = \partial\Omega_0$.

As regards the perturbation term $\xi^\varepsilon(t)$, we shall consider two types of mild noises:

- **Type (MN1)** This is proposed in [17]. $\xi^\varepsilon(t) := \varepsilon^{-\gamma}\xi(\varepsilon^{-2\gamma}t)$, where $\xi(t)$ is a smooth stochastic process that is stationary and strongly mixing in a certain sense. Thus $\xi^\varepsilon(t)$ converges to white noise as $\varepsilon \to 0$ (convergence in probability).
- **Type (MN2)** This is proposed in [34]. $\xi^\varepsilon(t) = \dot{W}^\varepsilon(t)$, where $W^\varepsilon(t)$ is a mollified Brownian motion. Thus $\xi^\varepsilon(t) \to \dot{W}(t)$ as $\varepsilon \to 0$ (almost sure convergence).

Now we give more precise definition of these mild noises.

*Mild noise of type (MN1)*

Following Funaki [17], we consider a mild noise $\xi^\varepsilon$ given in the form

$$\xi^\varepsilon(t) := \varepsilon^{-\gamma_1}\xi(\varepsilon^{-2\gamma_1}t), \quad t \geq 0, \tag{1.10}$$

where $\gamma_1$ is a constant satisfying

$$0 < \gamma_1 < \frac{1}{3}, \tag{1.11}$$

and $\xi(t) = \xi_t$ is a stochastic process in $t$ that is stationary and strongly mixing. More specifically, let $F_{\xi_{t_1+\tau},\dots,\xi_{t_k+\tau}}$ be the distribution function of the $k$ random variables $\xi_{t_1+\tau},\dots,\xi_{t_k+\tau}$, then the stochastic process $\xi_t$ is called *stationary* if for all $k$, $\tau$ and for all $t_1,\dots,t_k$

$$F_{\xi_{t_1+\tau},\dots,\xi_{t_k+\tau}} = F_{\xi_{t_1},\dots,\xi_{t_k}}.$$

Let $(\Omega_{prob},\mathcal{F},\mathbb{P})$ be the probability space where $\xi_t$ is realized, with $\mathcal{F} := \sigma(\xi_r : \ 0 \leq r < +\infty)$ the $\sigma$-algebra generated by $\xi_r$ for $0 \leq r < +\infty$, and $\mathbb{P}$ the probability measure. Then $\mathcal{F}_{s,t} := \sigma(\xi_r : s \leq r \leq t)$ is the subalgebra of $\mathcal{F}$ generated by $\xi_r$ for $s \leq r \leq t$. We assume that the process $\xi_t$ is strongly mixing in the following sense: the mixing rate $\rho(t)$ defined by

$$\rho(t) := \sup_{s\geq 0} \sup_{A\in\mathcal{F}_{s+t,\infty},\, B\in\mathcal{F}_{0,s}} |\mathbb{P}(A\cap B) - \mathbb{P}(A)\mathbb{P}(B)|/\mathbb{P}(B), \quad t \geq 0,$$

satisfies

$$\int_0^\infty \rho(t)^{1/p}dt < +\infty \quad \text{ for some } \ p > 3/2.$$

In Funaki [17], this last condition is used to derive some estimates that are uniform in $\varepsilon$; see the proof of Proposition 4.1 and Lemma 5.3 in [17].

Furthermore, it is assumed that $t \mapsto \xi(t)$ is $C^1$ almost surely,

$$|\xi(t)| \leq M, \quad |\dot\xi(t)| \leq M, \quad E[\xi(t)] = 0,$$

for some deterministic constant $M$, with $\dot\xi := \frac{d\xi}{dt}$. Obviously, the above implies that

$$t \mapsto \xi^\varepsilon(t) \text{ is } C^1 \text{ almost surely,}$$

and that

$$|\xi^\varepsilon(t)| \leq M\varepsilon^{-\gamma_1}, \quad |\dot{\xi^\varepsilon}(t)| \leq M\varepsilon^{-3\gamma_1}. \tag{1.12}$$

In Funaki [17], these conditions are used to justify the limit interface equation (1.2) as $\varepsilon \to 0$, but, as we shall see, the estimate (1.12) will also be fundamental for our analysis of the initial formation of layers (the generation of interface).

Notice that the coefficient $\varepsilon^{-\gamma_1}$ in the definition (1.10) implies that $\xi^\varepsilon(t)$ is unbounded as $\varepsilon \to 0$. As shown in [17], $\xi^\varepsilon(t)$ converges to an irregular white noise as $\varepsilon \to 0$ in a certain sense.

*Second type of noise (MN2)*

Following Weber [34], we define the mild noise $\xi^\varepsilon(t) = \xi^\varepsilon_t$ as the derivative of a mollified Brownian motion. More precisely, let $W(t)$, $t \geq 0$, be a Brownian motion defined on the space $(\Omega_{prob},\mathcal{F},\mathbb{P})$. (Here, as usual, the dependence of $W$ on the sample points $\omega \in \Omega_{prob}$ is not shown explicitly.) For technical reasons, $W(t)$ is extended over $\mathbb{R}$ by considering an independent Brownian motion

$\widetilde{W}(t)$, $t \geq 0$, and setting $W(t) = \widetilde{W}(-t)$ for $t < 0$. Then $W(t)$, $t \in \mathbb{R}$, is a Gaussian process, with independent stationary increments and a distinguished point $W(0) = 0$ almost surely. Also, let $\rho : \mathbb{R} \to \mathbb{R}_+$ be a mollifying smooth and symmetric kernel, with $\rho = 0$ outside $[-1, 1]$ and $\int_{\mathbb{R}} \rho = 1$. The approximated Brownian motion $W^\varepsilon(t), t \geq 0$, is defined as usual by

$$W^\varepsilon(t) := W * \rho^\varepsilon(t) := \int_{-\infty}^{\infty} \rho^\varepsilon(t-s) W(s) ds, \tag{1.13}$$

where $\rho^\varepsilon(\tau) := \varepsilon^{-\gamma_2} \rho(\varepsilon^{-\gamma_2} \tau)$ for some constant $\gamma_2$ satisfying

$$0 < \gamma_2 < \frac{2}{3}. \tag{1.14}$$

Note that the Brownian motion for negative times is needed only in the expression (1.13), so only the negative times in $(-\varepsilon^{\gamma_2}, 0]$ will play a role. The constant $\gamma_2$ determines how quickly $W^\varepsilon$ converges to the true integrated white noise as $\varepsilon \to 0$. Since $W(t)$ is Hölder continuous almost surely, $W^\varepsilon(t)$ is a smooth function of $t$ almost surely. The noise $\xi^\varepsilon(t)$ is then defined as the derivative of $W^\varepsilon(t)$:

$$\xi^\varepsilon(t) = \dot{W}^\varepsilon(t). \tag{1.15}$$

In [34, Propositions 1.2 and 1.3], the author derives estimates for $\xi^\varepsilon(t)$ and its derivative $\dot{\xi}^\varepsilon(t)$ in the form

$$|\xi^\varepsilon(t)| \leq M \varepsilon^{-\tilde{\gamma}/2} \quad \text{a.s.}, \qquad |\dot{\xi}^\varepsilon(t)| \leq C \varepsilon^{-3\tilde{\gamma}/2} \quad \text{a.s.} \qquad (\gamma_2 < {}^\forall \tilde{\gamma} < 2/3),$$

for some deterministic constant $M$ (as in the definition of (MN1) by Funaki [17]) but random constant $C = C(\omega)$ depending on the realization, by using Lévy's well-known result on the modulus of continuity of Brownian motion:

$$\mathbb{P}\left[ \limsup_{\delta \to 0} \frac{1}{g(\delta)} \max_{\substack{0 \leq s < t \leq T \\ t-s \leq \delta}} |W(t) - W(s)| = 1 \right] = 1,$$

where the modulus of continuity is given by $g(\delta) = \sqrt{2\delta \log(\frac{1}{\delta})}$. Actually the very same argument as in [34] gives the following slightly more refined estimates, whose proof is omitted as it is straightforward—roughly speaking it suffices to set $\delta = \varepsilon^{\gamma_2}$ in $g(\delta)$.

**Proposition 1.2** (Estimates of the noise term)**.** *For any $T > 0$, there exist a non-random constant $M > 0$, a random constant $C = C(\omega)$, and (random) $\varepsilon_0 > 0$ such that, for all $0 < \varepsilon \leq \varepsilon_0$ and all $0 \leq t \leq T$,*

$$|\xi^\varepsilon(t)| \leq M \varepsilon^{-\gamma_2/2} |\log \varepsilon|^{1/2} \quad \text{a.s.}, \quad |\dot{\xi}^\varepsilon(t)| \leq C \varepsilon^{-3\gamma_2/2} |\log \varepsilon|^{1/2} \quad \text{a.s.} \tag{1.16}$$

This is an analogue of (1.12) and will be fundamental in the study of the emergence of interface.

**Remark 1.3.** *The two types of mild noises, (MN1) and (MN2), are smooth stochastic processes, and they both satisfy similar estimates* (1.12) *and* (1.16)*. However, there are notable differences between the two. In particular, (MN1) is not necessarily a Gaussian noise but is more general, and all the convergence results in* [17] *as well as those in the present paper for the (MN1) type noises can only hold in law; see Theorem 3.2. On the other hand, stronger almost sure convergence results hold for the (MN2) type noises as shown in* [34] *and also in Theorem 3.3 of the present paper.*

**Remark 1.4.** *In this paper, the notations for order estimates such as $\mathcal{O}(\varepsilon)$, $o(\varepsilon)$, $\mathcal{O}(\varepsilon^\alpha)$, etc, generally represent path-wise estimates, therefore the constants involved in these estimates may be random, that is, they depend on each realization.*

1.2. **Deterministic and stochastic Allen–Cahn equations.** It has long been known among physicists that the interface motion of Allen–Cahn type equations (namely, reaction-diffusion equations with balanced or nearly balanced bistable nonlinearities) is well approximated by the mean curvature flow or its variant. A pioneering study was done by Allen and Cahn [3], who proposed what is now known as the Allen–Cahn equation as a model for the dynamics of crystal structures in alloys and derived from it a curvature-dependent motion law of interfaces by formal analysis. Another early pioneering work in this direction is found in [21].

Mathematically, this problem can be formulated as follows. Consider the following reaction-diffusion equation, where $f$ is a balanced bistable nonlinearity (with $u = a_\pm$ being the two stable zeros) and $\varepsilon$ is a small parameter, while the term $g^\varepsilon$ represents a small perturbation:

$$\partial_t u = \Delta u + \frac{1}{\varepsilon^2}\left(f(u) + \varepsilon g^\varepsilon(x,t)\right). \tag{1.17}$$

Then the solution quickly develops a transition layer (or interface) between the regions where $u \approx a_-$ and $u \approx a_+$. Furthermore, a formal analysis shows that the thickness of this transition layer is of order $\mathcal{O}(\varepsilon)$ and that the position of this transition layer for $\varepsilon$ sufficiently small is well approximated by a time-dependent hypersurface $\Gamma_t^\varepsilon$ whose motion law is given by

$$V = (n-1)\kappa - \frac{1}{\varepsilon}\, c\left(\varepsilon g^\varepsilon(x,t)\right) \quad \text{on } \ \Gamma_t^\varepsilon, \ t > 0, \tag{1.18}$$

where $V$ denotes the normal velocity of $\Gamma_t^\varepsilon$, $\kappa$ the mean curvature, $n$ the space dimension, and $c(\delta)$, $\delta > 0$, denotes the speed of the travelling wave for the one-dimensional diffusion equation

$$\partial_t u = \partial_{zz} u + f(u) + \delta \qquad (z \in \mathbb{R}, \ t \in \mathbb{R}). \tag{1.19}$$

Consequently, as $\varepsilon \to 0$, the transition layer is expected to converge to a hypersurface $\Gamma_t$ of thickness 0 (the *sharp interface limit*) whose motion is governed by the equation

$$V = (n-1)\kappa + c(x,t) \quad \text{on } \ \Gamma_t, \ t > 0, \tag{1.20}$$

where $c(x,t)$ is the limit of the term $-\varepsilon^{-1} c\left(\varepsilon g^\varepsilon(x,t)\right)$ as $\varepsilon \to 0$. Since around 1990, many mathematicians started to give rigorous justification of the above formal scenario. As there is a huge literature, we shall name only a few.

First, in the case of deterministic problems, the perturbation term $g^\varepsilon$ in (1.17) is typically a smooth function that remains uniformly bounded as $\varepsilon \to 0$ (hence the term $c(x,t)$ in (1.20) is also bounded). The sharp-interface limit of (1.17) for $n = 1$ is studied in [8, 10]. Note that the curvature effect does not appear if $n = 1$, therefore this case is in a regime totally different from the case $n \geq 2$. For the multi-dimensional case $n \geq 2$, we refer, among others, to [7], [9], [30, 31], [1]. The above works study both the *generation of interface* at the early stage and the *motion of interface* at the later stage.

Next, in the stochastic case where $g^\varepsilon$ is a random noise, additional difficulties appear. First difficulty concerns the well-posedness. If $g^\varepsilon$ is a space-time white noise, then (1.17) is well-posed if $n = 1$ (see [14, 16]) but ill-posed if $n \geq 2$ (see [33, 12, 18]). Thus, if $n \geq 2$, the noise term $g^\varepsilon$ had better be $x$-independent, or at least a colored noise. The second difficulty is the interpretation of the limit equation (1.20), as the term $c(x,t)$ is typically replaced by white noise:

$$V = (n-1)\kappa + c\dot{W}_t \quad \text{on } \ \Gamma_t, \ t > 0. \tag{1.21}$$

Systematic analysis of equations of this type is found in [36, 13]. The third difficulty is in the rigorous derivation of the sharp-interface limit of (1.17). In addition to randomness, the noise term $g^\varepsilon$ is unbounded, or at least unbounded as $\varepsilon \to 0$, which adds extra technical difficulty.

As regards the case $n = 1$, the sharp-interface limit of (1.17), with $g^\varepsilon$ being white noise, is studied in [16], [6] for solutions with initial data close to a Heaviside function. The paper [35] studies the same one-dimensional problem via an invariant measure approach. As there is no curvature effect in (1.21) if $n = 1$, the dynamics is totally different from the case $n \geq 2$.

As for the case $n \geq 2$, the aforementioned work of Funaki [17] considers (1.1) with a mild noise $\xi^\varepsilon(t)$ (or equivalently $g^\varepsilon = \xi^\varepsilon(t)$ in (1.17)) and proves that its sharp-interface limit is (1.21), provided that $n = 2$ and $\Gamma_t$ is strictly convex. (See also [18, Sec.4.2].) This result is extended in [34] to general dimension $n \geq 2$ without the convexity restriction. Note that these results are obtained only for solutions whose initial data already have well-developed transition layers.

As we already mentioned, there is also a large literature on the sharp-interface limit of (1.17) or more general equations in the framework of viscosity solutions. This is a different approach from ours. For more details, see Remark 1.1 and the works listed therein.

1.3. **Organization of the present paper.** As mentioned above, Funaki [17] and Weber [34] proved that the sharp-interface limit of the stochastic Allen–Cahn equation (1.1) is given by (1.21) for a special class of initial data that have well-developed transition layers of $\mathcal{O}(\varepsilon^\alpha)$ thickness $(0 < \alpha < 1)$. The objective of the present paper is to show that the same results hold for a large class of solutions whose initial data do not depend on $\varepsilon$ and to obtain finer estimates for the convergence process. More precisely, we shall show the following:

(a) Generation of interface: Under rather mild non-degeneracy conditions on the initial data $u_0$, we show that the solution of (1.1) develops a transition layer of $\mathcal{O}(\varepsilon)$ thickness within a very short time of order $\mathcal{O}(\varepsilon^2|\ln\varepsilon|)$.
(b) Motion of interface: We prove that the transition layer of the solution $u^\varepsilon$ of (1.1) keep the optimal $\mathcal{O}(\varepsilon)$ thickness during the time evolution and we obtain fine estimates for the distance between the transition layer of $u^\varepsilon$ and the solution of (1.21).
(c) Rigidity of profile: We prove that the profile of the solution $u^\varepsilon$ around its transition layer converges to that of the squeezed travelling wave as $\varepsilon \to 0$.

A more precise account of statement (c) is given in (3.10). It implies that the layer profile is quite robust and is not destroyed by the mild noise $\xi^\varepsilon(t)$ despite its unboundedness as $\varepsilon \to 0$.

The present paper is organized as follows. In Section 2, we give precise definition of the limit motion law (1.21), the one given by [17] and the other by [34], separately.

In Section 3, we state our main results. Theorem 3.1 covers statements (a) and (b) above, both for the (MN1) and the (MN2) type mild noises. Letting $\varepsilon \to 0$ in statement (b) above, we immediately see that the sharp interface limit of (1.1) is (1.21), thus confirming that the same results as in [17] and [34] hold for our general class of solutions (Theorems 3.2 and 3.3). Finally, Theorem 3.4 asserts the statement (c) above.

In Section 4, we prove the first part of Theorem 3.1 on statement (a) above (Theorem 4.1). The method of the proof is basically the same as in [1] for the deterministic problem. However, since $\xi^\varepsilon(0) = \mathcal{O}(\varepsilon^{-\gamma_1})$ is generally unbounded as $\varepsilon \to 0$, initial sudden shift of order $\mathcal{O}(\varepsilon^{1-\gamma_1})$ can occur, while in the deterministic case, the initial shift is of order $\mathcal{O}(\varepsilon)$. We therefore have to handle this problem more carefully.

In Sections 5 and 6, we prove the second part of Theorem 3.1 concerning statement (b) above. As $\xi^\varepsilon(t)$ is unbounded as $\varepsilon \to 0$, we cannot use the same method as in [1]. We shall employ a different, more efficient pair of super- and subsolutions for the proof. Section 5 is devoted to the construction of such a pair, and we complete the proof of Theorem 3.1 in Section 6.

In Section 7, we prove Theorem 3.4 concerning statement (c) above. The method of proof is the same as in [2]. More precisely, we use rescaling arguments and the Liouville type theorem for entire solutions of the Allen–Cahn type equations on $\mathbb{R}^n$.

## 2. On stochastic motion by mean curvature

As mentioned briefly in subsection 1.2, our strategy for establishing the motion law (1.2) for the sharp-interface limit of (1.1) consists of the following two steps:

Step 1 In view of the fact that the motion of the transition layer arising in the deterministic problem (1.17) is well approximated by the interface equation (1.18) so long as $g^\varepsilon(x,t)$ remains bounded as $\varepsilon \to 0$, we anticipate that the motion of the transition layer of (1.1) is also well approximated by the following equation despite the unboundedness of $\xi^\varepsilon$ as $\varepsilon \to 0$:

$$V = (n-1)\kappa - \frac{c(\varepsilon\xi^\varepsilon(t))}{\varepsilon}, \tag{2.1}$$

where $c(\delta)$, $\delta > 0$, denotes the speed of the travelling wave for the one-dimensional equation (1.19). As in the deterministic problems as well as in [17] and [34], the validity of (2.1) will be shown by the super-subsolution method. In order to obtain our optimal $\mathcal{O}(\varepsilon)$ error estimate, we shall need to choose an appropriate pair of super-subsolutions as shown in subsection 5.3. Because of the unboundedness of $\xi^\varepsilon(t)$, we need to take an approach that is different from the one in [1].

Step 2 It is known that the above $c(\delta)$ is a smooth function with $c(0) = 0$ and $\partial_\delta c(0) = -c_0$ for some $c_0 > 0$. Therefore, $-\varepsilon^{-1}c(\varepsilon\xi^\varepsilon(t)) \approx c_0\xi^\varepsilon(t)$ for sufficiently small $\varepsilon > 0$, that is,

$$V \approx (n-1)\kappa + c_0\xi^\varepsilon(t).$$

Since $\xi^\varepsilon(t)$ converges to a white noise in a certain sense, one can formally expect that the above interface equation converges to an equation of the form (1.2) as $\varepsilon \to 0$.

In order to prove the above claim in Step 2 rigorously, we first need to give a precise definition of the motion law (1.2) for the limit interface. The interpretation of this motion law differs between the results for the (MN1) type noise and those for the (MN2) type noise. Roughly speaking, the former converts (1.2) into an SPDE for the curvature, while the latter employs the level-set approach for (1.2). In what follows, we explain these in more detail.

### 2.1. Motion law for the (MN1) type noise (for $n = 2$).

In this subsection we adopt the interpretation of the motion law (1.2) given by Funaki [17]. Note that this interpretation is valid as long as the random curve $\Gamma_t$ remains strictly convex and does not touch the boundary $\partial\Omega$.

It is shown in [17] that the sharp-interface limit of (1.1) is given by

$$V = \kappa + c_0\alpha_0\dot{W}_t \quad \text{on} \quad \Gamma_t, \tag{2.2}$$

at least in some restrictive sense, where $c_0, \alpha_0$ are positive constants that are to be specified in (5.4) and (3.4). The derivation of (2.2) is done via equation (2.1) with $n = 2$. The question here is how to interpret equation (2.2).

Before explaining the idea of [17], let us first recall the general procedure of parametrizing a strictly convex closed curve $\Gamma$ in terms of the Gauss map $\Gamma \to [0, 2\pi)$. The position $x$ on $\Gamma$ is denoted by $x(\theta)$ with $\theta \in [0, 2\pi)$ if the angle between the fixed direction $\boldsymbol{e}_1 = (1,0)$ and the outward normal $n(x)$ at $x$ to $\Gamma$ equals $\theta$. Denote by $\kappa = \kappa(\theta) > 0$ the (inward) curvature of $\Gamma$ at

$x = x(\theta)$. Since $\theta$ denotes the normal angle, we have $dx_1 = -\sin\theta\, ds$, $dx_2 = \cos\theta\, ds$ along the curve $\Gamma$, where $s$ denotes the arclength parameter. This, together with $d\theta/ds = \kappa$, implies

$$dx_1 = -\frac{\sin\theta}{\kappa}\, d\theta, \qquad dx_2 = \frac{\cos\theta}{\kappa}\, d\theta.$$

Consequently, we have

$$x_1(\theta) = x_1(0) - \int_0^\theta \frac{\sin\theta'}{\kappa(\theta')}\, d\theta', \qquad x_2(\theta) = x_2(0) + \int_0^\theta \frac{\cos\theta'}{\kappa(\theta')}\, d\theta'. \tag{2.3}$$

In particular, if $\Gamma$ is a closed curve, it holds that

$$\int_0^{2\pi} \frac{\sin\theta}{\kappa(\theta)}\, d\theta = \int_0^{2\pi} \frac{\cos\theta}{\kappa(\theta)}\, d\theta = 0. \tag{2.4}$$

Conversely, if a positive function $\kappa(\theta)$ satisfying (2.4) is given, then (2.3) defines a closed convex curve $\Gamma$ whose curvature coincides with $\kappa(\theta)$. Such a curve is unique up to translation.

Next suppose that $\Gamma_t$ is a $t$-dependent strictly convex curve, and let $\kappa(\theta, t)$ be the curvature of $\Gamma_t$ as defined above. Then it is well known that $\kappa(\theta, t)$ satisfies the following equation:

$$\partial_t \kappa = \kappa^2 \left(\partial_{\theta\theta} V + V\right), \tag{2.5}$$

where $V$ denotes the inward normal velocity of the curve $\Gamma_t$. See [4] for the derivation of (2.5).

Once the solution $\kappa(\theta, t)$ of (2.5) is obtained, one can recover the curve $\Gamma_t$ by combining (2.3) with the following formula, which determines the position of $x(\theta, t)$ at $\theta = 0$:

$$\frac{d}{dt} x_1(0, t) = -V(0, t), \qquad \frac{d}{dt} x_2(0, t) = -V_\theta(0, t).$$

We omit the proof of the above formula as it is elementary. The negative sign in front of $V$ comes from the fact that $V$ denotes the inward normal velocity of the curve. This and (2.3) yield:

$$\begin{aligned} x_1(\theta, t) &= x_1(0, 0) - \int_0^t V(0, \tau) d\tau - \int_0^\theta \frac{\sin\theta'}{\kappa(\theta', t)}\, d\theta', \\ x_2(\theta, t) &= x_2(0, 0) - \int_0^t V_\theta(0, \tau) d\tau + \int_0^\theta \frac{\cos\theta'}{\kappa(\theta', t)}\, d\theta'. \end{aligned} \tag{2.6}$$

Now we explain the idea in [17]. Substituting (2.2) *formally* into (2.5), we obtain the following nonlinear stochastic partial differential equation for $\kappa = \kappa(\theta, t)$:

$$\partial_t \kappa = \kappa^2 \partial_{\theta\theta} \kappa + \kappa^3 + c_0 \alpha_0 \kappa^2 \circ \dot{W}_t, \quad \theta \in [0, 2\pi),\ 0 < t < \sigma, \tag{2.7}$$

where $\circ$ denotes the Stratonovich stochastic integral, and the stopping time $\sigma$ is to be specified below. Once the curvature $\kappa(\theta, t)$ is obtained via (2.7), one can determine a $t$-dependent curve $\Gamma_t = \{x_t(\theta) \in \mathbb{R}^2 \cong \mathbb{C},\ \theta \in [0, 2\pi)\}$ uniquely by substituting (2.2) into (2.6), or, equivalently, by the formula [17, (1.10)], and we define the solution of (2.2) by this curve $\Gamma_t$. Finally the stopping time $\sigma$ is given by $\sigma = \lim_{N\to\infty} \sigma_N$, where

$$\sigma_N := \inf\{t > 0,\ \bar{\kappa}_t > N \text{ or } \mathrm{dist}(\Gamma_t, \partial\Omega) < 1/N\}, \quad N > 0,$$

where $\bar{\kappa}_t = \max_{\theta\in[0,2\pi)} \max\{\kappa(\theta, t), \kappa^{-1}(\theta, t), |\partial_\theta \kappa(\theta, t)|\}$.

Next we consider approximation of this motion for small $\varepsilon > 0$. As explained in Step 1 above, we make an ansatz that the motion of the transition layer of (1.1) is well approximated by the interface equation (2.1) with $n = 2$, namely

$$V = \kappa - \frac{c(\varepsilon\xi^\varepsilon(t))}{\varepsilon}, \tag{2.8}$$

where $c(\delta)$, $\delta > 0$, denotes the speed of the travelling wave for the one-dimensional equation (1.19) (see (5.2)). Substituting (2.8) into (2.5) yields the following equation which is equivalent to (2.8) so long as the curve is smooth and strictly convex:

$$\partial_t \kappa = \kappa^2 \partial_{\theta\theta}\kappa + \kappa^3 - \frac{c(\varepsilon\xi^\varepsilon(t))}{\varepsilon}\,\kappa \quad \theta \in [0, 2\pi),\ 0 < t < \sigma^\varepsilon, \tag{2.9}$$

where $\sigma^\varepsilon$ denotes the stopping time that is to be specified below.

The validity of (2.8) will be shown as follows. We consider a family of hypersurfaces $(\gamma_t^\varepsilon)_{0\le t<\sigma^\varepsilon}$ starting from a certain cruve $\gamma_0^\varepsilon$ that satisfies

$$\gamma_0^\varepsilon \to \Gamma_0 \ \text{ as } \varepsilon \to 0, \ \text{ in the } C^{2,\alpha} \text{ sense} \tag{2.10}$$

and evolving with the law (2.8). We then construct a super- and a subsolution $u_\varepsilon^+, u_\varepsilon^-$ for (1.1) to show that the position of the transition layer of the actual solution $u^\varepsilon$ (which we shall call $\Gamma_t^\varepsilon$ later) is confined within an $\mathcal{O}(\varepsilon)$ neighbourhood of $\gamma_t^\varepsilon$ for $0 \le t < \sigma^\varepsilon$. This is how the validity of (2.8) will be proved. As for the stopping time, we set $\sigma^\varepsilon = \lim_{N\to\infty} \sigma_N^\varepsilon$, where

$$\sigma_N^\varepsilon := \inf\{t > 0, \quad \bar\kappa_t^\varepsilon > N \text{ or } \operatorname{dist}(\gamma_t^\varepsilon, \partial\Omega) < 1/N\}, \quad N > 0, \tag{2.11}$$

and

$$\bar\kappa_t^\varepsilon := \max_{\theta\in[0,2\pi)} \max\{\kappa^\varepsilon(\theta,t), (\kappa^\varepsilon)^{-1}(\theta,t), |\partial_\theta\kappa^\varepsilon(\theta,t)|\}, \tag{2.12}$$

where $\kappa^\varepsilon$ is the curvature of $\gamma_t^\varepsilon$. Thus the smoothness of the curve $\gamma_t^\varepsilon$ and hence the validity of (2.9) are guaranteed for $0 \le t < \sigma^\varepsilon$.

Finally, the derivation of the motion law (2.8) in the sharp-interface limit as $\varepsilon \to 0$ will be done in the framework of the curvature equations (2.7) and (2.9). More precisely, let $u_\varepsilon^+, u_\varepsilon^+$ be the aforementioned super- and subsolutions of (1.1) that traps the actual solution $u^\varepsilon$ in between. Denote by $\gamma_t^{\varepsilon,+}, \gamma_t^{\varepsilon,-}$ the interface (i.e., the position of the transition layer) of $u_\varepsilon^+, u_\varepsilon^+$, respectively. Then the interface $\Gamma_t^\varepsilon$ of the actual solution $u^\varepsilon$, as well as the curve $\gamma_t^\varepsilon$, is trapped between $\gamma_t^{\varepsilon,+}$ and $\gamma_t^{\varepsilon,-}$. The equations of motion for the curves $\gamma_t^{\varepsilon,+}, \gamma_t^{\varepsilon,-}$ are very close to (2.8) with very small (explicit) perturbation terms that go to 0 as $\varepsilon \to 0$. Consequently, the curvature equations corresponding to $\gamma_t^{\varepsilon,+}, \gamma_t^{\varepsilon,-}$ are also very close to (2.9).

Since $-\varepsilon^{-1}c(\varepsilon\xi^\varepsilon(t)) \sim c_0\,\xi^\varepsilon(t)$ for sufficiently small $\varepsilon > 0$ as mentioned in Step 2 above, and since $\xi^\varepsilon(t)$ converges to $\alpha_0 \dot W_t$ in law, one can expect that the solution $\kappa_t^\varepsilon$ of (2.9) and also the solutions of the modified equations corresponding to $\gamma_t^{\varepsilon,+}, \gamma_t^{\varepsilon,-}$ all converge in law to the solution of (2.7) as $\varepsilon \to 0$. In [17] such convergence is justified by using the so-called martingale method for the case $\gamma_0^\varepsilon = \Gamma_0$. In order to obtain a more refined estimate as in our Theorem 3.2, we have to choose $\gamma_0^\varepsilon$ that is slightly different from $\Gamma_0$, therefore we shall need extra arguments to deal with this problem. We note that, for all but countable many $N > 0$, we have $\sigma_N^\varepsilon \to \sigma_N$ as $\varepsilon \to 0$.

In view of the fact that the actual interface $\Gamma_t^\varepsilon$ is trapped between $\gamma_t^{\varepsilon,+}$ and $\gamma_t^{\varepsilon,-}$, whose spatial gap tends to 0 and whose motion laws both converge to (2.2) via (2.7), we can regard (2.2) as the law of motion in the sharp-interface limit of (1.1).

2.2. **Motion law for the (MN2) type noise.** Here we consider a general spatial dimension $n \geq 2$ and do not require that the initial interface $\Gamma_0$ be strictly convex. The precise meaning of the motion law (1.21) in the context of the (MN2) type noise can be clarified by using the results of Dirr, Luckhaus and Novaga [13], which are based on the level-set methods. In this subsection we summarize their results and apply them to our problem. More precisely, we refer to [13, Theorem 3.1] for the existence of solution hypersurfaces and to [13, Corollary 4.2] for the estimate of the deviation from the original problem, when the white noise is smoothly approximated and when the initial hypersurface is slightly shifted.

Let $c_0 > 0$ be a given constant (which will be taken as in (5.4) in our context). Since the initial hypersurface $\Gamma_0 = \partial\Omega_0$ is of class $C^{2,\alpha}$, there is a positive stopping time $\tau = \tau(\Gamma_0) = \tau(\omega, \Gamma_0)$ depending on the $C^{2,\alpha}$-norm of $\Gamma_0$, and a family of hypersurfaces $(\Gamma_t)_{0\leq t<\tau} = (\Gamma_t(\omega))_{0\leq t<\tau(\omega,\Gamma_0)}$ of class $C^{2,\alpha}$, such that, for any $X_0 \in \Gamma_0$, there is a process $X(\cdot)$ with $X(t) = X(t,\omega) \in \Gamma_t = \Gamma_t(\omega)$ for almost all $\omega \in \Omega_{prob}$ which solves the Itô equation

$$dX = \nu(X(t,\omega),t)(n-1)\kappa(X(t,\omega),t)dt + \nu(X(t,\omega),t)c_0 dW, \quad X(0) = X_0, \tag{2.13}$$

where $\kappa(y,t)$ and $\nu(y,t)$ are respectively the mean curvature and the inner normal at $y \in \Gamma_t$. This is the sense we adopt for the motion law

$$V = (n-1)\kappa + c_0 \dot{W}_t, \tag{2.14}$$

or $dV = (n-1)\kappa dt + c_0 dW_t$, which we call the stochastic motion by mean curvature.

Next we consider approximations of this motion as follows. As we have done for the case of the (MN1) type noise, we consider an approximate equation (2.1). We then consider a family of hypersurfaces $(\gamma_t^\varepsilon)_{0\leq t<\tau^\varepsilon}$ that starts from a certain hypersurface $\gamma_0^\varepsilon$ satisfying (2.10) (which will be specified as in (4.5) in our context) and evolving with the law (2.1). As before, by constructing a pair of super-subsolutions $u_\varepsilon^+, u_\varepsilon^-$ of (1.1), we can show that the interface $\Gamma_t^\varepsilon$ of the actual solution $u^\varepsilon$ is trapped within a small neighbourhood of $\gamma_t^\varepsilon$.

Since the mild noise $\xi^\varepsilon(t)$ is defined as the derivative of a mollified Brownian motion $W^\varepsilon(t)$ and thanks to the above assumptions, we have that, for any $T > 0$, the random functions $t \mapsto \int_0^t -\frac{c(\varepsilon\xi^\varepsilon(s))}{\varepsilon} ds$ converge almost surely to $t \mapsto c_0 W(t)$ in $C^{0,\alpha}([0,T])$ for any $0 < \alpha < \frac{1}{2}$ (see [34, Lemma 3.3]). Hence, by [13, Corollary 4.2], there is a time $T > 0$ such that

$$\sup_{0\leq t\leq T} \|d(t,x) - d^\varepsilon(t,x)\|_{C^{2,\alpha}} \to 0, \quad \text{as } \varepsilon \to 0, \tag{2.15}$$

where $d(\cdot,t)$, $d^\varepsilon(\cdot,t)$ denote the signed distance functions to $\Gamma_t$, $\gamma_t^\varepsilon$, respectively. This, together with the fact that $\gamma_t^\varepsilon$ is a good approximation of the actual interface $\Gamma_t^\varepsilon$, shows that $\Gamma_t^\varepsilon$ converges to $\Gamma_t$, whose law of motion is given by (2.14) via (2.13).

## 3. Main results

We first remark that, for each given realization, the mild noise $\xi^\varepsilon(t)$ is a smooth function of $t$, therefore the solution $u^\varepsilon(x,t)$ of (1.1) is classical. Throughout this paper, the calculations involving $u^\varepsilon$ are done in the standard point-wise sense, thus they hold true almost surely.

Our first main result, Theorem 3.1, is on the rapid formation of a transition layer of $\mathcal{O}(\varepsilon)$ thickness at the initial stage, and on the thickness and the location of this transition later at the later stage. More precisely, the theorem asserts that the transition layer keeps the optimal $\mathcal{O}(\varepsilon)$ thickness during the time evolution and that it is located within an $\mathcal{O}(\varepsilon)$ neighbourhood of the $t$-dependent family of hypersurfaces $(\gamma_t^\varepsilon)_{t\geq 0}$, which is defined as follows. The initial hypersurface

$\gamma_0^\varepsilon$ is defined in (4.5) and is a slight shift of the initial interface $\Gamma_0$ defined in (1.7) (we hope that the reason for such a shift will become transparent for the reader in Section 4).

Let the family $(\gamma_t^\varepsilon)$ evolve with the law of motion

$$V = (n-1)\kappa - \frac{c(\varepsilon \xi^\varepsilon(t))}{\varepsilon} \quad \text{on } \gamma_t^\varepsilon, \tag{3.1}$$

where $c(\delta)$ is the speed of the bistable travelling wave $m(z;\delta)$ defined in (5.2). Recalling Section 2, if the noise is of the (MN1) type then this family is defined for $0 < t \leq \sigma_N^\varepsilon$, $N > 0$ arbitrary, whereas if the noise is of the (MN2) type this family is defined for $0 < t \leq \tau^\varepsilon$. In the latter case, let $T > 0$ be given as in (2.15). Also $\Omega_t^\varepsilon$ denotes the region enclosed by $\gamma_t^\varepsilon$.

**Theorem 3.1** (Emergence and motion of $\mathcal{O}(\varepsilon)$ layers)**.** *Let the nonlinearity $f$ and the initial data $u_0$ satisfy the assumptions of subsection 1.1, and the mild noise be of (MN1) or (MN2) type. In the former case, let $N > 0$ be given. Let $u^\varepsilon(x,t)$ be the solution of (1.1). Let $\eta \in (0, \eta_0 := \min(a - a_-, a_+ - a))$ be arbitrary and define $\mu$ as the derivative of $f(u)$ at the unstable zero $u = a$, that is*

$$\mu = f'(a) > 0. \tag{3.2}$$

*Then there exist positive constants $\varepsilon_0$ and $C$ such that, for all $\varepsilon \in (0, \varepsilon_0)$ and for all $t^\varepsilon \leq t \leq \sigma_N^\varepsilon$—if noise is of (MN1) type—or for all $t^\varepsilon \leq t \leq T$—if noise is of (MN2) type—where*

$$t^\varepsilon := \mu_\varepsilon^{-1} \varepsilon^2 |\ln \varepsilon|, \ \text{with } \mu_\varepsilon \to \mu \text{ as } \varepsilon \to 0,$$

*we have*

$$u^\varepsilon(x,t) \in \begin{cases} [a_- - \eta, a_+ + \eta] & \text{if } \ x \in \mathcal{N}_{C\varepsilon}(\gamma_t^\varepsilon) \\ [a_- - \eta, a_- + \eta] & \text{if } \ x \in \Omega_t^\varepsilon \setminus \mathcal{N}_{C\varepsilon}(\gamma_t^\varepsilon) \\ [a_+ - \eta, a_+ + \eta] & \text{if } \ x \in (\Omega \setminus \overline{\Omega_t^\varepsilon}) \setminus \mathcal{N}_{C\varepsilon}(\gamma_t^\varepsilon), \end{cases} \tag{3.3}$$

*where $\mathcal{N}_r(\gamma_t^\varepsilon) := \{x \in \Omega : dist(x, \gamma_t^\varepsilon) < r\}$ denotes the $r$-neighbourhood of $\gamma_t^\varepsilon$. Here, $\varepsilon_0$, $\sigma_N^\varepsilon$, and $T$ depend on the realization of the noise and are thus $\omega$-dependent.*

The above theorem implies that, once the transition layer of the solution $u^\varepsilon(x,t)$ forms at the initial stage, its thickness remains $\mathcal{O}(\varepsilon)$ as time passes. Note that the $\mathcal{O}(\varepsilon)$ thickness estimate is optimal. Indeed, if we change the variables $x, t$ to $z := \varepsilon^{-1}x, \tau := \varepsilon^{-2}t$, we get

$$\partial_\tau u = \Delta_z u + f(u) + \varepsilon \xi^\varepsilon(t).$$

The third term on the right-hand side of the above equation tends to 0 as $\varepsilon \to 0$, therefore we have a uniform estimate for $|\nabla_z u|$. This implies $|\nabla_x u| = \mathcal{O}(\varepsilon^{-1})$. Hence the transition layer cannot be narrower than $\mathcal{O}(\varepsilon)$. This optimal $\mathcal{O}(\varepsilon)$ estimate is known for deterministic problems (see [1]), but is new for the stochastic equastion (1.1).

The next two theorems are concerned with the motion law of the sharp interface limit $\varepsilon \to 0$. These theorems, which follow from Theorem 3.1, extend the results of [17], [34] on solutions with well-prepared ($\varepsilon$-dependent) initial data to solutions with rather general $\varepsilon$-independent initial data.

**Theorem 3.2** (Extension of Funaki [17] to general initial data)**.** *Let the nonlinearity $f$ and the initial data $u_0$ satisfy the assumptions of subsection 1.1. Let the mild noise be of (MN1) type and also $u^\varepsilon(x,t)$ be the solution of (1.1). Assume further that $n = 2$ and that $\Omega_0$ is strictly convex. Following subsection 2.1, let $(\Gamma_t)_{0 \leq t < \sigma := \lim_{N\to\infty} \sigma_N}$ evolving by*

$$V = \kappa + (c_0 \alpha_0) \dot{W}_t,$$

*with $c_0 > 0$ the constant defined in* (5.4)*, and*

$$\alpha_0 := \sqrt{2\int_0^\infty E[\xi_0\xi_t]dt}. \tag{3.4}$$

*Define $T^\varepsilon := t^\varepsilon + \frac{2}{\beta}\varepsilon^2|\ln\varepsilon|$, where $\beta > 0$ is as in Proposition 5.2. Then there exists random motion of curves $(\Gamma_t^\varepsilon)_{0\le t<\sigma^\varepsilon:=\lim_{N\to\infty}\sigma_N^\varepsilon}$, with stopping times $\sigma_N^\varepsilon$ as in* [17, (1.9)]*, satisfying the following two conditions:*

(*i*) *Let $N > 0$ be given. For $0 \le t \le \sigma_N^\varepsilon - T^\varepsilon$, let $x \mapsto \Phi^\varepsilon(x,t)$ be the step function with value $a_-$ in the region enclosed by $\Gamma_t^\varepsilon$ and $a_+$ elsewhere. Then*

$$\sup_{0\le t\le\sigma_N^\varepsilon - T^\varepsilon} \|u^\varepsilon(\cdot, T^\varepsilon + t) - \Phi^\varepsilon(\cdot,t)\|_{L^2(\Omega)} \to 0 \quad \text{in probability, as } \varepsilon\to 0.$$

(*ii*) *$\Gamma_t^\varepsilon$ converges to $\Gamma_t$ as $\varepsilon \to 0$ in the following sense: for any $T > 0$ and all but countable many $N \in \mathbb{R}^+$, the joint distribution of $(\sigma_N^\varepsilon, \Gamma^\varepsilon_{t\wedge\sigma_N^\varepsilon})$ on $\mathbb{R}^+ \times C([0,T], C([0,2\pi),\mathbb{R}^2))$ converges, as $\varepsilon \to 0$, to that of $(\sigma_N, \Gamma_{t\wedge\sigma_N})$.*

**Theorem 3.3** (Extension of Weber [34] to general initial data)**.** *Let the nonlinearity $f$ and the initial data $u_0$ satisfy the assumptions of subsection 1.1. Let the mild noise be of (MN2) type. Let $u^\varepsilon(x,t)$ be the solution of* (1.1)*. Following subsection 2.2, let $(\Gamma_t)_{0\le t<\tau(\Gamma_0)}$ evolve by*

$$dV = (n-1)\kappa dt + c_0 dW_t,$$

*with $c_0 > 0$ the constant defined in* (5.4)*. For $0 \le t < \tau(\Gamma_0)$, let $x \mapsto \Phi(x,t)$ be the step function with value $a_-$ in the region enclosed by $\Gamma_t$ and $a_+$ elsewhere. Let $T > 0$ be as in* (2.15)*.*

*Then there exists a stopping time $\sigma$, with $t^\varepsilon < \sigma \le \min(\tau(\Gamma_0), T)$, such that*

$$\sup_{t^\varepsilon\le t\le\sigma} \|u^\varepsilon(\cdot,t) - \Phi(\cdot,t)\|_{L^2(\Omega)} \to 0 \quad \text{almost surely, as } \varepsilon\to 0,$$

*where $t^\varepsilon$ is as in Theorem 3.1.*

Next we study the profile of the solution $u^\varepsilon$ near the interface. As shown in [1, Section 2] for the deterministic problems, the formal asymptotic expansion of $u^\varepsilon(x,t)$ near the interface $\Gamma_t^\varepsilon$ is given in the form

$$u^\varepsilon(x,t) \sim U_0\left(\frac{d(x,t)}{\varepsilon}\right) + \varepsilon U_1\left(\frac{d(x,t)}{\varepsilon}\right) + \cdots, \tag{3.5}$$

where $d(\cdot,t)$ denotes the signed distance function to $\Gamma_t^\varepsilon$, and $U_0(z)$ is the unique solution of the stationary problem

$$\begin{cases} U_0'' + f(U_0) = 0 \\ U_0(-\infty) = a_-\,, \quad U_0(0) = a\,, \quad U_0(\infty) = a_+\,. \end{cases} \tag{3.6}$$

The existence of such $U_0$ is guaranteed by the integral condition (1.5). Theorem 3.4 below proves rigorously the validity of the leading order term $U_0$ in the asymptotic expansion (3.5). The same result was first shown in [2] for deterministic problems of the form (1.17). Theorem 3.4 below shows that the leading order term is not destroyed by a random noise.

Before stating the theorem, let us introduce some notation. The level surface of the solution $u^\varepsilon$ is defined by

$$\Gamma_t^\varepsilon := \{x \in \Omega :\ u^\varepsilon(x,t) = a\} \tag{3.7}$$

and the *signed distance function associated with* $\Gamma^\varepsilon$ by

$$\overline{d^\varepsilon}(x,t) := \begin{cases} -\operatorname{dist}(x,\Gamma_t^\varepsilon) & \text{if } u^\varepsilon(x,t) < a \\ \ \ \operatorname{dist}(x,\Gamma_t^\varepsilon) & \text{if } u^\varepsilon(x,t) > a\,. \end{cases} \tag{3.8}$$

Recall that if noise is of the (MN2) type, then $T>0$ was defined in (2.15). To unify the notation in the following, if noise is of the (MN1) type, then, for a given $N>0$, we select $0<T<\sigma_N$ (see subsection 2.1). It therefore follows from (3.3) that $\Gamma_t^\varepsilon \subset \mathcal{N}_{C\varepsilon}(\gamma_t^\varepsilon)$ for all $t^\varepsilon \le t \le T$, so that

$$|\overline{d^\varepsilon}(x,t) - d^\varepsilon(x,t)| \le C\varepsilon \quad \forall (x,t) \in \overline{\Omega} \times [t^\varepsilon, T], 0<\varepsilon << 1. \tag{3.9}$$

**Theorem 3.4** (Profile in the layers)**.** *Let the assumptions of Theorem 3.1 hold. Fix $\rho>1$ and $0<T'<T$. Then*

(i) *If $\varepsilon>0$ is small enough then, for any $t\in[\rho t^\varepsilon, T']$, the level set $\Gamma_t^\varepsilon$ is a smooth hypersurface and can be expressed as a graph over $\gamma_t^\varepsilon$.*

(ii) *We have*

$$\lim_{\varepsilon\to 0} \sup_{\rho t^\varepsilon \le t \le T',\, x\in\bar\Omega} \left| u^\varepsilon(x,t) - U_0\left(\frac{\overline{d^\varepsilon}(x,t)}{\varepsilon}\right)\right| = 0 \quad \text{a.s.}, \tag{3.10}$$

*where $\overline{d^\varepsilon}$ denotes the signed distance function associated with $\Gamma^\varepsilon$.*

The rest of the paper is organized as follows. In Section 4, we prove the emergence of internal layers (or *generation of interface*) for problem (1.1). In Section 5, we construct an appropriate pair of sub- and super-solutions to prove the validity of the interface equation (2.1) with optimal $\mathcal{O}(\varepsilon)$ error estimates. In Section 6, by combining the results in the earlier sections, we prove Theorem 3.1, then also Theorems 3.2 and 3.3. Finally, we prove Theorem 3.4 in Section 7.

## 4. Rapid emergence of $\mathcal{O}(\varepsilon)$ layers

This section deals with the emergence of internal layers (or the generation of interface) that occurs very quickly. More precisely, given a virtually arbitrary initial data that satisfies mild nondegeneracy conditions, we prove that the solution $u^\varepsilon(x,t)$ quickly becomes close to $a_\pm$ in most part of $\Omega$ and develops an internal layer of thickness $\mathcal{O}(\varepsilon)$. This scenario looks similar to the generation of interface in deterministic problems, but there is one important difference that we have to deal with carefully. More precisely, the initial value of the mild noise, $\xi^\varepsilon(0)$, is not bounded but is of order $\mathcal{O}(\varepsilon^{-\gamma})$ for some $0<\gamma<1/3$, where

$$\begin{cases} 0<\gamma_1<\gamma<\frac{1}{3} \text{ if the noise is of the (MN1) type} \\ 0<\frac{\gamma_2}{2}<\gamma<\frac{1}{3} \text{ if the noise is of the (MN2) type.} \end{cases}$$

Consequently, there is an abrupt shift of the interface of order $\mathcal{O}(\varepsilon^{1-\gamma})$ right after $t=0$ before the formation of the internal layer. Therefore, a layer of thickness $\mathcal{O}(\varepsilon)$ does not form around $\Gamma_0$ but around a slightly shifted smooth hypersurface $\gamma_0^\varepsilon$ that lies in an $\mathcal{O}(\varepsilon^{1-\gamma})$ neighbourhood of $\Gamma_0$,

The reason for such an abrupt initial drift of the interface is the following. For $0\le t\le T$, the mean value theorem provides a $0<\theta<1$ such that

$$\xi^\varepsilon(t) = \xi^\varepsilon(0) + \dot\xi^\varepsilon(\theta t)t = \xi^\varepsilon(0) + o(\varepsilon), \tag{4.1}$$

as long as $0\le t\le \mathcal{O}(\varepsilon^2|\ln\varepsilon|)$, where we have used (1.12) under the noise assumption (MN1), and (1.16) under the noise assumption (MN2). Once the crucial observation (4.1) is made, the treatment of the $o(\varepsilon)$ term mainly follows from the generation of interface property performed in [1, Section 4], whereas the $\xi^\varepsilon(0) = o(\varepsilon^{-\gamma})$ term explains the initial shift.

In order to take advantage of observation (4.1), we define

$$f^\varepsilon(u) := f(u) + \varepsilon\xi^\varepsilon(0). \tag{4.2}$$

In view of assumptions (1.3) and (1.4) on $f$, and since $\varepsilon\xi^\varepsilon(0) = o(\varepsilon^{1-\gamma}) \to 0$, we have, for $\varepsilon > 0$ small enough, that $f^\varepsilon$ is still of the bistable type, in the sense that

$$f^\varepsilon \text{ has exactly three zeros } \; a_-^\varepsilon < a^\varepsilon < a_+^\varepsilon, \tag{4.3}$$

where $a_-^\varepsilon = a_- + o(\varepsilon^{1-\gamma})$, $a^\varepsilon = a + o(\varepsilon^{1-\gamma})$, $a_+^\varepsilon = a_+ + o(\varepsilon^{1-\gamma})$, and

$$\frac{d}{du} f^\varepsilon(a_\pm^\varepsilon) \to f'(a_\pm) < 0, \quad \mu_\varepsilon := \frac{d}{du} f^\varepsilon(a^\varepsilon) \to \mu = f'(a) > 0. \tag{4.4}$$

We now define

$$\gamma_0^\varepsilon := \{x \in \Omega : \; u_0(x) = a^\varepsilon\}, \tag{4.5}$$

which consists in a $o(\varepsilon^{1-\gamma})$ shift of the initial interface $\Gamma_0$ defined in (1.7). In view of assumptions in subsection 1.1, $\gamma_0^\varepsilon$ is a smooth hypersurface without boundary and properties analogous to (1.9) and (1.8) hold true with obvious changes. In particular, thanks to the compactness of $\Gamma_0$, (1.9) is transferred into

$$\nabla u_0(x) \cdot n^\varepsilon(x) \geq \theta' > 0 \;\text{ for any } x \in \gamma_0^\varepsilon, \tag{4.6}$$

for all $\varepsilon > 0$ small enough. We can now state our generation of interface result.

**Theorem 4.1** (Emergence of $\mathcal{O}(\varepsilon)$ layers around $\gamma_0^\varepsilon$)**.** *Let the nonlinearity $f$ and the initial data $u_0$ satisfy the assumptions of subsection 1.1. Let the mild noise be of (MN1) or (MN2) type. Let $u^\varepsilon(x,t)$ be the solution of (1.1). Define $\eta_0 := \min(a - a_-, a_+ - a)$, and let $\eta \in (0, \eta_0)$ be arbitrary.*

*Then there exist positive constants $\varepsilon_0$ and $M_0$ such that, for all $\varepsilon \in (0, \varepsilon_0)$,*

*(i) for all $x \in \Omega$,*

$$a_- - \eta \leq u^\varepsilon(x, \mu_\varepsilon^{-1}\varepsilon^2|\ln\varepsilon|) \leq a_+ + \eta, \tag{4.7}$$

*(ii) for all $x \in \Omega$ such that $|u_0(x) - a^\varepsilon| \geq M_0\varepsilon$, we have that*

$$\text{if } u_0(x) \geq a^\varepsilon + M_0\varepsilon \text{ then } u^\varepsilon(x, \mu_\varepsilon^{-1}\varepsilon^2|\ln\varepsilon|) \geq a_+ - \eta, \tag{4.8}$$

$$\text{if } u_0(x) \leq a^\varepsilon - M_0\varepsilon \text{ then } u^\varepsilon(x, \mu_\varepsilon^{-1}\varepsilon^2|\ln\varepsilon|) \leq a_- + \eta. \tag{4.9}$$

*Proof.* In view of the crucial observation (4.1) and definition (4.2), the Allen–Cahn equation (1.1) is recast (for small enough times)

$$\partial_t u = \Delta u + \frac{1}{\varepsilon^2}\left(f^\varepsilon(u) - \varepsilon g^\varepsilon(t)\right), \quad 0 < t \leq \mu_\varepsilon^{-1}\varepsilon^2|\ln\varepsilon|, \quad x \in \Omega,$$

where the perturbation term

$$g^\varepsilon(t) := -\xi^\varepsilon(t) + \xi^\varepsilon(0),$$

satisfies $\|g^\varepsilon\|_{L^\infty(0,\mu_\varepsilon^{-1}\varepsilon^2|\ln\varepsilon|)} = o(\varepsilon)$, as $\varepsilon \to 0$. Moreover, using (1.12) under the noise assumption (MN1), and (1.16) under the noise assumption (MN2), we get that, in any case,

$$\|\dot{g}^\varepsilon\|_{L^\infty(0,\mu_\varepsilon^{-1}\varepsilon^2|\ln\varepsilon|)} = \mathcal{O}\left(\varepsilon^{-1}\right), \quad \text{as } \varepsilon \to 0.$$

After writing the problem in such a form and as far as the perturbation term is concerned, we are in a "neighbourhood" of the Allen–Cahn equation ($P^\varepsilon$) studied in [1], since the above estimate corresponds to assumption (1.3) in [1].

Nevertheless, we need to handle the following change: $f$ in [1] is replaced by $f^\varepsilon$ in our setting. This difference implies that $a$ in [1] is replaced by $a^\varepsilon$ and is the reason why the generation occurs around $\gamma_0^\varepsilon$ (and not around $\Gamma_0$). Nevertheless, it is completely transparent that $f^\varepsilon$ is still of the bistable type *uniformly with respect to small* $\varepsilon > 0$ (this property is used in order to derive certain estimates). More precisely, (4.3) and (4.4) correspond to assumption (1.1) in [1], uniformly with respect to small $\varepsilon > 0$. Similarly, the non degeneracy assumption (4.6), when crossing the initial interface $\gamma_0^\varepsilon$, is uniform with respect to small $\varepsilon > 0$ and corresponds to assumption (1.10) in [1].

We can then construct the analogous of the sub- and supersolutions of [1, Section 4], namely

$$w^{\pm}(x,t) = Y^\varepsilon\left(\frac{t}{\varepsilon^2}, u_0(x) \pm \varepsilon^2 C(e^{\tilde{\mu}_\varepsilon \frac{t}{\varepsilon^2}} - 1); \pm\varepsilon\right), \tag{4.10}$$

where $C > 0$ is a large constant, $\tilde{\mu}_\varepsilon$ is a very small perturbation of $\mu_\varepsilon$, and $Y^\varepsilon(\tau, \xi; \delta)$ is the solution of the Cauchy problem

$$\begin{cases} Y_\tau^\varepsilon(\tau,\xi;\delta) = f^\varepsilon(Y^\varepsilon(\tau,\xi;\delta)) + \delta & \text{for } \tau > 0 \\ Y^\varepsilon(0,\xi;\delta) = \xi. \end{cases} \tag{4.11}$$

Notice that, in this very early stage of emergence of the layers, the above sub- and super-solutions are obtained by considering only the nonlinear reaction term, that is diffusion is neglected.

One can then follow the lengthy arguments of [1, Section 4] taking $f^\varepsilon$ in place of $f$. The aforementioned crucial observations enable us to modify properly the several proofs in [1, Section 4], and details are omitted. We then derive (4.7), (4.8), (4.9), which concludes the proof. □

## 5. Propagation of $\mathcal{O}(\varepsilon)$ layers

In this section, we construct a pair of sub- and super-solutions whose role is to capture in an $\mathcal{O}(\varepsilon)$ sandwich the layers of the solution $u^\varepsilon(x,t)$, while they are propagating. In order to proceed to the aforementioned construction, we need to define first properly some travelling waves and a signed distance function used in the definition of this pair.

5.1. **Some travelling waves.** For $\delta_0 > 0$ small enough and any $|\delta| \leq \delta_0$, the function $u \mapsto f(u) + \delta$ is still of bistable type, we denote by $a_-(\delta) < a(\delta) < a_+(\delta)$ its three zeros, and there is $\overline{C} > 0$ such that, for all $\delta_1 \neq \delta_2$ with $|\delta_1| \leq \delta_0$, $|\delta_2| \leq \delta_0$,

$$\frac{a_\pm(\delta_2) - a_\pm(\delta_1)}{\delta_2 - \delta_1} \geq \overline{C}, \quad \frac{a(\delta_2) - a(\delta_1)}{\delta_2 - \delta_1} \leq -\overline{C}. \tag{5.1}$$

Let $c(\delta)$, $m(z;\delta)$ be the speed and the profile of the unique travelling wave associated with the one dimensional problem

$$\partial_t v = v_{zz} + f(v) + \delta, \;\; t > 0, \; z \in \mathbb{R}.$$

In other words, we have

$$\begin{aligned} &m_{zz}(z;\delta) + c(\delta) m_z(z;\delta) + f(m(z;\delta)) + \delta = 0, \quad z \in \mathbb{R}, \\ &m(-\infty;\delta) = a_-(\delta), \quad m(0;\delta) = a(\delta), \quad m(+\infty;\delta) = a_+(\delta). \end{aligned} \tag{5.2}$$

Notice in particular that the assumption of balanced nonlinearity (1.5) implies $c(0) = 0$. Moreover, the following estimates are well-known (see in [11], [17] or [34]).

**Lemma 5.1** (Estimates on travelling waves)**.** *There exist constants $\delta_0 > 0$, $C > 0$, $\lambda > 0$ such that, for all $|\delta| \leq \delta_0$,*

$$\begin{aligned} &0 < a_+(\delta) - m(z;\delta) \leq Ce^{-\lambda|z|}, \quad z \geq 0, \\ &0 < m(z;\delta) - a_-(\delta) \leq Ce^{-\lambda|z|}, \quad z \leq 0, \\ &0 < m_z(z;\delta) \leq Ce^{-\lambda|z|}, \quad z \in \mathbb{R}, \\ &|m_{zz}(z;\delta)| \leq Ce^{-\lambda|z|}, \quad z \in \mathbb{R}, \\ &|m_\delta(z;\delta)| \leq C, \quad z \in \mathbb{R}, \end{aligned} \tag{5.3}$$

*and*

$$\partial_\delta c(0) = -c_0 := -\frac{a_+ - a_-}{\displaystyle\int_{a_-}^{a_+} \sqrt{2F(u)}\,du} < 0, \quad F(u) := \int_u^{a_+} f(z)\,dz. \tag{5.4}$$

5.2. **Signed distance functions.** We recall that the family of hypersurfaces $(\gamma_t^\varepsilon)$ follows the law (3.1) with initial data $\gamma_0^\varepsilon$ defined in (4.5). If the noise is of the (MN1) type then it follows from (2.11) and (2.12) that, up to reducing $\varepsilon_0$ if necessary,

$$\mathcal{K} := \sup_{0<\varepsilon<\varepsilon_0} \sup_{0\leq t\leq \sigma_N^\varepsilon} \sup_{y\in\gamma_t^\varepsilon} \sup_{1\leq i\leq n-1} |\kappa_i^\varepsilon(y,t)| < \infty,$$

with $\kappa_i^\varepsilon(y,t)$ the $i$-th principal curvature of $\gamma_t^\varepsilon$ at point $y$. On the other hand, if the noise is of the (MN2) type then it follows from (2.15) that, up to reducing $\varepsilon_0$ if necessary,

$$\mathcal{K} := \sup_{0<\varepsilon<\varepsilon_0} \sup_{0\leq t\leq T} \sup_{y\in\gamma_t^\varepsilon} \sup_{1\leq i\leq n-1} |\kappa_i^\varepsilon(y,t)| < \infty.$$

In the sequel we unify the notations by letting $\mathcal{T} = \sigma_N^\varepsilon$, $\mathcal{T} = T$ if the noise is of the (MN1) type, (MN2) type respectively.

Let $\Omega_t^\varepsilon$ denote the region enclosed by $\gamma_t^\varepsilon$. We then define the associated signed distance function by

$$\tilde{d}^\varepsilon(x,t) := \begin{cases} -\text{dist}(x,\gamma_t^\varepsilon) & \text{for } x \in \Omega_t^\varepsilon, \\ +\text{dist}(x,\gamma_t^\varepsilon) & \text{for } x \in \Omega \setminus \overline{\Omega_t^\varepsilon}. \end{cases} \tag{5.5}$$

For $d_0 > 0$, choose an increasing function $\varphi \in C^\infty(\mathbb{R})$ satisfying

$$\varphi(s) = \begin{cases} -2d_0 & \text{if } s \leq -2d_0, \\ s & \text{if } |s| \leq d_0, \\ 2d_0 & \text{if } s \geq 2d_0. \end{cases}$$

If $d_0$ is sufficiently small, then, for any $0 < \varepsilon < \varepsilon_0$,

$$d^\varepsilon(x,t) := \varphi(\tilde{d}^\varepsilon(x,t))$$

is smooth in $\Omega \times (0,\mathcal{T})$, satisfies $d^\varepsilon(x,t) = 0$ for $x \in \gamma_t^\varepsilon$,

$$|\nabla d^\varepsilon(x,t)| = 1 \quad \text{in } \{(x,t) : |d^\varepsilon(x,t)| < d_0\}. \tag{5.6}$$

Also, since the inward normal velocity $V$ and the mean curvature $\kappa$ are equal to $\partial_t d^\varepsilon$ and $\frac{\Delta d^\varepsilon}{n-1}$, equation (3.1) is recast as

$$\partial_t d^\varepsilon(y,t) = \Delta d^\varepsilon(y,t) - \frac{c(\varepsilon\xi^\varepsilon(t))}{\varepsilon} \quad \text{on } \{(y,t) : y \in \gamma_t^\varepsilon\}. \tag{5.7}$$

5.3. **An $\mathcal{O}(\varepsilon)$-sandwich of the layers.** Equipped with the above material, we are now in the position to construct sub-and supersolutions for equation (1.1) in the form

$$u_\varepsilon^\pm(x,t) := m\left(\frac{d^\varepsilon(x,t)\pm\varepsilon p(t)}{\varepsilon};\varepsilon\xi^\varepsilon(t)\right)\pm q(t), \tag{5.8}$$

where

$$p(t) := -e^{-\beta t/\varepsilon^2}+e^{Lt}+K, \quad q(t) := \sigma\big(\beta e^{-\beta t/\varepsilon^2}+\varepsilon^2 Le^{Lt}\big), \tag{5.9}$$

where $\beta$, $\sigma$, $K$ and $L$ are positive constants to be chosen. Notice that $q=\sigma\varepsilon^2 p_t$. Notice also that, initially, the vertical shift $q(0)$ is $\mathcal{O}(1)$ but, as soon as $t>0$, $q(t)$ becomes $\mathcal{O}(\varepsilon^2)$. Furthermore, it is clear from the definition of $u_\varepsilon^\pm$ that, as soon as $t>0$, $\lim_{\varepsilon\to 0} u_\varepsilon^\pm(x,t)=a_-$, respectively $a_+$, if $x\in\Omega_t^\varepsilon$, respectively $x\in\Omega\setminus\overline{\Omega_t^\varepsilon}$.

**Proposition 5.2** (Sub- and supersolutions for the propagation)**.** *Choose $\beta>0$ and $\sigma>0$ appropriately. Then for any $K>1$, there exist constants $\varepsilon_0>0$ and $L>0$ such that, for any $\varepsilon\in(0,\varepsilon_0)$, the functions $(u_\varepsilon^-,u_\varepsilon^+)$ are a pair of sub- and super-solutions for equation* (1.1) *in the domain $\Omega\times(0,\mathcal{T})$, that is*

$$\mathcal{L}u_\varepsilon^+ := \partial_t u_\varepsilon^+ - \Delta u_\varepsilon^+ - \frac{1}{\varepsilon^2}f(u_\varepsilon^+) - \frac{1}{\varepsilon}\xi^\varepsilon(t)\geq 0, \quad \mathcal{L}u_\varepsilon^-\leq 0,$$

*in $\Omega\times(0,\mathcal{T})$.*

*Proof.* We only give the proof of the inequality for $u_\varepsilon^+$, since the one for $u_\varepsilon^-$ follows the same argument. In the sequel, $m$ and its derivatives are evaluated at

$$(z^*;\delta^*) := \left(\frac{d^\varepsilon(x,t)+\varepsilon p(t)}{\varepsilon};\varepsilon\xi^\varepsilon(t)\right)$$

which belongs to $\mathbb{R}\times(-\delta_0,\delta_0)$ if $\varepsilon>0$ is small enough. Straightforward computations combined with

$$f(m+q)=f(m)+qf'(m)+\frac{1}{2}q^2f''(\theta), \qquad \text{for some } m<\theta=\theta(x,t)<u_\varepsilon^+,$$

and equation (5.2) yield $\mathcal{L}u_\varepsilon^+=E_1+E_2+E_3+E_4$, with

$$\begin{aligned}
E_1 &= -\frac{1}{\varepsilon^2}q\left(f'(m)+\frac{1}{2}qf''(\theta)\right)+m_zp_t+q_t\\
E_2 &= (1-|\nabla d^\varepsilon|^2)\frac{m_{zz}}{\varepsilon^2}\\
E_3 &= \left(\partial_t d^\varepsilon(x,t)-\Delta d^\varepsilon(x,t)+\frac{c(\varepsilon\xi^\varepsilon(t))}{\varepsilon}\right)\frac{m_z}{\varepsilon}\\
E_4 &= \varepsilon\dot{\xi}^\varepsilon(t)m_\delta.
\end{aligned}$$

Let us first present some useful inequalities. By assumption (1.4), there are $b>0$, $\rho>0$ such that

$$f'(m(z;\delta))\leq-\rho \qquad \text{if} \quad m(z;\delta)\in[a_--b,a_-+b]\cup[a_+-b,a_++b]. \tag{5.10}$$

On the other hand, since the region $\{(z;\delta)\in\mathbb{R}\times(-\delta_0,\delta_0):\, m(z;\delta)\in[a_-+b,\,a_+-b]\,\}$ is compact, there is $a_1>0$ such that

$$m_z(z;\delta)\geq a_1 \qquad \text{if} \quad m(z;\delta)\in[a_-+b,\,a_+-b]. \tag{5.11}$$

We now select

$$\beta = \frac{\rho}{4}, \quad 0 < \sigma \le \min(\sigma_0, \sigma_1, \sigma_2), \tag{5.12}$$

where

$$\sigma_0 := \frac{a_1}{\rho + \|f'\|_{L^\infty(a_- -1, a_+ +1)}}, \quad \sigma_1 := \frac{1}{2(\beta+1)}, \quad \sigma_2 := \frac{4\beta}{\|f''\|_{L^\infty(a_- -1, a_+ +1)}(\beta+1)}.$$

Combining (5.10), (5.11) and $0 < \sigma \le \sigma_0$, we obtain

$$m_z(z;\delta) - \sigma f'(m(z;\delta)) \ge \sigma\rho, \quad \forall (z;\delta) \in \mathbb{R} \times (-\delta_0, \delta_0). \tag{5.13}$$

Now let $K > 1$ be arbitrary. In what follows we shall show that $\mathcal{L}u_\varepsilon^+ \ge 0$ provided that the constants $\varepsilon_0$ and $L$ are appropriately chosen. We go on under the following assumption (to be checked at the end)

$$\varepsilon_0^2 L e^{L\mathcal{T}} \le 1. \tag{5.14}$$

Then, given any $\varepsilon \in (0, \varepsilon_0)$, we have, since $\sigma \le \sigma_1$, $0 \le q(t) \le \frac{1}{2}$, that

$$a_- - 1 \le u_\varepsilon^\pm(x,t) \le a_+ + 1. \tag{5.15}$$

Using the expressions for $p$ and $q$, the "favourable" term $E_1$ is recast as

$$E_1 = \frac{\beta}{\varepsilon^2} e^{-\beta t/\varepsilon^2}(I - \sigma\beta) + Le^{Lt}(I + \varepsilon^2\sigma L),$$

where

$$I = m_z(z^*;\delta^*) - \sigma f'(m(z^*;\delta^*)) - \frac{\sigma^2}{2} f''(\theta)(\beta e^{-\beta t/\varepsilon^2} + \varepsilon^2 L e^{Lt}).$$

In virtue of (5.13), (5.15) and (5.14), we have $I \ge \sigma\rho - \frac{\sigma^2}{2}\|f''\|_{L^\infty(a_- -1, a_+ +1)}(\beta+1)$. Since $0 < \sigma \le \sigma_2$, we obtain $I \ge 2\sigma\beta$. Consequently, we have

$$E_1 \ge \frac{\sigma\beta^2}{\varepsilon^2} e^{-\beta t/\varepsilon^2} + 2\sigma\beta L e^{Lt} \ge 2\sigma\beta L e^{Lt}.$$

Next, in view of (5.6), $E_2 = 0$ in the region $|d^\varepsilon(x,t)| \le d_0$. Next we consider the region where $|d^\varepsilon(x,t)| \ge d_0$. We deduce from Lemma 5.1 that

$$|E_2| \le \frac{C}{\varepsilon^2} e^{-\lambda |d^\varepsilon(x,t) + \varepsilon p(t)|/\varepsilon} \le \frac{C}{\varepsilon^2} e^{-\lambda(d_0/\varepsilon - p(t))}.$$

We remark that $0 < K - 1 \le p \le e^{L\mathcal{T}} + K$. Consequently, if we assume (to be checked at the end)

$$e^{L\mathcal{T}} + K \le \frac{d_0}{2\varepsilon_0}, \tag{5.16}$$

then $\dfrac{d_0}{\varepsilon} - p(t) \ge \dfrac{d_0}{2\varepsilon}$, so that $|E_2| \le \frac{C}{\varepsilon^2} e^{-\lambda d_0/(2\varepsilon)} = \mathcal{O}(1)$, as $\varepsilon \to 0$.

Let us now turn to the term $E_3$. In the region where $|d^\varepsilon(x,t)| \ge \min(d_0, \frac{1}{2\mathcal{K}}) > 0$ (away from the interface), argument similar as those for $E_2$ yield $|E_3| = \mathcal{O}(1)$ as $\varepsilon \to 0$ (thanks to the exponential decay of the wave). In the region where $|d^\varepsilon(x,t)| \le \min(d_0, \frac{1}{2\mathcal{K}})$, let us pick a $y \in \gamma_t^\varepsilon$ such that $|d^\varepsilon(x,t)| = dist(x,y)$. In view of (5.7) and $\partial_t d^\varepsilon(x,t) = \partial_t d^\varepsilon(y,t)$ we get

$$E_3 = (\Delta d^\varepsilon(y,t) - \Delta d^\varepsilon(x,t)) \frac{m_z}{\varepsilon}.$$

But it follows from [19, Lemma 14.17] that

$$\begin{aligned}|\Delta d^\varepsilon(y,t)-\Delta d^\varepsilon(x,t)| &= \left|\sum_{i=1}^{n-1}\kappa_i^\varepsilon(y,t)-\sum_{i=1}^{n-1}\frac{\kappa_i^\varepsilon(y,t)}{1-d^\varepsilon(x,t)\kappa_i^\varepsilon(y,t)}\right|\\ &\leq |d^\varepsilon(x,t)|\sum_{i=1}^{n-1}\frac{(\kappa_i^\varepsilon)^2(y,t)}{|1-d^\varepsilon(x,t)\kappa_i^\varepsilon(y,t)|}\\ &\leq 2|d^\varepsilon(x,t)|\sum_{i=1}^{n-1}(\kappa_i^\varepsilon)^2(y,t)\end{aligned}$$

since $|d^\varepsilon(x,t)|\leq\frac{1}{2\mathcal{K}}$, and $|\kappa_i^\varepsilon(y,t)|\leq\mathcal{K}$. As a result we have $|E_3|\leq 2(n-1)\mathcal{K}^2|d^\varepsilon(x,t)|=:C|d^\varepsilon(x,t)|$, so that

$$\begin{aligned}|E_3| &\leq C\frac{|d^\varepsilon(x,t)|}{\varepsilon}m_z\left(\frac{d^\varepsilon(x,t)+\varepsilon p(t)}{\varepsilon};\varepsilon\xi^\varepsilon(t)\right)\\ &\leq C\sup_{z\in\mathbb{R},|\delta|\leq\delta_0}|zm_z(z;\delta)|+C\varepsilon|p(t)|\sup_{z\in\mathbb{R},|\delta|\leq\delta_0}|m_z(z;\delta)|\\ &\leq C_3+C_3'(e^{Lt}+K),\end{aligned}$$

for some constants $C_3>0$, $C_3'>0$ and where we have used Lemma 5.1.

Last, it follows from (1.12), (1.16) and Lemma 5.1 that $|E_4|\to 0$, as $\varepsilon\to 0$, uniformly in $\Omega\times(0,\mathcal{T})$.

Putting the above estimates all together, we arrive at

$$\mathcal{L}u_\varepsilon^+\geq(2\sigma\beta L-C_3')e^{Lt}-\mathcal{O}(1)$$

which is nonnegative, if $L>0$ is sufficiently large, and $\varepsilon_0>0$ sufficiently small to validate assumptions (5.14) and (5.16). The theorem is proved. □

## 6. Description of the $\mathcal{O}(\varepsilon)$ layers and their convergence

6.1. **Proof of Theorem 3.1.** Let $\eta\in(0,\eta_0)$ be given. Let us select $\beta>0$ and $\sigma>0$ that satisfy (5.12)—so that Proposition 5.2 is available—and $\beta\sigma\leq\eta/3$. By the emergence of the layers property, we are equipped with small $\varepsilon_0>0$ and a $M_0>0$ such that (4.7), (4.8), (4.9) hold with $\beta\sigma/2$ playing the role of $\eta$. On the other hand, in view of (4.6), there is $M_1>0$ such that we have the following correspondence

$$\begin{aligned}&\text{if}\quad d^\varepsilon(x,0)\geq M_1\varepsilon\quad\text{then}\quad u_0(x)\geq a^\varepsilon+M_0\varepsilon\\ &\text{if}\quad d^\varepsilon(x,0)\leq -M_1\varepsilon\quad\text{then}\quad u_0(x)\leq a^\varepsilon-M_0\varepsilon,\end{aligned}\tag{6.1}$$

where we recall that $d^\varepsilon(x,0)$ denotes the signed distance function associated with the hypersurface $\gamma_0^\varepsilon:=\{x:\,u_0(x)=a^\varepsilon\}$. Now we define functions $H^+(x),H^-(x)$ by

$$H^+(x)=\begin{cases}a_++\sigma\beta/2 & \text{if } d^\varepsilon(x,0)\geq -M_1\varepsilon\\ a_-+\sigma\beta/2 & \text{if } d^\varepsilon(x,0)<-M_1\varepsilon,\end{cases}$$
$$H^-(x)=\begin{cases}a_+-\sigma\beta/2 & \text{if } d^\varepsilon(x,0)\geq M_1\varepsilon\\ a_--\sigma\beta/2 & \text{if } d^\varepsilon(x,0)< M_1\varepsilon.\end{cases}$$

Then from the above observations we see that, after a very short time $\mathcal{O}(\varepsilon^2|\ln\varepsilon|)$, we have an $\mathcal{O}(\varepsilon)$ sandwich of the layers, namely

$$H^-(x) \le u^\varepsilon(x, \mu_\varepsilon^{-1}\varepsilon^2|\ln\varepsilon|) \le H^+(x) \qquad \text{for } x \in \Omega. \tag{6.2}$$

We now would like to use the sub and supersolutions (5.8) for the propagation described at Section 5. Observe that

$$u_\varepsilon^\pm(x,0) = m\left(\frac{d^\varepsilon(x,0) \pm K}{\varepsilon}; \varepsilon\xi^\varepsilon(0)\right) \pm \sigma(\beta + \varepsilon^2 L),$$

so that it follows from $\varepsilon\xi^\varepsilon(0) = \mathcal{O}(\varepsilon^{1-\gamma}) \to 0$ and Lemma 5.1 on travelling waves $m(z;\delta)$ that we can select $K >> M_1$ so that

$$u_\varepsilon^-(x,0) \le H^-(x) \le u^\varepsilon(x, \mu_\varepsilon^{-1}\varepsilon^2|\ln\varepsilon|) \le H^+(x) \le u_\varepsilon^+(x,0) \qquad \text{for } x \in \Omega. \tag{6.3}$$

Let us now choose $\varepsilon_0 > 0$ and $L > 0$ so that Proposition 5.2 applies. It therefore follows from the comparison principle that

$$u_\varepsilon^-(x,t) \le u^\varepsilon(x, t + t^\varepsilon) \le u_\varepsilon^+(x,t) \quad \text{for } x \in \Omega,\ 0 \le t \le \mathcal{T} - t^\varepsilon, \tag{6.4}$$

where $t^\varepsilon = \mu_\varepsilon^{-1}\varepsilon^2|\ln\varepsilon|$.

To conclude, in view of $\varepsilon\xi^\varepsilon(t) = \mathcal{O}(\varepsilon^{1-\gamma}) \to 0$ and Lemma 5.1 on travelling waves, we can select $\varepsilon_0 > 0$ small enough and $C > 0$ large enough so that, for all $\varepsilon \in (0, \varepsilon_0)$, all $0 \le t \le \mathcal{T} - t^\varepsilon$,

$$m(C - e^{L\mathcal{T}} - K; \varepsilon\xi^\varepsilon(t)) \ge a_+ - \frac{\eta}{2} \quad \text{and} \quad m(-C + e^{L\mathcal{T}} + K; \varepsilon\xi^\varepsilon(t)) \le a_- + \frac{\eta}{2}. \tag{6.5}$$

Using inequalities (6.4), expressions (5.8) for $u_\varepsilon^\pm$, estimates (6.5) and $\sigma\beta \le \eta/3$ we then see that, for all $\varepsilon \in (0, \varepsilon_0)$ and all $0 \le t \le \mathcal{T} - t_\varepsilon$, we have

$$\begin{array}{llll} \text{if} & d^\varepsilon(x,t) \ge C\varepsilon & \text{then} & u^\varepsilon(x, t+t^\varepsilon) \ge a_+ - \eta \\ \text{if} & d^\varepsilon(x,t) \le -C\varepsilon & \text{then} & u^\varepsilon(x, t+t^\varepsilon) \le a_- + \eta, \end{array} \tag{6.6}$$

and $u^\varepsilon(x, t+t^\varepsilon) \in [a_- - \eta, a_+ + \eta]$, which completes the proof of Theorem 3.1. □

6.2. **Proof of Theorem 3.2.** In order to prove Theorem 3.2, we apply [17, Theorem 1.1] which is concerned with well-prepared initial data. We therefore need to show that, at some small time $T^\varepsilon$, the solution $u^\varepsilon(\cdot, T^\varepsilon)$ satisfies [17, Assumption (3.8)], namely

$$m\left(\frac{d(x,0) - \varepsilon^\alpha}{\varepsilon}; \varepsilon\xi^\varepsilon(0) - \varepsilon^b\right) \le u^\varepsilon(x, T^\varepsilon) \le m\left(\frac{d(x,0) + \varepsilon^\alpha}{\varepsilon}; \varepsilon\xi^\varepsilon(0) + \varepsilon^b\right), \tag{6.7}$$

where $d(\cdot, 0)$ denotes the signed distance function to $\Gamma_0$, $\alpha > 0$ and $1 + \gamma < b < 2$. Here, and in the sequel, inequalities need to hold for $\varepsilon > 0$ sufficiently small.

We select

$$T^\varepsilon := t^\varepsilon + \bar{t}^\varepsilon, \quad \bar{t}^\varepsilon := \frac{2}{\beta}\varepsilon^2|\ln\varepsilon|, \tag{6.8}$$

where $\beta > 0$ is as in Proposition 5.2. Taking advantage of the sandwich equation (6.4) and definition (5.8), in order to prove (6.7), it is enough to obtain

$$m\Bigg(\underbrace{\frac{d^\varepsilon(x,\bar{t}^\varepsilon) + \varepsilon p(\bar{t}^\varepsilon)}{\varepsilon}}_{=:z_1}; \underbrace{\varepsilon\xi^\varepsilon\left(\bar{t}^\varepsilon\right)}_{=:\delta_1}\Bigg) + q\left(\bar{t}^\varepsilon\right) \le m\Bigg(\underbrace{\frac{d(x,0) + \varepsilon^\alpha}{\varepsilon}}_{=:z_2}; \underbrace{\varepsilon\xi^\varepsilon(0) + \varepsilon^b}_{=:\delta_2}\Bigg), \tag{6.9}$$

where $d^\varepsilon(x,t)$ denotes the signed distance function to $\gamma_t^\varepsilon$, starting from $\gamma_0^\varepsilon$ defined in (4.5) and having the law of motion (3.1). As we have seen in Section 4, $\gamma_0^\varepsilon$ lies in a $o(\varepsilon^{1-\gamma})$ neighbourhood of $\Gamma_0$, and so does $\gamma_{\overline{t}^\varepsilon}^\varepsilon$ since $\overline{t}^\varepsilon = \mathcal{O}(\varepsilon^2|\ln\varepsilon|)$. Also, from the expression of $p(t)$ in (5.9), we have $p\left(\overline{t}^\varepsilon\right) = \mathcal{O}(1)$. As a result,

$$\varepsilon z_1 \leq d(x,0) + o(\varepsilon^{1-\gamma}) \leq d(x,0) + \frac{1}{2}\varepsilon^\alpha,$$

by further requiring $0 < \alpha < 1-\gamma$. Now, from (1.12), we get, for $\varepsilon > 0$ sufficiently small,

$$\delta_1 = \varepsilon\left(\xi^\varepsilon(0) + \mathcal{O}\left(\frac{\varepsilon^2|\ln\varepsilon|}{\varepsilon^{3\gamma_1}}\right)\right) = \varepsilon\xi^\varepsilon(0) + o(\varepsilon^{3(1-\gamma)}) \leq \varepsilon\xi^\varepsilon(0) + \frac{1}{2}\varepsilon^b,$$

since $3(1-\gamma) > 2 > b$. Since both $z \mapsto m(z;\delta)$ and $\delta \mapsto m(z;\delta)$ are nondecreasing, in order to prove (6.9) it is enough to show

$$m\left(\frac{d(x,0)+\frac{1}{2}\varepsilon^\alpha}{\varepsilon};\varepsilon\xi^\varepsilon(0) + \frac{1}{2}\varepsilon^b\right) + q\left(\overline{t}^\varepsilon\right) \leq m\left(\frac{d(x,0)+\varepsilon^\alpha}{\varepsilon};\varepsilon\xi^\varepsilon(0)+\varepsilon^b\right). \tag{6.10}$$

From (5.9), we have $q\left(\overline{t}^\varepsilon\right) = \mathcal{O}(\varepsilon^2)$. We now distinguish between the following three cases in order to derive (6.10).

First, if $d(x,0) \geq -\frac{1}{2}\varepsilon^\alpha$, it is enough to have

$$a_+\left(\varepsilon\xi^\varepsilon(0) + \frac{1}{2}\varepsilon^b\right) + \mathcal{O}(\varepsilon^2) \leq m\left(\frac{1}{2}\frac{1}{\varepsilon^{1-\alpha}};\varepsilon\xi^\varepsilon(0)+\varepsilon^b\right),$$

and thus, from Lemma 5.1,

$$a_+\left(\varepsilon\xi^\varepsilon(0) + \frac{1}{2}\varepsilon^b\right) + \mathcal{O}(\varepsilon^2) \leq a_+\left(\varepsilon\xi^\varepsilon(0)+\varepsilon^b\right) - Ce^{-\frac{\lambda}{2\varepsilon^{1-\alpha}}},$$

which is true in view of (5.1), (recall $b<2$).

Next, if $d(x,0) \leq -\varepsilon^\alpha$, by a similar argument it would be sufficient to have

$$a_-\left(\varepsilon\xi^\varepsilon(0) + \frac{1}{2}\varepsilon^b\right) + Ce^{-\frac{\lambda}{2\varepsilon^{1-\alpha}}} + \mathcal{O}(\varepsilon^2) \leq a_-\left(\varepsilon\xi^\varepsilon(0)+\varepsilon^b\right),$$

which is true in view of (5.1), (recall $b<2$).

Last, if $-\varepsilon^\alpha < d(x,0) < -\frac{1}{2}\varepsilon^\alpha$, we need

$$a\left(\varepsilon\xi^\varepsilon(0) + \frac{1}{2}\varepsilon^b\right) + \mathcal{O}(\varepsilon^2) \leq a\left(\varepsilon\xi^\varepsilon(0)+\varepsilon^b\right),$$

which is true in view of (5.1), (recall $b<2$).

This proves (6.10) and thus (6.7), which concludes the proof of Theorem 3.2. □

6.3. **Proof of Theorem 3.3.** Combining Theorem 3.1 and estimate (2.15), we get Theorem 3.3 by reproducing the arguments of [34, Proof of Theorem 1.1]. Details are omitted. □

## 7. Profile in the layers

Equipped with Theorem 3.1, we can now prove the validity of the first term of the asymptotic expansions *inside* the layers, namely Theorem 3.4. The proof consists in using the stretched variables, a blow-up argument and the result of [5], as performed in the deterministic case [2].

Before going further, we recall that a solution of an evolution equation is called *eternal* (or an *entire* solution) if it is defined for all positive and negative time. We follow this terminology to refer to a solution $w(z,\tau)$ of

$$w_\tau = \Delta_z w + f(w)\,, \quad z\in\mathbb{R}^n,\ \tau\in\mathbb{R}\,. \tag{7.1}$$

Stationary solutions and travelling waves are examples of eternal solutions. We quote below a result of Berestycki and Hamel [5] asserting that "any planar-like eternal solution is actually a planar wave". More precisely, the following holds (for $z\in\mathbb{R}^n$ we write $z=(z^{(1)},\cdots,z^{(n)})$).

**Lemma 7.1** ([5, Theorem 3.1])**.** *Let $w(z,\tau)$ be an eternal bounded solution of* (7.1) *satisfying*

$$\liminf_{z^{(n)}\to\infty}\ \inf_{z'\in\mathbb{R}^{n-1},\,\tau\in\mathbb{R}} w(z,\tau)>a\,,\quad \limsup_{z^{(n)}\to-\infty}\ \sup_{z'\in\mathbb{R}^{n-1},\,\tau\in\mathbb{R}} w(z,\tau)<a\,, \tag{7.2}$$

*where $z':=(z^{(1)},\cdots,z^{(n-1)})$. Then there exists a constant $z^*\in\mathbb{R}$ such that*

$$w(z,\tau)=U_0(z^{(n)}-z^*)\,,\quad z\in\mathbb{R}^n\,,\ \tau\in\mathbb{R}\,.$$

7.1. **Proof of (ii) in Theorem 3.4.** Let $\rho>1$ and $0<T'<T$ be given. Assume by contradiction that (3.10) does not hold. Then there is $\eta>0$ and sequences $\varepsilon_k\downarrow 0$, $t_k\in[\rho t^{\varepsilon_k},T']$, $x_k\in\bar\Omega$ $(k=1,2,...)$ such that

$$\left|u^{\varepsilon_k}(x_k,t_k)-U_0\left(\frac{\overline{d^{\varepsilon_k}}(x_k,t_k)}{\varepsilon_k}\right)\right|\geq 2\eta\,. \tag{7.3}$$

In view of (3.3), (3.9) and $U_0(\pm\infty)=a_\pm$, for (7.3) to hold it is necessary to have

$$d^\varepsilon(x_k,t_k)=\mathcal{O}(\varepsilon_k)\,,\qquad \text{as }\ k\to\infty\,. \tag{7.4}$$

Recall that $d^\varepsilon(\cdot,t)$ denotes the signed distance function to $\gamma_t^\varepsilon$ as defined in subsection 5.2, whereas $\overline{d^\varepsilon}(\cdot,t)$ denotes that to $\Gamma_t^\varepsilon$ defined in (3.8).

If $u^{\varepsilon_k}(x_k,t_k)=a$, then this would mean that $x_k\in\Gamma_{t_k}^{\varepsilon_k}$, in which case the left-hand side of (7.3) would be 0 (since $U_0(0)=a$), which is impossible. Hence $u^{\varepsilon_k}(x_k,t_k)\neq a$. By extracting a subsequence if necessary, we may assume without loss of generality that $u^{\varepsilon_k}(x_k,t_k)-a$ has a constant sign for $k=0,1,2,\ldots$, say

$$u^{\varepsilon_k}(x_k,t_k)>a\quad(k=0,1,2,\ldots), \tag{7.5}$$

which then implies that $\overline{d^{\varepsilon_k}}(x_k,t_k)>0$ $(k=0,1,2,\ldots)$. Since the mean curvature of $\gamma_t^\varepsilon$ is uniformly bounded for $0\leq t\leq T'$, $0<\varepsilon<<1$, there is a small $\delta>0$ such that each $x$ in a $\delta$-tubular neighbourhood of $\gamma_t^\varepsilon$ has a unique orthogonal projection on $\gamma_t^\varepsilon$. Since the sequence $(x_k)$ remains very close to $\gamma_{t_k}^{\varepsilon_k}$ by (7.4), each $x_k$ (with sufficiently large $k$) has a unique orthogonal projection $p_k=p^{\varepsilon_k}(x_k,t_k)\in\gamma_{t_k}^{\varepsilon_k}$. Let $y_k$ be a point on $\Gamma_{t_k}^{\varepsilon_k}$ that has the smallest distance from $x_k$. If such a point is not unique, we choose one such point arbitrarily. Then we have

$$u^{\varepsilon_k}(y_k,t_k)=a\quad(k=0,1,2,\ldots), \tag{7.6}$$

$$\overline{d^{\varepsilon_k}}(x_k,t_k)=\|x_k-y_k\|\,, \tag{7.7}$$

(7.8) $$u^{\varepsilon_k}(x,t_k) > a \quad \text{if } \|x-x_k\| < \|y_k - x_k\|\,.$$

$$x_k - p_k \perp \gamma^{\varepsilon_k}_{t_k} \quad \text{at } p_k \in \gamma^{\varepsilon_k}_{t_k}\,,$$

Furthermore, (7.4) and (3.9) imply

(7.9) $$\|x_k - p_k\| = \mathcal{O}(\varepsilon_k)\,, \quad \|y_k - p_k\| = \mathcal{O}(\varepsilon_k) \quad (k=0,1,2,\ldots).$$

We now rescale the solution $u^\varepsilon$ around $(p_k, t_k)$ and define

(7.10) $$w^k(z,\tau) := u^{\varepsilon_k}(p_k + \varepsilon_k \mathcal{R}_k z, t_k + \varepsilon_k^2 \tau)\,,$$

where $\mathcal{R}_k$ is a matrix in $SO(n,\mathbb{R})$ that rotates the $z^{(n)}$ axis onto the normal at $p_k \in \gamma^{\varepsilon_k}_{t_k}$, that is,

$$\mathcal{R}_k : (0,\ldots,0,1)^T \mapsto n^{\varepsilon_k}(p_k,t_k)\,,$$

where $(\ )^T$ denotes a transposed vector and $n^\varepsilon(p,t)$ the outward normal unit vector at $p\in\gamma^\varepsilon_t$. Since $\gamma^\varepsilon_t$ (hence the points $p_k$) is uniformly separated from $\partial\Omega$ by some positive distance, there exists $c>0$ such that $w^k$ is defined (at least) on the box

$$B^k := \left\{ (z,\tau)\in\mathbb{R}^n\times\mathbb{R} \,:\, \|z\| \le \frac{c}{\varepsilon_k}, \quad -(\rho-1)\mu_{\varepsilon_k}^{-1}|\ln\varepsilon_k| \le \tau \le \frac{T-T'}{\varepsilon_k^2} \right\}\,,$$

where we recall that $\mu_\varepsilon \to \mu = f'(a) > 0$ as $\varepsilon\to 0$. Since $u^\varepsilon$ satisfies the equation in (1.1), we see that $w^k$ satisfies

(7.11) $$w^k_\tau = \Delta_z w^k + f(w^k) + \varepsilon_k \xi^{\varepsilon_k}(t_k + \varepsilon_k^2\tau) \quad \text{ in } B^k\,.$$

Moreover, if $(z,\tau)\in B^k$ then $t^{\varepsilon_k} \le t_k + \varepsilon_k^2\tau \le T$. Therefore (3.3) implies

(7.12) $$\begin{cases} d^{\varepsilon_k}(p_k + \varepsilon_k\mathcal{R}_k z, t_k + \varepsilon_k^2\tau) \le -C\varepsilon_k & \Rightarrow \quad w^k(z,\tau) \le a_- + \eta\,,\\ d^{\varepsilon_k}(p_k + \varepsilon_k\mathcal{R}_k z, t_k + \varepsilon_k^2\tau) \ge C\varepsilon_k & \Rightarrow \quad w^k(z,\tau) \ge a_+ - \eta\,,\end{cases}$$

as long as $(z,\tau)\in B^k$. Now we recall that the rotation by $\mathcal{R}_k$ of the $z^{(n)}$ axis is normal to $\gamma^{\varepsilon_k}_{t_k}$ at $p_k$, and that the mean curvature of $\gamma^\varepsilon_t$ is uniformly bounded for $0\le t\le T'$, $0<\varepsilon<<1$. Also the normal speed of $\gamma^\varepsilon_t$, given by $V=(n-1)\kappa - \frac{c(\varepsilon\xi^\varepsilon_t)}{\varepsilon}$, is $\mathcal{O}(\varepsilon^{-\gamma'})$ for some $0<\gamma'<\frac13$ in view of $c(\delta) = -c_0\delta + \mathcal{O}(\delta^2)$ as $\delta\to 0$, and (1.12) (if (MN1) noise) or Proposition 1.2 (if (MN2) noise). As a result $d^\varepsilon(x,t)$ satisfies

$$|d^\varepsilon(x,t) - d^\varepsilon(x,t')| \le \frac{\tilde C}{\varepsilon^{\gamma'}}|t-t'|, \quad 0\le t,t'\le T', 0<\varepsilon<<1,$$

for some $\tilde C>0$. From these observations and (7.12), we see that there exists a constant $K>0$, which is independent of $k$, such that

(7.13) $$z^{(n)} \le -K \Rightarrow w^k(z,\tau) \le a_- + \eta\,, \quad z^{(n)} \ge K \Rightarrow w^k(z,\tau) \ge a_+ - \eta\,,$$

for all $(z,\tau)\in B^k$ with $\|z\| \le \sqrt{1/\varepsilon_k}$ and $|\tau| \le 1/({\varepsilon_k}^{1-\gamma'})$.

Now, since $w^k$ solves (7.11), the uniform (w.r.t. $k\ge 0$) boundedness of $w^k$ and standard parabolic estimates, along with the derivative bounds on $\tau\mapsto\varepsilon_k\xi^{\varepsilon_k}(t_k+\varepsilon_k^2\tau)$ (see (1.12) or Proposition 1.2), imply that $w^k$ is uniformly bounded in $C^{2+\gamma,1+\frac{\gamma}{2}}_{loc}(B^1)$. We can therefore extract from $(w^k)$ a subsequence that converges to some $w$ in $C^{2,1}_{loc}(B^1)$. By repeating this on all $B^k$, we can find a subsequence of $(w^k)$ that converges to some $w$ in $C^{2,1}_{loc}(\mathbb{R}^n\times\mathbb{R})$ (note that $\cup_{k\ge 0}B^k = \mathbb{R}^n\times\mathbb{R}$). Passing to the limit in (7.11) yields

$$w_\tau = \Delta_z w + f(w) \quad \text{ on } \mathbb{R}^n\times\mathbb{R}\,.$$

Hence we have constructed an eternal solution $w(z,\tau)$ which—in view of (7.13)—satisfies (7.2). Lemma 7.1 then implies that

$$w(z,\tau) = U_0(z^{(n)} - z^*) \tag{7.14}$$

for some $z^* \in \mathbb{R}$.

Now we define sequences of points $(z_k)$, $(\tilde{z}_k)$ by

$$z_k := \frac{1}{\varepsilon_k}\mathcal{R}_k^{-1}(x_k - p_k)\,, \quad \tilde{z}_k := \frac{1}{\varepsilon_k}\mathcal{R}_k^{-1}(y_k - p_k)\,.$$

By (7.9), these sequences are bounded, so we may assume without loss of generality that they converge:

$$z_k \to z_\infty\,, \quad \tilde{z}_k \to \tilde{z}_\infty\,, \quad \text{as} \;\; k \to \infty\,.$$

By the definition of the $z$ coordinates, $z_\infty$ must lie on the $z^{(n)}$ axis, that is,

$$z_\infty = (0,\dots,0,z_\infty^{(n)})^T\,.$$

It follows from (7.6) and (7.8) that

$$w(\tilde{z}_\infty,0) = a\,, \qquad w(z,0) \geq a \;\text{ if }\; \|z - z_\infty\| \leq \|\tilde{z}_\infty - z_\infty\|\,. \tag{7.15}$$

Note that by (7.14), the level set $w(z,0) = a$ coincides with the hyperplane $z^{(n)} = z^*$, and recall that ${U_0}' > 0$. Therefore, in view of (7.14) and (7.15), we have either $\tilde{z}_\infty = z_\infty$, or that the ball of radius $\|\tilde{z}_\infty - z_\infty\|$ centered at $z_\infty$ is tangential to the hyperplane $z^{(n)} = z^*$ at $\tilde{z}_\infty$. This implies that $\tilde{z}_\infty$, as well as $z_\infty$, must also lie on the $z^{(n)}$ axis. Therefore

$$\tilde{z}_\infty = (0,\dots,0,z^*)^T\,,$$

and the inequality $w(z_\infty,0) \geq a$ implies that $z_\infty^{(n)} \geq z^*$. On the other hand equality (7.7) implies $\overline{d^{\varepsilon_k}}(x_k,t_k)/\varepsilon_k = \|x_k - y_k\|/\varepsilon_k = \|z_k - \tilde{z}_k\| \to \|z_\infty - \tilde{z}_\infty\| = z_\infty^{(n)} - z^*$. The assumption (7.3) then yields

$$\begin{aligned} 0 &= \Big|w(z_\infty,0) - U_0(z_\infty^{(n)} - z^*)\Big| \\ &= \Big|\lim_{k\to\infty} u^{\varepsilon_k}(x_k,t_k) - U_0\Big(\lim_{k\to\infty}\frac{\overline{d^{\varepsilon_k}}(x_k,t_k)}{\varepsilon_k}\Big)\Big| \\ &\geq 2\eta\,. \end{aligned}$$

This contradiction proves statement (ii) of Theorem 3.1. □

7.2. **Proof of (i) in Theorem 3.4.** The proof of (i) below uses an argument similar to the proof of Corollary 4.8 in [29]. Fix $\rho > 1$ and $0 < T' < T$. For a given $\eta \in (0, \min(a - a_-, a_+ - a))$ define $\varepsilon_0 > 0$ and $C > 0$ as in Theorem 3.1. Then we claim that

$$\liminf_{\varepsilon\to 0} \inf_{x\in\mathcal{N}_{C\varepsilon}(\gamma_t^\varepsilon),\, \rho t^\varepsilon \leq t \leq T'} \nabla u^\varepsilon(x,t)\cdot n^\varepsilon(p(x,t),t) > 0\,, \tag{7.16}$$

where $n^\varepsilon(p,t)$ denotes the outward unit normal vector at $p \in \gamma_t^\varepsilon$.

Indeed, assume by contradiction that there exist sequences $\varepsilon_k \downarrow 0$, $t_k \in [\rho t^{\varepsilon_k}, T']$, $x_k \in \mathcal{N}_{C\varepsilon_k}(\gamma_{t_k}^{\varepsilon_k})$ $(k = 1,2,\dots)$ such that

$$\nabla u^{\varepsilon_k}(x_k,t_k)\cdot n^{\varepsilon_k}(p_k,t_k) \leq 0\,,$$

where $p_k = p(x_k, t_k)$. By rescaling around $(p_k, t_k)$ and using arguments similar to those in the proof of (ii), one can find a point $z_\infty$ with $|z_\infty^{(n)}| \le C$ such that

$$U_0{}'(z_\infty^{(n)}) \le 0\,,$$

which contradicts to the fact that $U_0{}' > 0$ and establishes (7.16). Since, in view of Theorem 3.1, $\Gamma_t^\varepsilon \subset \mathcal{N}_{C\varepsilon}(\gamma_t^\varepsilon)$, the estimate (7.16) implies that $\nabla u^\varepsilon(x,t) \neq 0$ for all $x \in \Gamma_t^\varepsilon$; hence by the implicit function theorem, $\Gamma_t^\varepsilon$ is a smooth hypersurface in a neighbourhood of any point on it. The fact that $\Gamma_t^\varepsilon$ can be expressed as a graph over $\gamma_t^\varepsilon$ also follows from (7.16). This proves the statement (i) of Theorem 3.1. □

**Acknowledgments.** M. Alfaro is supported by the ANR I-SITE MUSE, project MICHEL 170544IA (n$^o$ ANR-IDEX-0006). H. Matano is supported by KAKENHI (16H02151). The research work was supported by the Hellenic Foundation for Research and Innovation (H.F.R.I.) under the 'First Call for H.F.R.I. Research Projects to support Faculty members and Researchers and the procurement of high-cost research equipment grant' (Project Number: HFRI-FM17-45).

## References

[1] M. Alfaro, D. Hilhorst, H. Matano, *The singular limit of the Allen–Cahn equation and the FitzHugh-Nagumo system*, J. Differential Equations **245** (2) (2008), 505–565.
[2] M. Alfaro, H. Matano, *On the validity of formal asymptotic expansions in Allen–Cahn equation and FitzHugh-Nagumo system with generic initial data*, DCDS-Series B, **17** (6) (2012), 1639–1649.
[3] S. M. Allen, J. W. Cahn, *A microscopic theory for antiphase boundary motion and its application to antiphase domain coarsening*, Acta Metallurgica **27** (6) (1979), 1085–1095.
[4] S. Angenent, M. Gurtin, *Multiphase thermomechanics with interfacial structure. 2. Evolution of an isothermal interface*, Arch. Rational Mech. Anal. **108** (1989), 323-–391.
[5] H. Berestycki, F. Hamel, *Generalized travelling waves for reaction-diffusion equations* in Perspectives in Nonlinear Partial Differential Equations, in honor of Haïm Brezis, Contemp. Math. **446**, Amer. Math. Soc., Providence, RI, 2007, 101–123.
[6] S. Brassesco, A. De Masi, E. Presutti, *Brownian fluctuations of the interface in the $D = 1$ Ginzburg-Landau equation with noise*, Ann. Inst. H. Poincaré Probab. Statist. **31** (1) (1995), 81–118.
[7] L. Bronsard, R. V. Kohn, *Motion by mean curvature as the singular limit of Ginzburg–Landau dynamics*, J. Differential Equations **90** (2) (1991), 211–237.
[8] J. Carr, R. Pego, *Metastable patterns in solutions of $u_t = \varepsilon^2 u_{xx} - f(u)$*, Comm. Pure Appl. Math. **42**(1989), 523–576.
[9] X. Chen, *Generation and propagation of interfaces for reaction-diffusion equations*, J. Differential Equations **96** (1) (1992), 116–141.
[10] X. Chen, *Generation, propagation, and annihilation of metastable patterns*, J. Differential Equations **206** (2) (2004), 399–437.
[11] X. Chen, D. Hilhorst, E. Logak, *Asymptotic behavior of solutions of an Allen–Cahn equation with a nonlocal term*, Nonlinear Anal. **28** (7) (1997), 1283–1298.
[12] G. Da Prato, J. Zabczyk, Stochastic equations in infinite dimensions, *Cambridge University Press, Cambridge 1992.*
[13] N. Dirr, S. Luckhaus, M. Novaga, *A short selection principle in case of fattening for curvature flow*, Calc. Var. PDE **13** (4) (2001), 405–425.
[14] W. G. Faris, G. Jona-Lasinio, *Large fluctuations for a nonlinear heat equation with noise*, J. Phys. A **15** (10) (1982), 3025–3055.
[15] P. C. Fife, J. B. McLeod, *The approach of solutions of nonlinear diffusion equations to travelling front solutions*, ARMA, **65** (4) (1977), 335–361.
[16] T. Funaki, *The scaling limit for a stochastic PDE and the separation of phases*, Probab. Theory Related Fields, **102** (2) (1995), 221–288.

[17] T. Funaki, *Singular Limit for Stochastic Reaction-Diffusion Equation and Generation of Random Interfaces*, Acta Mathemaica Sinica, **15** (3) (1999), 407–438.
[18] T. Funaki, *Lectures on Random Interfaces*, Springer Briefs in Probability and Mathematical Sciences, Springer-Verlag, 2016.
[19] D. Gilbarg, N. Trudinger, *Elliptic Partial Differential Equations of Second Order*, Springer-Verlag: Berlin, 1977.
[20] T. Ilmanen, *Convergence of the Allen–Cahn equation to Brakke's motion by mean curvature*, J. Differential Geom. **38** (2) (1993), 417–461.
[21] K. Kawasaki, T. Ohta, *Kinetic drumhead model of interface I*, Progress of Theoretical Physics **67** (1982) 147–163.
[22] K. Lee, *Generation and motion of interfaces in one-dimensional stochastic Allen–Cahn equation*, J. Theor. Probab. **31** (2018), 268–293.
[23] K. Lee, *Generation of interfaces for multi-dimensional stochastic Allen-Cahn equation with a noise smooth in space*, to appear in Stochastics.
[24] P.-L. Lions, P. E. Souganidis, *Fully nonlinear stochastic partial differential equations* , C. R. Acad. Sci. Paris Sér. I Math. **326** (1998), 1085–1092.
[25] P.-L. Lions, P. E. Souganidis, *Fully nonlinear stochastic partial differential equations: non-smooth equations and applications*, C. R. Acad. Sci. Paris Sér. I Math. **327** (1998), 735–741.
[26] P.-L. Lions, P. E. Souganidis, *Fully nonlinear stochastic pde with semilinear stochastic dependence*, C. R. Acad. Sci. Paris Sér. I Math. **331** (2000), 617–624.
[27] P.-L. Lions, P. E. Souganidis, *Uniqueness of weak solutions of fully nonlinear stochastic partial differential equations*, C. R. Acad. Sci. Paris Sér. I Math. **331** (2000), 783–790.
[28] P.-L. Lions, P. E. Souganidis, *The asymptotics of stochastically perturbed reaction-diffusion equations and front propagation*, arXiv:1909.05673v1, (2019).
[29] H. Matano, M. Nara, *Large time behavior of disturbed planar fronts in the Allen–Cahn equation*, J. Differential Equations **251** (12) (2011), 3522–3557.
[30] P. de Mottoni, M. Schatzman, *Development of interfaces in* $\mathbb{R}^n$, Proc. Roy. Soc. Edinburgh **116A** (3–4) (1990), 207–220.
[31] P. de Mottoni, M. Schatzman, *Geometrical evolution of developed interfaces*, Trans. Amer. Math. Soc. **347** (5) (1995), 1533–1589.
[32] H. Spohn, *Large scale dynamics of interacting particles*, Springer 1991.
[33] J. Walsh, An introduction to stochastic partial differential equations, *École d'Été de Probabilités de Saint Flour XIV-1984 265–439.*
[34] H. Weber, *On the short time asymptotic of the stochastic Allen–Cahn equation*, Ann. Inst. Henri Poincaré Probab. Stat. **46** (4) (2010), 965–975.
[35] H. Weber, *Sharp interface limit for invariant measures of a stochastic Allen–Cahn equation*, Comm. Pure Appl. Math. **63** (8) (2010), 1071–1109.
[36] N. K. Yip, *Stochastic motion by mean curvature*, Arch. Rational Mech. Anal., **144** (4) (1998), 313–355.

Matthieu Alfaro, Université de Rouen Normandie, CNRS, Laboratoire de Mathématiques Raphaël Salem, Saint-Etienne-du-Rouvray, France & BioSP, INRAE, 84914, Avignon, France.
*Email address*: matthieu.alfaro@univ-rouen.fr

Dimitra Antonopoulou, Department of Mathematical and Physical Sciences, University of Chester, CH2 4NU, Chester, UK, and, Institute of Applied and Computational Mathematics, FORTH, GR–711 10 Heraklion, Greece.
*Email address*: d.antonopoulou@chester.ac.uk

Georgia Karali, Department of Mathematics and Applied Mathematics, University of Crete, GR–714 09 Heraklion, Greece, and, Institute of Applied and Computational Mathematics, FORTH, GR–711 10 Heraklion, Greece.
*Email address*: gkarali@uoc.gr

Hiroshi Matano, Meiji Institute for Advanced Study of Mathematical Sciences, Meiji University, 4-21-1 Nakano, Tokyo 164-8525, Japan.
*Email address*: matano@meiji.ac.jp