

# Convergence of Diffusion Generated Motion to Motion by Mean Curvature

Drew Swartz

IRI, Chicago, IL, USA

Nung Kwan Yip

Department of Mathematics, Purdue University, USA

September 15, 2017

## Abstract

We provide a new proof of convergence to motion by mean curvature (MMC) for the Merriman-Bence-Osher (MBO) thresholding algorithm. The proof is elementary and does not rely on maximum principle for the scheme. The strategy is to construct a natural ansatz of the solution and then estimate the error. The proof thus also provides a convergence rate. Only some weak integrability assumptions of the heat kernel, but not its positivity, is used. Currently the result is proved in the case when smooth and classical solution of MMC exists.

## 1 Introduction

Motion by mean curvature (MMC) is a fundamental geometric motion arising in a broad range of scientific disciplines. Besides its intrinsic geometric interests, in applications, it arises naturally in modeling the evolution of interfaces such as grain boundaries which are subject to the effect of surface tension. It also appears in various aspects of image de-noising algorithms. Mathematically, MMC describes the evolution of a manifold with its normal velocity  $V_N$  equal to its mean curvature, i.e.

$$V_N = H = \frac{1}{n} \sum_{i=1}^n \kappa_i, \quad (1)$$

where the  $\kappa_i$ 's are the principal curvatures of the manifold at a point. This evolution can also be thought of as the geometric analogue of the heat equation. Specifically, given an initial  $n$ -dimensional embedded manifold  $M_0$  in  $\mathbb{R}^{n+1}$ , the time dependent manifolds  $M_t = F(M_0, t)$  with  $F : M_0 \times \mathbb{R}_+ \rightarrow \mathbb{R}^{n+1}$  solves the MMC equation (1) with initial data  $M_0$  if  $F$  satisfies the following equation,

$$\partial_t F = \Delta_{M_t} F, \quad F(\cdot, 0) = F_0, \quad (2)$$

where  $\Delta_{M_t}$  denotes the Laplace-Beltrami operator on the manifold  $M_t$ . Alternatively, MMC can be interpreted as the  $L^2$ -gradient flow, or steepest descent for the area functional of a manifold.

It is known that this geometric flow can lead to singularity formation and topological changes for the underlying evolving manifold. Thus it is desirable to use mathematical formulations which can handle such events. One method is the *phase-field* approach. The underlying equation is typically given by the *Allen-Cahn equation*

$$\frac{\partial u}{\partial t} = \Delta u - \frac{1}{\epsilon^2} W'(u),$$

where  $W(u) = (1 - u^2)^2$  is the double-well potential (with 1 and  $-1$  being the two minima of  $W$ ). As  $u$  evolves under this equation, the domain will separate into two regions/phases where  $u$  is approximately equal to 1 in one region and  $-1$  in the other. Between these two regions,  $u$  will have a diffuse transition layer of thickness  $\mathcal{O}(\epsilon)$ . For  $\epsilon \ll 1$ , this layer will in turn evolve approximately by its mean curvature. Convergence of this motion to MMC as  $\epsilon \rightarrow 0$  has been shown rigorously by several authors using various approaches, for example, [21, 30, 31] (varifold formulation), [5] (energy approach), [9] (asymptotic expansion), [6] (sub- and super-solutions), [16] (viscosity solution).

Another formulation for MMC is to make use of a level set function  $u : \mathbb{R}^{n+1} \times \mathbb{R}_+ \rightarrow \mathbb{R}$  which solves the following equation,

$$\frac{\partial u}{\partial t} = |\nabla u| \operatorname{div} \left( \frac{\nabla u}{|\nabla u|} \right).$$

Each  $c$ -level set of  $u$ ,  $\{u = c\}$ , or  $\partial\{u \geq c\}$ , then evolves by its mean curvature (in the viscosity sense). The theory surrounding this equation has been developed independently by Evans-Spruck ([17]) and Chen-Giga-Goto ([7]). See [2] for more recent development, in particular the relation between the phase-field and level set formulations.

The *diffusion generated motion* for approximating mean curvature flow was first proposed by Merriman-Bence-Osher in [28] which hereby from now on will be called the MBO algorithm or scheme. Essentially it is a time-splitting algorithm for the Allen-Cahn equation. Its description is as follows.

Let  $h > 0$  be a small positive number, called the time step. Given an initial set  $\Omega_0 \subseteq \mathbb{R}^{n+1}$ , a new set  $\Omega_h$  and its boundary  $\Gamma_h = \partial\Omega_h$  are constructed by the following two simple procedures:

**Step 1 - linear diffusion:** given a set  $\Omega_0 \subseteq \mathbb{R}^{n+1}$  and let

$$U_0(x) = 2 \cdot \mathbb{1}_{\Omega_0}(x) - 1 = \begin{cases} 1 & \text{for } x \in \Omega_0, \\ -1 & \text{for } x \notin \Omega_0. \end{cases} \quad (3)$$

Then solve the linear heat equation:

$$U_t - \Delta U = 0, \text{ in } \mathbb{R}^{n+1} \times (0, h), \quad U(x, 0) = U_0(x). \quad (4)$$

**Step 2 - thresholding:** project  $U(\cdot, h)$ , the solution from **Step 1**, onto  $\{-1, 1\}$  by setting

$$U(x, h^+) = \begin{cases} 1 & \text{if } U(x, h) \geq 0, \\ -1 & \text{if } U(x, h) < 0. \end{cases} \quad (5)$$

Then define

$$\Omega_h = \{x : U(x, h^+) \geq 0\} \quad \text{and} \quad \Gamma_h = \partial\Omega_h. \quad (6)$$

Note that  $U(x, h^+) = 2\mathbb{1}_{\Omega_h}(x) - 1$ .

The algorithm then repeats **Steps 1** and **2** but using the result from **Step 2** as the initial data for **Step 1**. The following collection of sets are thus generated:

$$\left\{ \Omega_{kh} = \{x : U(x, kh^+) \geq 0\}, \quad \Gamma_{kh} = \partial\Omega_{kh} \right\}_{k=0,1,2,\dots}. \quad (7)$$

It is proved in [15] and [1] that as  $h \rightarrow 0$ , the sequence (7) converges to a solution of MMC in the viscosity sense. See also [22, 26, 23] for generalizations to incorporate general kernels and anisotropy effects. The recent works [13, 12, 25] have recast thresholding schemes into a variational setting.

In this paper, we provide a new convergence proof of the algorithm. Our approach is elementary and does not rely on the theory of viscosity solution which depends very much on comparison

principle. Furthermore, it provides a convergence rate. Even though so far it has only been applied to the case when smooth and classical solution of MMC exists, it has the potential to be used in the study of thresholding schemes for other geometric motions. These include fourth order flows such as Willmore and surface diffusion flows [14], and higher co-dimension mean curvature flows such as filament motions in  $\mathbb{R}^3$  [29]. Furthermore, the works [3, 13], for anisotropic MMC shows that the convolution kernel used in the thresholding scheme are necessarily non-positive. Consequently comparison principle does not hold. Our proof can offer reasonable generalizations to these settings.

## 2 Main Results and Outline of Proof

Recall that the MBO scheme produces, for each  $h > 0$ , a sequence of sets and their boundaries (7). Our main result is that the  $\Gamma_{kh}$  converges to the solution of MMC as long as the classical solution exists. The precise statement of result is given in the following. All the definitions used will be elaborated afterwards.

**Theorem 1.** *Let  $\Gamma_0$  be a compact, smooth embedded  $n$ -dimensional manifold in  $\mathbb{R}^{n+1}$ . Then there is a time  $T > 0$  for which the classical solution of MMC starting from  $\Gamma_0$  exists on  $[0, T]$  such that the sequence  $\{\Gamma_{kh}\}_{0 \leq k \leq \lfloor \frac{T}{h} \rfloor}$  remains an embedded manifold and as  $h \rightarrow 0$ , converges to the solution of MMC in the Hausdorff distance and also in the Bounded-Variation (BV)-sense. The time  $T$  depends on the initial manifold  $\Gamma_0$  but not on the time step  $h$ .*

Next we give the definitions of Hausdorff distance, BV-convergence and some remarks about the theorem.

**Definition 1** (Hausdorff Distance  $d_{\mathcal{H}}(\cdot)$ ). *Let  $A$  and  $B$  be two subsets of  $\mathbb{R}^{n+1}$ . Then the Hausdorff distance between  $A$  and  $B$  is defined as:*

$$d_{\mathcal{H}}(A, B) = \max \left\{ \sup_{y \in B \setminus A} \text{dist}(y, A), \sup_{x \in A \setminus B} \text{dist}(x, B) \right\} \quad (8)$$

where  $\text{dist}(y, A) = \inf_{x \in A} |y - x|$  and likewise,  $\text{dist}(x, B) = \inf_{y \in B} |x - y|$ .

It is known that  $d_{\mathcal{H}}$  gives rise to a complete metric space. In addition, for  $\Omega_1, \Gamma_1 = \partial\Omega_1$  and  $\Omega_2, \Gamma_2 = \partial\Omega_2$ , it holds that  $d_{\mathcal{H}}(\Omega_1, \Omega_2) = d_{\mathcal{H}}(\Gamma_1, \Gamma_2)$ . Hence it does not matter if we are using  $\Omega$  or  $\Gamma$  to measure the Hausdorff distance.

To formulate the notion of convergence, we define

$$\Gamma_h(t) = M(\Gamma_{kh}, t - kh), \quad \text{for each } k \geq 1, \text{ and } kh \leq t < (k + 1)h \quad (9)$$

where  $M(\Gamma_{kh}, t - kh)$  is the solution at time  $t - kh$  of the classical MMC (2) with initial data  $\Gamma_{kh}$ . Note that due to the thresholding step,  $\Gamma_h(t)$  is discontinuous in time, in particular, at each  $t = kh$ :  $\Gamma_h(kh^-) \neq \Gamma_h(kh^+)$ . Let  $\Omega_h(t)$  be the interior of  $\Gamma_h(t)$ , i.e. the bounded subset of  $\mathbb{R}^{n+1}$  such that  $\Gamma_h(t) = \partial\Omega_h(t)$ . Then we define

$$\chi_h : \mathbb{R}^{n+1} \times \mathbb{R}_+ \longrightarrow \{0, 1\} : \quad \chi_h(x, t) = \mathbb{1}_{\Omega_h(t)}(x). \quad (10)$$

(The above definition is to facilitate a more efficient usage of integration by parts formula for MMC, in particular (104)-(105). But from the perspective of understanding the convergence statement, it is essentially the same as if we had used the ‘‘piece-wise-constant’’ definition:  $\Gamma_h(t) = \Gamma_{kh}, \Omega_h(t) = \Omega_{kh}$  for  $kh < t < (k + 1)h$ .)

**Definition 2** (BV-Convergence to MMC). [27] *There is a  $0 < T < \infty$  and a*

$$\chi^* \in L^\infty\left(0, T; BV(\mathbb{R}^{n+1}, \{0, 1\})\right) \quad (11)$$

*such that as  $h \rightarrow 0$ ,  $\chi_h$  converges to  $\chi^*$  in  $L^1(\mathbb{R}^{n+1} \times [0, T])$ . Furthermore there exists a function  $v \in L^1(0, T; L^1(|\nabla\chi^*|))$  such that for all  $\zeta \in C^\infty(\Lambda \times [0, T], \mathbb{R}^n)$  with  $\zeta = 0$  on  $\partial\Lambda \times [0, T]$  and  $\xi \in C^\infty(\Lambda \times [0, T], \mathbb{R})$  with  $\xi = 0$  on  $\partial\Lambda \times [0, T] \cup \Lambda \times \{0\}$ , the following two properties hold,*

$$\int_0^T \int_\Lambda \left( \operatorname{div}\zeta - \frac{\nabla\chi^*}{|\nabla\chi^*|} \nabla\zeta \frac{\nabla\chi^*}{|\nabla\chi^*|} \right) |\nabla\chi^*| = - \int_0^T \int_\Lambda v\zeta \cdot \nabla\chi^*, \quad (12)$$

$$\int_0^T \int_\Lambda \chi^* \partial_t \xi + \int_\Lambda \chi_0 \xi(0) = - \int_0^T \int_\Lambda v\xi |\nabla\chi^*|. \quad (13)$$

*In the above,  $\Lambda$  is a fixed and sufficiently large ball in  $\mathbb{R}^{n+1}$ .*

Some remarks are in place.

(i) The property of  $\chi^*$  above implies that for a.e.  $t$ ,  $\chi^*(t) = \mathbb{1}_{\Omega^*(t)}$  for some  $\Omega^*(t) \subseteq \mathbb{R}^{n+1}$  which is a *set of finite perimeter*. In fact, since the results are proved within the realm of classical solution, we actually have  $\chi^* \in C\left(0, T; BV(\mathbb{R}^{n+1}, \{0, 1\})\right)$ . Furthermore, the set  $\Omega^*(t)$ , or more exactly its boundary  $\Gamma^*(t) = \partial\Omega^*(t)$ , evolves smoothly in time. Identity (13) shows that  $v(t)$  is the velocity function of  $\Gamma^*(t)$  and (12) shows that  $v(t)$  is given by the mean curvature of  $\Gamma^*(t)$ . We refer to [27] and also Section 5 for more detail explanation about these concepts and notations.

(ii) For some quantitative estimate and reasoning, more precise condition on the initial manifold  $\Gamma_0$  will be given in Section 2.6.

(iii) The time  $T$  appearing in the Theorem depends only on geometric properties of the initial manifold  $\Gamma_0$  but not on  $h$ . We have not yet shown that  $T$  coincides with the maximal time  $T^*$  for which the smooth MMC flow starting from  $\Gamma_0$  exists. This is because with the current estimate, we have only  $C^{1,\alpha}$ -convergence (in space) of the  $\Gamma_h$ . We expect that this can be improved to  $C^{2,\alpha}$ -convergence so that the curvature of  $\Gamma_h$  will also converge. Then  $T$  would be the same as  $T^*$ . See the remark (ii) after **Theorem 4** for more detail discussion.

(iv) The convergence statement is that  $\Gamma_h$  converges to a solution of MMC in the BV-sense. It will be ideal if we can show that this solution coincides with the classical (strong) solution. Again, this can be done if we can demonstrate  $C^{2,\alpha}$ -convergence.

In order to present the strategy of our proof, we introduce the following notations. Let  $\Gamma$  be a smooth compact  $n$ -dimensional embedded manifold in  $\mathbb{R}^{n+1}$ . We use  $M(M_0, t)$ , or simply  $M_t$  if the choice of  $M_0$  is clear, to denote the *solution of MMC at time  $t$  starting from  $M_0$* . By regularity theory of MMC,  $M_t$  is a smooth embedded manifold for small  $t > 0$ . Each manifold  $\Gamma_{kh}$ , obtained from the MBO scheme will be called a *numerical manifold*.

The proof of Theorem 1 relies on two main results: *consistency* and *stability* statements. The former states that for  $h \ll 1$ , the Hausdorff distance between  $M(\Gamma_{kh}, h)$  and  $\Gamma_{(k+1)h}$  is of order  $o(h)$ . The exact order of  $o(h)$  will be made precise during the proof. The latter states that the curvatures of the numerical manifolds  $\Gamma_{kh}$  are uniformly bounded so that the error does not accumulate over repeated iterations. We next give an outline of the proof.

## 2.1 Step I. Construction of ansatz.

Recall that at each **Step 1** of the MBO scheme, we solve the linear heat equation

$$U_t - \Delta U = 0, \quad \text{for } (x, t) \in \mathbb{R}^{n+1} \times (0, h), \quad (14)$$

$$U(\cdot, 0) = \begin{cases} 1 & \text{if } x \in \Omega_0 \\ -1 & \text{if } x \notin \Omega_0 \end{cases}, \quad (15)$$

where  $\Omega_0$  is an open set in  $\mathbb{R}^{n+1}$  with smooth boundary  $\Gamma_0 = \partial\Omega_0$ . The main idea is to formulate an appropriate ansatz for the solutions to (14)-(15) which can be easily compared to the solution of MMC.

For this, we will make use of  $M(\Gamma_0, t)$ , the solution of MMC starting from  $\Gamma_0$ . By regularity of MMC, for a short time, the solution  $M(\Gamma_0, t)$  will remain smooth and continue to bound a set  $\Omega(t)$ . Now let  $r(\cdot, t)$  be the signed distance function to  $M(\Gamma_0, t)$ , with the convention that  $r > 0$  in the interior  $\Omega(t)$  of  $M(\Gamma_0, t)$ :

$$r(x, t) = \text{sdist}(x, M(\Gamma_0, t)) = \begin{cases} \text{dist}(x, M(\Gamma_0, t)), & \text{for } x \in \Omega(t), \\ -\text{dist}(x, M(\Gamma_0, t)), & \text{for } x \notin \Omega(t). \end{cases} \quad (16)$$

Then we decompose  $U$  as

$$U = U^0 + U^1 \quad (17)$$

where the leading order term  $U^0$  is given by

$$U^0(x, t) = \frac{2}{\sqrt{\pi}} \int_0^{\frac{r(x,t)}{2\sqrt{t}}} \exp(-y^2) dy. \quad (18)$$

Note that  $U^0$  is the modified error function which solves the following one dimensional linear heat equation in  $(r, t)$ :

$$U_t^0 = \partial_{rr} U^0, \quad U^0(r, 0) = 2 \cdot \mathbb{1}_{(0, \infty)}(r) - 1. \quad (19)$$

For future reference, we write down the following formula:

$$\partial_r U^0 = \frac{1}{\sqrt{\pi t}} \exp\left(-\frac{r^2}{4t}\right). \quad (20)$$

The error term of the above ansatz is given by the function  $U^1 = U - U^0$ . Since  $U$  solves the linear heat equation, we have that

$$U_t^1 - \Delta U^1 = -(U_t^0 - \Delta U^0).$$

Furthermore, as the initial data for  $U$  is the same as that for  $U^0$ , we have  $U^1(\cdot, 0) \equiv 0$ . Hence by means of variation of parameters,  $U^1$  can be represented as

$$U^1(x, t) = - \int_0^t \int_{\mathbb{R}^{n+1}} G(x-y, t-\tau) (\partial_\tau - \Delta) U^0(y, \tau) dy d\tau, \quad (21)$$

where  $G(x, t) = \frac{e^{-\frac{|x|^2}{4t}}}{(4\pi t)^{\frac{n+1}{2}}}$  is the heat kernel defined on  $\mathbb{R}^{n+1} \times \mathbb{R}_+$ . The explicit representations (18) and (21) allow us to carry out a detailed analysis of the solution to (14)-(15).

## 2.2 Step II. Consistency estimate.

The consistency result in this paper is phrased in terms of the Hausdorff distance  $d_{\mathcal{H}}$  between  $\Gamma_h$  and  $M(\Gamma_0, h)$ . This is stated by the following estimate (**Theorem 2**):

$$d_{\mathcal{H}}(\Gamma_h, M(\Gamma_0, h)) \leq \|\mathcal{B}_0\|^2 \mathcal{O}(h^{\frac{3}{2}}). \quad (22)$$

Here  $\|\mathcal{B}_0\|^2$  is a norm placed on the Weingarten map  $\mathcal{B}_0$  for  $\Gamma_0$  which incorporates curvature information of  $\Gamma_0$  (see **Section 2.5** for definition). Heuristically this means that  $\Gamma_h$  lies within a tubular neighborhood of  $M(\Gamma_0, h)$  of radius  $\|\mathcal{B}_0\|^2 \mathcal{O}(h^{\frac{3}{2}})$ .

The proof of (22) makes use of the following formula (**Lemma 1**, (43)):

$$(\partial_t - \Delta)U^0(x, t) = \frac{1}{\sqrt{\pi t}} \exp\left(-\frac{r(x, t)^2}{4t}\right) \sum_{i=1}^n \frac{r(x, t)\kappa_i^2(\Pi(x), t)}{1 - r(x, t)\kappa_i(\Pi(x), t)}$$

where the  $\kappa_i(\Pi(x), t)$ 's are the principal curvatures of the point  $\Pi(x)$  on  $M(\Gamma_0, t)$  which is closest to  $x$ . By (21), the above gives  $\|U^1\|_{L^\infty(\mathbb{R}^{n+1})} \leq \|\mathcal{B}_0\|^2 \mathcal{O}(h)$ .

Next, near  $M(\Gamma_0, h)$ ,  $U^0(\cdot, h)$  roughly equals  $\frac{r(\cdot, h)}{\sqrt{h}}$ . Hence,  $\Gamma_h$ , which is given by the zero set of  $U^0(\cdot, h) + U^1(\cdot, h)$ , corresponds to

$$r(\cdot, h) = \|\mathcal{B}_0\|^2 \mathcal{O}(h^{\frac{3}{2}}) \ll \mathcal{O}(h) \quad (23)$$

giving the consistency of the scheme (see Figure 1).

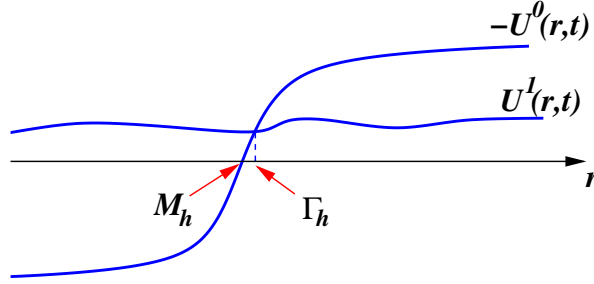


Figure 1: Illustration of  $U^0$ ,  $U^1$ . Note that  $U^0$  vanishes at  $M_h = M(\Gamma_0, h)$ , the solution at time  $h$  of MMC, starting from  $M_0$ . The intersection between  $-U^0$  and  $U^1$  gives  $\Gamma_h$ .

The convergence rate in (23) can in fact be improved to  $\mathcal{O}(h^2)$ . Estimate (23) simply makes use of the  $L^\infty$ -norm of  $(\partial_t - \Delta)U^0(x, t)$ . The improved rate comes by performing a more precise point-wise analysis. This will involve more analytical computation but it is quite similar to what is done in the stability estimates – see Remark 1 in Section 4.3.

### 2.3 Step III. Stability Estimates

Note that the consistency statement (22) involves the curvature of  $\Gamma_h$ . In order to ensure that such a statement can be extended to multiple iterations, we need to estimate the curvature of  $\Gamma_h$  in terms of the curvature of  $\Gamma_0$ .

The first step in doing this is to describe  $\Gamma_h$  as a graph over  $M_h = M(\Gamma_0, h)$ . This is achievable due to the fact that near  $M_h$ , we have

$$\nabla U \approx \partial_r U^0 = \frac{1}{\sqrt{\pi h}} \exp\left(-\frac{r^2}{4h}\right) \gg 1 \quad (\text{since } r \sim \mathcal{O}(h^{\frac{3}{2}})).$$

Hence Implicit Function Theorem (IFT) gives us that locally  $\Gamma_h$  can be written as a graph of a function  $g$  over the tangent plane  $\mathbb{T}_{M_h}(p_0)$  at some point  $p_0$  on  $M_h$  (see Figure 2). The next step is to relate the curvature of  $\Gamma_h$  to the second derivatives of  $g$ .

The necessary computations are also facilitated by IFT. Letting  $\mathcal{B}_h$  be the Weingarten map of  $\Gamma_h$ , and  $\bar{x}$  be the coordinates of the tangent plane  $\mathbb{T}_{M_h}(p_0)$ , we have (61, 62, 63)

$$\mathcal{B}_h = \nabla_{\bar{x}}^2 g + \nabla_{\bar{x}}^2 U^1 \mathcal{O}(\sqrt{h}) + \|\mathcal{B}_0\|^3 \mathcal{O}(h). \quad (24)$$

Careful analysis of  $\nabla_{\bar{x}}^2 U^1$  via (21) and (43) gives (**Lemma 7**):

$$\|\nabla_{\bar{x}}^2 U^1\|_{L^\infty} \leq \|\mathcal{B}_0\|^3 \mathcal{O}(\sqrt{h}) \quad (25)$$

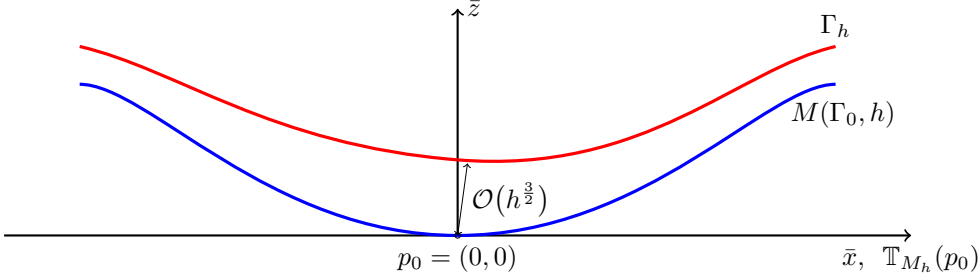


Figure 2: The manifolds  $M_h = M(\Gamma_0, h)$  and  $\Gamma_h$ . Note that  $d_{\mathcal{H}}(M_h, \Gamma_h) = \mathcal{O}(h^{\frac{3}{2}})$  and  $\Gamma_h$  locally can be written as a graph over  $\mathbb{T}_{M_h}$ :  $\Gamma_h = \{(\bar{x}, g(\bar{x})) : \bar{x} \in \mathbb{T}_{M_h}\}$ .

leading to the following *one-step stability estimate* (**Theorem 3**):

$$\|\mathcal{B}_h\| \leq \|\mathcal{B}_0\| + \|\mathcal{B}_0\|^3 \mathcal{O}(h). \quad (26)$$

The most daunting computation is the estimate (25). This is because (21) expresses  $U^1$  in terms of higher order derivatives of  $U^0$ . The analysis thus needs to make crucial use of some regularity theory of MMC, in particular the decay estimates for the Weingarten map of  $M_t$  (**Lemmas 2 and 3**).

Once we have (25), it can be iterated by means of a discrete Gronwall inequality. This leads to that the curvatures of our numerical manifolds  $\Gamma_{kh}$  are bounded uniformly over multiple iterations of the scheme (**Theorem 4**).

## 2.4 Step IV. Convergence.

By the consistency statement and curvature bound from the previous steps, together with some geometric argument to exclude the occurrence of self-intersection, the sequence of  $\Gamma_{nh}$  can be shown to converge to a limit manifold in the Hausdorff distance and also the  $C^{1,\alpha}$ -norm. The final step is to identify the equation satisfied by the limit.

We find the weak formulation of MMC using BV-functions or sets of finite perimeter as used in [27] the most convenient for our purpose. Using the notation of Theorem 1, we show that  $\chi_{kh}$  converges to a limiting function  $\chi_t^* \in BV(\mathbb{R}^{n+1}; \{0, 1\})$  as  $kh \rightarrow t$ . Furthermore,  $\chi^*$  solves MMC in the sense of (12)-(13). The key step in establishing this is to prove that the area converges in the sense that

$$\int |\nabla \chi_{kh}| \rightarrow \int |\nabla \chi_t^*|. \quad (27)$$

(The above is assumed in [27].) The main ingredient in doing this is the *Ball Lemma* (**Lemma 8**) by which we may place a tubular neighborhood with uniform thickness such that  $\Gamma_{kh}$  remains embedded.

## 2.5 Some notations from geometry of surfaces

The reference for this section is [10], in particular Appendix A. From our perspective and application point of view, we take the definition of the mean curvature  $H$  as the negative first variation of the area functional. But for the sake of performing analytical computation, we will relate  $H$  to the Weingarten map of a manifold. Essentially  $H$  is defined to be the trace of the Weingarten map.

Recall that for an  $n$ -dimensional manifold  $M = \partial\Omega$  embedded in  $\mathbb{R}^{n+1}$  given by an embedding map:  $M = F(D)$  where  $F : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^{n+1}$ , the second fundamental form is the symmetric bilinear form on the tangent bundle  $TM$  of  $M$  given by

$$A_{ij} = \langle \partial_i \mathbf{N}, \partial_j F \rangle = -\langle \mathbf{N}, \partial_{ij}^2 F \rangle \quad (28)$$

where  $\mathbf{N}$  is the unit outward normal for  $M$ . Inherently related is the *Weingarten map*, which is the mapping  $L$  from  $TM$  to  $TM$  determined by,

$$L(\mathbf{u}) = -\nabla_{\mathbf{u}} \mathbf{N}. \quad (29)$$

In the coordinate system determined by  $F$ , the matrix corresponding to the Weingarten map is given by,

$$A_j^i = g^{ik} A_{kj}, \quad (30)$$

where  $(g^{ij}) = (g_{ij})^{-1} = (\langle \partial_i F, \partial_j F \rangle)^{-1}$ . The eigenvalues  $\kappa_1, \dots, \kappa_n$  of  $L$  are called the principal curvatures of  $M$ .

With the above notations, the mean curvature  $H$  of  $M$  is given by

$$H = \frac{1}{n} \operatorname{div}_M \mathbf{N} := \frac{1}{n} \sum_{i,j} g^{ij} \langle \partial_i \mathbf{N}, \partial_j F \rangle \quad (31)$$

which can be related to the trace of the Weingarten map of  $M$  as follows,

$$H = \frac{1}{n} \sum_i A_i^i = \frac{1}{n} (\kappa_1 + \dots + \kappa_n). \quad (32)$$

Of particular relevance is the case when  $F$  is the graph of a function  $f$  over  $\mathbb{R}^n$ , i.e.  $F(x) = (x, f(x))$  for  $x \in \mathbb{R}^n$ . In this case, we have

$$A_{ij} = \frac{\nabla^2 f}{\sqrt{1 + |\nabla f|^2}},$$

$$(g_{ij}) = I + \nabla f \otimes \nabla f,$$

and

$$(A_j^i) = \left( I - \frac{\nabla f \otimes \nabla f}{1 + |\nabla f|^2} \right) \frac{\nabla^2 f}{\sqrt{1 + |\nabla f|^2}}. \quad (33)$$

The mean curvature  $H$  is then given by,

$$H = \frac{1}{n} \sum_{i=1}^n A_i^i = \operatorname{div} \frac{\nabla f}{\sqrt{1 + |\nabla f|^2}}. \quad (34)$$

For the rest of this paper, we will use  $A_M$  (or simply  $A$ ) to denote the *Weingarten map* of some general manifold  $M$ . But to emphasize the importance of the numerical manifolds  $\Gamma_{kh}$ 's, we will use  $\mathcal{B}_{kh}$  to denote their Weingarten maps. For the analysis in the rest of the paper, we will use the following norm for the Weingarten map of a manifold  $M$ ,

$$|A(p)| = \sqrt{\operatorname{tr}(AA^T)} = \sqrt{\sum_{i,j} A_i^j(p) A_j^i(p)}, \quad \text{and} \quad \|A\| = \sup_{p \in M} |A(p)|. \quad (35)$$

A useful observation is that if the Weingarten map is bounded, then we can have a quantitative estimate about the size of the region over which the manifold can be written as a graph. To be



specific, let  $p \in M$ ,  $\mathbb{T}_p(M)$  be its tangent plane, and locally near  $p$ ,  $f$  be the graph of  $M$  over  $\mathbb{T}_p(M)$ . From (33), we have

$$\nabla^2 f = \sqrt{1 + |\nabla f|^2} (I + \nabla f \otimes \nabla f) (A_j^i).$$

Hence  $\|\nabla^2 f\| \leq C (1 + |\nabla f|^2)^{\frac{3}{2}} \|A\|$ . Upon integrating in space, there is a universal constant  $C_* = C_*(n)$  such that

$$\left\{ x \in \mathbb{T}_p(M), |x| \leq \frac{C_*}{\|A\|} \right\} \subseteq \{x : |\nabla f(x)| < \infty\}. \quad (36)$$

The above implies that the connected component of  $M \cap B_{\frac{C_*}{\|A\|}}(p)$  containing  $p$  is completely given by the graph of  $f$ .

## 2.6 Time Step Size $h$ in Relation to the Geometry of Initial Manifold $\Gamma_0$

Here we discuss the requirement for the initial manifold  $\Gamma_0$  which is a compact smooth embedded  $n$ -dimensional manifold in  $\mathbb{R}^{n+1}$ . The time step  $h$  will be sent to zero in the convergence statement. But to be quantitative, we will specify its smallness with respect to two geometric quantities pertaining to  $\Gamma_0$ . In the following, we introduce a small constant  $\delta$  such that  $h \leq \delta$ .

The first requirement is a local property. The value of  $\delta$  satisfies

$$\|\mathcal{B}_0\|(\delta |\log \delta|)^{\frac{1}{4}} \leq 1. \quad (37)$$

The second is a more global condition. To describe this precisely, we first define the following *Ball Property*.

**Definition 3. Ball Property.** *Given an embedded  $n$ -dimensional manifold  $M$  which is the boundary of a subset  $\Omega \subseteq \mathbb{R}^{n+1}$ , i.e.,  $M = \partial\Omega$ , we say that  $M$  satisfies the ball property with radius  $r$  if for every  $p \in M$ , there are two balls  $B_{r,p}^{\text{int}}$  and  $B_{r,p}^{\text{ext}}$  (interior and exterior) with radius  $r$  such that  $B_{r,p}^{\text{int}} \subseteq \Omega$  and  $B_{r,p}^{\text{ext}} \subseteq \mathbb{R}^{n+1} \setminus \Omega$  and*

$$B_{r,p}^{\text{int}} \cap M = B_{r,p}^{\text{ext}} \cap M = \{p\}. \quad (38)$$

Note that the above condition is stronger than simply requiring that the curvature of  $M$  is bounded from above by  $\frac{1}{r}$ . It is some kind of “uniform embeddedness” condition.

Now let  $R(\Gamma_0)$  be the maximal radius for which  $\Gamma_0$  satisfies the ball property. Then we require that there is a small constant  $\rho$  such that,

$$(\delta |\log \delta|)^{\frac{1}{4}} \leq \rho \leq \frac{R(\Gamma_0)}{2}. \quad (39)$$

The stability analysis to be carried out in Section 4 which includes the regularity statement and the Ball Lemma will imply that for  $h \leq \delta$  and  $k = \mathcal{O}(\frac{1}{h})$ , we have that

$$\|\mathcal{B}_{kh}\|(h |\log h|)^{\frac{1}{4}} \leq 1$$

and the  $\Gamma_{kh}$ 's will all be embedded manifolds.

The two conditions (37) and (39) will be assumed for the rest of the paper.

## 2.7 Notations and Conventions

Throughout the estimates in the paper, constants and bounded functions will be grouped together as  $C$  and  $\mathcal{O}(1)$  respectively. These are terms bounded by constants that do not depend on  $\|\mathcal{B}_0\|$  or  $h$ . From one line to the next, the  $C$  and  $\mathcal{O}(1)$  terms may change, but we will still use  $C$  and  $\mathcal{O}(1)$ . We will also use the following conventions:

(i)  $A = \mathcal{O}(1)$  or  $A(\cdot) = \mathcal{O}(1)f(\cdot)$ : there is some constant  $C$  such that

$$|A| \leq C \quad \text{or} \quad |A(\cdot)| \leq C|f(\cdot)|.$$

(ii)  $A \lesssim B$  or  $A \gtrsim B$ : there is some constant  $C$  such that

$$A \leq CB \quad \text{or} \quad A \geq CB$$

(iii)  $A \ll 1$  or  $A \gg 1$ :  $A$  is a sufficiently small or large constant.

(iv)  $A \approx B$ : the following limiting behavior holds:

$$\lim \frac{A}{B} = 1.$$

The meaning of the limit will be specified or clear from the context.

### 3 Ansatz and Its Consistency

In this section, we prove the consistency of the scheme by analyzing the ansatz  $U^0$  and the error term  $U^1 = U - U^0$  defined in (18) and (17).

Recall that the initial manifold  $\Gamma_0$  is a smooth, embedded  $n$ -dimensional manifold in  $\mathbb{R}^{n+1}$ . By the regularity of MMC, there exists a positive number  $\rho = \rho(\Gamma_0) > 0$  such that for any  $0 < t < h \ll 1$ , the signed distance function  $r$  to  $M(\Gamma_0, t)$  is a smooth function in the  $\rho$ -tubular neighborhood  $T_\rho(\Gamma_0)$  of  $\Gamma_0$ ,

$$T_\rho(\Gamma_0) = \{q \in \mathbb{R}^{n+1} : \text{dist}(q, \Gamma_0) \leq \rho\}. \quad (40)$$

Hence inside  $T_\rho(\Gamma_0)$ , the  $U^0$  and  $U^1$  are smooth functions of  $(x, t)$ .

Inside  $T_\rho(\Gamma_0)$ , we have the following representation formula for  $U^1$  in terms of the heat kernel  $G$  on  $\mathbb{R}^{n+1} \times \mathbb{R}_+$  (c.f. [24]):

$$\begin{aligned} U^1(x, t) &= \int_0^t \int_{T_\rho(\Gamma_0)} G(x - y, t - \tau) (\partial_t - \Delta) U^0(y, \tau) dy d\tau \\ &\quad + \int_0^t \int_{\partial T_\rho(\Gamma_0)} G(x - y, t - \tau) \frac{\partial U^1}{\partial \nu}(y, \tau) - U^1(y, \tau) \frac{\partial G}{\partial \nu}(x - y, t - \tau) dS d\tau. \end{aligned} \quad (41)$$

where  $\nu$  is the unit outward normal to  $\partial T_\rho(\Gamma_0)$ .

We give some remarks about the above ansatz.

(i) Recall that  $\rho$  is some small constant satisfying the condition (39). Hence the signed distance function to  $M(\Gamma_0, t)$ ,  $r(\cdot, t)$ , is smooth in the set  $\{r(\cdot, t) < 2\rho\}$  (for  $0 \leq t \leq h \leq \delta$ ). It will be a consequence of Lemma 8 that the same  $\rho$  will work for multiple iterations, i.e. a tubular neighborhood of radius  $\rho$  may be placed about  $\Gamma_{kh}$  with the ansatz constructed in the same manner (see Theorem 4).

(ii) The second term in (41), integration on the boundary  $\partial T_\rho(\Gamma_0)$ , produces exponentially small terms. Specifically, the following estimates hold for  $U$  and  $U^0$  on  $\partial T_\rho(\Gamma_0)$ ,

$$|U - U^0| \lesssim e^{-\frac{\rho^2}{16h}}, \quad |\partial_r U^0| \lesssim \frac{1}{\sqrt{h}} e^{-\frac{\rho^2}{16h}}, \quad \text{and} \quad |\nabla U| \lesssim \frac{1}{\sqrt{h}} e^{-\frac{\rho^2}{8h}}.$$

Furthermore, the area of  $\partial T_\rho(\Gamma_0)$  is roughly two of that of  $\Gamma_0$ . Hence we will simply omit it in our analysis.

For the following, given any constant  $c$ , we define  $\{r = c\}$  to be the manifold which is at a distance  $c$  to  $M_t$ :

$$\{r = c\} = \{y \in \mathbb{R}^{n+1} : \text{sdist}(y, M(\Gamma_0, t)) = c\}. \quad (42)$$

We note that for  $c \ll 1$ ,  $\{r = c\}$  is a smooth manifold. Let also  $H_{\{r=c\}} : \{r = c\} \rightarrow \mathbb{R}$  be the mean curvature function of  $\{r = c\}$ .

**Lemma 1.** *For any  $(x, t) \in T_\rho(\Gamma_0) \times [0, h]$ , the following formula holds,*

$$(\partial_t - \Delta)U^0(x, t) = \left( H_{\{r=r(x,t)\}}(x) - H_{\{r=0\}}(\Pi(x)) \right) \partial_r U^0(x, t), \quad (43)$$

where  $\Pi(x)$  is the point in  $M(\Gamma_0, t)$  closest to  $x$ . In addition, we have

$$H_{\{r=r(x,t)\}}(x) - H_{\{r=0\}}(\Pi(x)) = r(x, t) \psi \left( r(x, t), \kappa_1(\Pi(x), t), \kappa_2(\Pi(x), t), \dots, \kappa_n(\Pi(x), t)) \right) \quad (44)$$

where the  $\kappa_i$  are the principal curvatures of  $M(\Gamma_0, t)$  evaluated at  $\Pi(x)$ , and

$$\psi(r, \kappa_1, \kappa_2, \dots, \kappa_n) = \sum_{i=1}^n \frac{\kappa_i^2}{1 - r\kappa_i}. \quad (45)$$

See Figure 3 for an illustration of some of the notations appearing above.

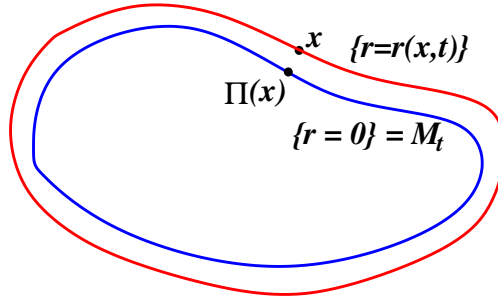


Figure 3:  $\{r = r(x, t)\} = \{y \in \mathbb{R}^{n+1} : \text{sdist}(y, M(\Gamma_0, t)) = r(x, t)\}$ .  $\Pi(x)$  is the projection of  $x$  onto  $M_t$

The above will be proved in **Section A.1**. We will frequently abuse the notation by simply writing  $\psi(x, t)$  in place of  $\psi(r(x, t), \kappa_1(\Pi(x), t), \kappa_2(\Pi(x), t), \dots, \kappa_n(\Pi(x), t))$ . By the formula (20) for  $\partial_r U^0$ , the expression  $(\partial_t - \Delta)U^0(x, t)$  takes the following form

$$\psi(r, \kappa) \frac{r(x, t)}{\sqrt{\pi t}} \exp \left( -\frac{r^2(x, t)}{4t} \right). \quad (46)$$

We also note the following estimates:

$$\|\psi\| \lesssim \|A_{M(\Gamma_0, t)}\|^2 \text{ for } |r| \ll \frac{1}{\|A_{M(\Gamma_0, t)}\|}, \quad \text{and} \quad \left| \frac{r}{\sqrt{t}} \exp \left( -\frac{r^2}{4t} \right) \right| = \mathcal{O}(1) \text{ for all } r, t. \quad (47)$$

With the above, we now proceed to prove the following consistency statement for the MBO scheme.

**Theorem 2.** *For any  $h \leq \delta$ , there is a constant  $C$  depending only on the spatial dimension such that the following estimate holds,*

$$d_{\mathcal{H}}(\Gamma_h, M(\Gamma_0, h)) \lesssim \|\mathcal{B}_0\|^2 h^{\frac{3}{2}} \quad (48)$$

where we have used  $\mathcal{B}_0$  to denote  $A_{\Gamma_0}$ .

*Proof.* By (46) and (47), we have for  $0 < t < h$  that

$$|(\partial_t - \Delta)U^0(x, t)| = \left| \psi(r, \kappa) \frac{r(x, t)}{\sqrt{\pi t}} \exp\left(-\frac{r^2(x, t)}{4t}\right) \right| \lesssim \|A_{M(\Gamma_0, t)}\|^2.$$

Substituting this estimate into the integral in (41) and by the  $L^\infty$ - $L^\infty$  estimate of the heat operator, we obtain that,

$$\begin{aligned} |U^1(x, h)| &\leq \left| \int_0^h \int_{T_\rho(\Gamma_0)} G(x-y, h-\tau) (\partial_t - \Delta)U^0(y, t) dy d\tau \right| \\ &\leq \int_0^h \|G(\cdot, h-\tau)\|_{L^1(\mathbb{R}^{n+1})} \|(\partial_t - \Delta)U^0(\cdot, \tau)\|_{L^\infty(T_\rho(\Gamma_0))} d\tau \\ &\lesssim \sup_{\tau \in [0, h]} \|A_{M(\Gamma_0, \tau)}\|^2 h. \end{aligned}$$

By Lemma 2 (which appears in the next section), we have

$$\sup_{t \in [0, h]} \|A_{M(\Gamma_0, t)}\| \leq (\|\mathcal{B}_0\| + C\|\mathcal{B}_0\|^3 h) = \|\mathcal{B}_0\| (1 + C\|\mathcal{B}_0\|^2 h) \lesssim \|\mathcal{B}_0\|$$

leading to

$$\|U^1(x, h)\|_{L^\infty(T_\rho(\Gamma_0))} \lesssim \|\mathcal{B}_0\|^2 h. \quad (49)$$

Finally, by the representation formula (18) of  $U^0$ , for  $r \ll 1$ , we have

$$U^0(x, t) = \frac{2}{\sqrt{\pi}} \int_0^{\frac{r}{2\sqrt{t}}} \exp(-y^2) dy \approx \frac{r}{\sqrt{\pi t}}. \quad (50)$$

Hence  $U^0 + U^1$  can be zero only if  $r$  satisfies

$$|r(\cdot, h)| \lesssim \|\mathcal{B}_0\|^2 h^{\frac{3}{2}}.$$

□

## 4 Stability

The stability estimates of this section will allow us to extend the consistency estimate from the previous section to multiple iterations in the MBO algorithm. The first, key step is show that the curvature of  $\Gamma_h$  can be controlled by the curvature of  $\Gamma_0$ . Then we prove a geometric result preventing  $\Gamma_h$  to have self-intersection and hence  $\Gamma_h$  remains embedded. Lastly, tying this together with a discrete non-linear Gronwall inequality, we are able to show that the curvatures of the  $\Gamma_{kh}$ 's are uniformly bounded over multiple iterations. The crucial computation and analysis rely on the regularity property of MMC.

We now state the two main results in this section. (We recall the notations about differential geometry from **Section 2.5**.)

**Theorem 3. Stability over one time step.** *There is a constant  $C$  depending on the spatial dimension such that for  $h \leq \delta$ , we have*

$$\|\mathcal{B}_h\| \leq \|\mathcal{B}_0\| (1 + C\|\mathcal{B}_0\|^2 h). \quad (51)$$

**Theorem 4. Stability over multiple time steps.** *Let  $\rho$  be some small number. For any constant*

$$C_0 \in \left( \|\mathcal{B}_0\|, \min \left\{ \frac{1}{(\delta |\log \delta|)^{\frac{1}{4}}}, \frac{1}{2\rho} - 1 \right\} \right], \quad (52)$$

*there exists a time  $T = T(\Gamma_0, C_0)$  independent of  $h$ , such that for  $0 \leq k \leq \lfloor \frac{T}{h} \rfloor$ , the following two statements hold.*

- $\Gamma_{kh}$  is an embedded manifold. More specifically, there is a uniform radius  $r_0 = r_0(\Gamma_0, C_0)$  so that  $\Gamma_{kh}$  satisfies the Ball Property (see Definition 3) with radius  $r_0$ .
- $\|\mathcal{B}_{kh}\| \leq C_0$ .

(Note that for  $\rho$  small enough, the interval for the choice of  $C_0$  is non-empty which by (39) can in fact be further simplified to  $C_0 \in \left( \|\mathcal{B}_0\|, \frac{1}{2\rho} - 1 \right]$ .)

Some remarks are in place.

(i) The specific manner in which  $r_0$  and  $T$  depend on  $\Gamma_0$ , and  $C_0$  will be apparent in the proof of Theorem 4. We note also that the established curvature bound can increase with each iteration, so that the larger the choice of  $C_0$ , the larger we may choose the convergence time  $T$ .

(ii) Theorem 4 essentially proves that  $\sup_k \|\Gamma_{kh}\|_{C^2} < \infty$  for appropriate range of  $k$ . This only implies that  $\Gamma_{kh}$  converges in  $C^{1,\alpha}$  in space. This is not sufficient to show that the convergence time  $T$  for our numerical scheme coincides with the maximum time  $T^*$  for the existence of classical solution. This is because the constant  $C$  in the discrete one-time-step-stability estimate (51) might not be optimal and can be different from the continuous time case. Note however that the scaling in the growth rate of the curvature is the same in both the discrete and continuous cases as illustrated by the MMC of a circle. Finite time blow-up in the estimate over multiple time steps will definitely occur, just as in the continuous case. But the two blow-up times might not be the same. We believe that with more refined analysis, we can in fact have  $\sup_k \|\Gamma_{kh}\|_{C^{2,\alpha}} < \infty$  so that the  $\Gamma_{kh}$  converges in  $C^{2,\alpha'}$  for  $\alpha' < \alpha$ , i.e. the curvature  $\mathcal{B}_{kh}$  also converges (in  $C^{\alpha'}$ ). Then such a discrepancy between  $T$  and  $T^*$  can be removed.

As mentioned before, the proof of Theorem 3 requires some regularity results of MMC. Specifically, we need the following two lemmas for surfaces following MMC. For the next two results, we use  $M_t$  to denote  $M(M_0, t)$ , the solution of MMC starting from  $M_0$ .

**Lemma 2. Bound on Curvature Growth of MMC.** *There is a constant  $C$  which depends only on the spatial dimension such that for  $0 \leq t \leq h$ , we have*

$$\|A_{M_t}\| \leq \|A_{M_0}\|(1 + C\|A_{M_0}\|^2 t).$$

**Lemma 3. Regularity of Higher Derivatives of MMC.** *Let  $\mathbf{N}$  be the normal vector of  $M_t$ . Suppose*

$$\|A_{M_t}\| \leq c_0 \text{ for } t \in [0, h]. \quad (53)$$

*Then for  $t \leq h$ , there is a constant  $C$  depending on the spatial dimension such that,*

$$\|\nabla_{M_t} A_{M_t}\| \leq \frac{C \cdot c_0}{\sqrt{t}}, \quad (54)$$

$$\text{and } \|\partial_t \mathbf{N}\|_{L^\infty(M_t)} \leq \frac{C \cdot c_0}{\sqrt{t}}. \quad (55)$$

*In the above,  $\|\nabla_{M_t} A_{M_t}\| = \sup_{M_t} \sqrt{|\nabla_{M_t} A_{M_t}|^2}$  where  $|\nabla_{M_t} A_{M_t}|^2$  is the squared norm of the tensor  $(\nabla_k A_{M_t})$ . (See [10, Appendix A] for detail explanation of the notations.)*

Note that assumption (53) holds by Lemma 2 together with assumption (37). The above Lemmas will be proved in **Sections A.3** and **A.4**.

The stability analysis makes use of the Implicit Function Theorem (IFT). From the formula (18) and (41) for  $U^0$  and  $U^1$ , it can be seen that,

$$|\partial_r U^0(x, t)| = \left| \frac{1}{\sqrt{\pi t}} \exp\left(-\frac{r^2(x, t)}{4t}\right) \right| \gtrsim \frac{1}{\sqrt{t}} \quad \text{for } |r(x, t)| \lesssim \sqrt{t},$$

and

$$\|\nabla U^1(\cdot, t)\|_{L^\infty(T_\rho(\Gamma_0))} \lesssim \|\mathcal{B}_0\|^2 \sqrt{t}.$$

We can then conclude that,

$$|\partial_r U(x, t)| > 0 \text{ when } |r(x, t)| \lesssim \sqrt{t}.$$

Hence, we can apply the IFT to locally write  $\Gamma_h$  as the graph of a function  $g$  over the tangent plane to  $M(\Gamma_0, h)$  at some reference point (see Fig. 2).

Next we give some preparations for the proof of Theorem 3.

#### 4.1 Further Notations and Preparations

For simplicity, we use  $M_t$  to denote  $M(\Gamma_0, t)$  for  $0 < t \leq h$ . The Weingarten maps of  $M_t$  and  $\Gamma_h$  are denoted by  $A_t$  and  $\mathcal{B}_h$ , respectively. We further set  $\mathcal{B}_0 = A_0$ .

The main goal of the next few sections is to estimate the derivatives of  $U$  at any point  $p \in \Gamma_h$ . By the consistency statement, we can assume that  $p \in T_\rho(\Gamma_0)$ . To simplify the computation, we introduce the following coordinate system. Let the origin be the point  $p_0 = \Pi_{M_h}(p) \in M_h$  which is closest to  $p$ . Then we write

$$\mathbb{R}^{n+1} \cong \mathbb{R}^n \times \mathbb{R} \cong \mathbb{T}_{M_h}(p_0) \times \mathbb{L}(\mathbf{N}(p_0)) \quad (56)$$

where  $\mathbb{T}_{M_h}(p_0)$  and  $\mathbb{L}(\mathbf{N}(p_0))$  are the tangent plane and the line spanned by the outward normal vector  $\mathbf{N}(p_0)$ , both to  $M_h$  at  $p_0$ . We use  $\bar{x} \in \mathbb{R}^n$  and  $\bar{z} \in \mathbb{R}$  to denote the coordinate variables of  $\mathbb{T}_{M_h}(p_0)$  and  $\mathbb{L}(\mathbf{N}(p_0))$ . Then we have the following representations,

$$x = (\bar{x}, \bar{z}) \in \mathbb{R}^{n+1}, \quad p_0 = (\bar{x} = 0, \bar{z} = 0), \quad (57)$$

and near  $p_0$ ,

$$\Gamma_h = \{(\bar{x}, g(\bar{x})) : \bar{x} \in \mathbb{T}_{M_h}(p_0)\}, \quad \text{in particular, } p = (0, g(0)) \in \mathbb{L}(\mathbf{N}(p_0)). \quad (58)$$

(See Figure 2.) With the above notations, we write

$$U = U(\bar{x}, \bar{z}, t), \quad U^0 = U^0(\bar{x}, \bar{z}, t), \quad U^1 = U^1(\bar{x}, \bar{z}, t), \quad \text{and } r = r(\bar{x}, \bar{z}, t).$$

We further note the following statements,

$$\nabla_{\bar{x}} \Big|_{p_0} = \nabla_{M_h} \Big|_{p_0}, \quad (59)$$

and on  $\mathbb{L}(\mathbf{N}(p_0))$  and in particular at  $p$ ,

$$\nabla_{\bar{x}} r = 0, \quad \partial_{\bar{z}} = \partial_r, \quad \partial_{\bar{z}} \nabla_{\bar{x}} r = 0, \quad \partial_{\bar{z}} r = -1, \quad \partial_{\bar{z}}^2 r = 0. \quad (60)$$

We now proceed to prove Theorem 3 by estimating the derivatives of  $U$  at  $p$ . First, by (33),  $\mathcal{B}_h$  can be expressed in terms of  $g$  via the following formula,

$$\mathcal{B}_h = \left( I - \frac{\nabla_{\bar{x}} g \otimes \nabla_{\bar{x}} g}{\sqrt{1 + |\nabla_{\bar{x}} g|^2}} \right) \frac{\nabla_{\bar{x}}^2 g}{\sqrt{1 + |\nabla_{\bar{x}} g|^2}}. \quad (61)$$

To analyze the above, we need the following expressions which are obtained by differentiating  $g$  using the implicit function (recall that  $U(\bar{x}, g(\bar{x}), h) = 0$ ):

$$\nabla_{\bar{x}} g(\bar{x}) = -\frac{1}{\partial_{\bar{z}} U(\bar{x}, g(\bar{x}), h)} \nabla_{\bar{x}} U(\bar{x}, g(\bar{x}), h), \quad (62)$$

$$\begin{aligned} \nabla_{\bar{x}}^2 g(\bar{x}) &= -\frac{1}{(\partial_{\bar{z}} U)} (\nabla_{\bar{x}}^2 U + \partial_{\bar{z}} \nabla_{\bar{x}} U \otimes \nabla_{\bar{x}} g) \\ &\quad + \frac{1}{(\partial_{\bar{z}} U)^2} (\partial_{\bar{z}} \nabla_{\bar{x}} U \otimes \nabla_{\bar{x}} U + \partial_{\bar{z}\bar{z}} U \nabla_{\bar{x}} g \otimes \nabla_{\bar{x}} U) \Big|_{(\bar{x}, g(\bar{x}), h)}. \end{aligned} \quad (63)$$

Theorem 3 is proved by obtaining precise estimates for  $\nabla_{\bar{x}}g, \nabla_{\bar{x}}^2g$  and hence  $\mathcal{B}_h$ . The central technical part is the analysis of the term  $\nabla_{\bar{x}}^2U^1$ .

We will frequently use the following estimates concerning the Green's function: for some constant  $C$  that depend only on the spatial dimension, it holds that

$$\|G(\cdot, t)\|_{L^1(\mathbb{R}^{n+1})} = 1 \ (\leq C) \text{ and } \|\nabla G(\cdot, t)\|_{L^1(\mathbb{R}^{n+1})} \leq \frac{C}{\sqrt{t}}. \quad (64)$$

## 4.2 Estimates for terms without $\nabla_{\bar{x}}^2U^1$

This section gives several preliminary estimates for various derivatives associated with  $U$ .

**Lemma 4** (Estimates for  $U^0$ ).

$$\nabla_{\bar{x}}U^0(p) = 0, \quad (65)$$

$$\nabla_{\bar{x}}^2U^0(p) = -\frac{1}{\sqrt{\pi h}} \exp\left(-\frac{g^2(0)}{4h}\right) \nabla_{\bar{x}}^2r(p) \quad (66)$$

$$\partial_{\bar{z}}\nabla_{\bar{x}}U^0(p) = 0, \quad (67)$$

$$\partial_{\bar{z}}U^0(p) = -\frac{1}{\sqrt{\pi h}} \exp\left(-\frac{g^2(0)}{4h}\right), \quad (68)$$

$$|\partial_{\bar{z}}^2U^0(p)| \lesssim \|\mathcal{B}_0\|^2. \quad (69)$$

*Proof.* Recall that  $U^0 = \frac{2}{\sqrt{\pi}} \int_0^{\frac{r}{2\sqrt{t}}} \exp(-y^2)dy$  and  $r(x, t) = \text{sdist}(x, M_t)$ . Then we have,

$$\begin{aligned} \nabla_{\bar{x}}U^0 &= \frac{1}{\sqrt{\pi t}} \exp\left(-\frac{r^2}{4t}\right) \nabla_{\bar{x}}r, \\ \nabla_{\bar{x}}^2U^0 &= \frac{1}{\sqrt{\pi t}} \left[ \exp\left(-\frac{r^2}{4t}\right) \nabla_{\bar{x}}^2r + \frac{r}{2t} \exp\left(-\frac{r^2}{4t}\right) \nabla_{\bar{x}}r \otimes \nabla_{\bar{x}}r \right], \\ \partial_{\bar{z}}\nabla_{\bar{x}}U^0 &= \frac{1}{\sqrt{\pi t}} \exp\left(-\frac{r^2}{4t}\right) \partial_{\bar{z}}\nabla_{\bar{x}}r - \frac{1}{\sqrt{\pi t}} \exp\left(-\frac{r^2}{4t}\right) \frac{r}{2t} \partial_{\bar{z}}r \nabla_{\bar{x}}r, \\ \partial_{\bar{z}}U^0 &= \frac{1}{\sqrt{\pi t}} \exp\left(-\frac{r^2}{4t}\right) \partial_{\bar{z}}r, \\ \partial_{\bar{z}}^2U^0 &= \frac{1}{\sqrt{\pi t}} \left[ \exp\left(-\frac{r^2}{4t}\right) \partial_{\bar{z}}^2r - \exp\left(-\frac{r^2}{4t}\right) \left(\frac{r}{2t}\right) (\partial_{\bar{z}}r)^2 \right] \end{aligned}$$

Noting the property (60), and the facts  $r(p) = g(0)$ , and  $\|g\|_{L^\infty} \lesssim \|\mathcal{B}_0\|^2 h^{\frac{3}{2}}$  (Theorem 2), all the claims in the Lemma follow.  $\square$

**Lemma 5** (Estimates for  $U^1$ ).

$$\|\nabla U^1\|_{L^\infty(T_\rho(\Gamma_0))} \lesssim \|\mathcal{B}_0\|^2 \sqrt{h}, \quad (70)$$

$$\|\nabla^2 U^1\|_{L^\infty(T_\rho(\Gamma_0))} \lesssim \|\mathcal{B}_0\|^2. \quad (71)$$

(In the above, the gradient operator  $\nabla$  is defined with respect to the spatial variable  $x \in \mathbb{R}^{n+1}$ .)

*Proof.* The statements are consequences of the  $L^1$ -estimates of the Green's function (64).

Taking the spatial gradient in (41), we get,

$$\nabla U^1(x, h) = \int_0^h \int_{T_\rho(\Gamma_0)} \left( \nabla_y G(x - y, h - \tau) \right) \left( \frac{1}{\sqrt{\pi \tau}} \exp\left(-\frac{r^2(y, \tau)}{4\tau}\right) r(y, \tau) \psi(y, \tau) \right) dy d\tau.$$

Therefore for all  $x \in T_\rho(\Gamma_0)$ ,

$$\begin{aligned} |\nabla U^1(x, h)| &\lesssim \int_0^h \|\nabla G(\cdot, h - \tau)\|_{L^1(\mathbb{R}^{n+1})} \left\| \frac{r(\cdot, \tau)}{\sqrt{\tau}} \exp\left(-\frac{r^2(\cdot, \tau)}{4\tau}\right) \psi(\cdot, \tau) \right\|_{L^\infty} d\tau \\ &\lesssim \int_0^h \frac{\|A_\tau\|^2}{\sqrt{h - \tau}} d\tau = \|\mathcal{B}_0\|^2 \mathcal{O}(\sqrt{h}) \end{aligned}$$

which is (70). (Note that we have used (47).)

Differentiate (41) again to get,

$$\begin{aligned} &|\nabla^2 U^1(x, h)| \\ &\leq \int_0^h \int_{T_\rho(\Gamma_0)} |\nabla_y G(x - y, h - \tau)| \left| \nabla_y \left( \frac{r(y, \tau)}{\sqrt{\pi\tau}} \exp\left(-\frac{r^2(y, \tau)}{4\tau}\right) \psi(y, \tau) \right) \right| dy d\tau. \end{aligned}$$

Note that for  $f(y) = ye^{-\frac{y^2}{4}}$ , we have

$$\nabla \left[ f\left(\frac{r}{\sqrt{\tau}}\right) \psi(y, \tau) \right] = f'\left(\frac{r}{\tau}\right) \frac{\nabla \tau}{\sqrt{\tau}} \psi + f\left(\frac{r}{\tau}\right) \nabla \psi$$

so that

$$\left| \nabla \left[ f\left(\frac{r}{\sqrt{\tau}}\right) \psi(y, \tau) \right] \right| \lesssim \frac{\|\psi\|_{L^\infty}}{\sqrt{\tau}} + \|\nabla \psi\|_{L^\infty} \lesssim \frac{\|A_\tau\|^2}{\sqrt{\tau}} + \|A_\tau \nabla A_\tau\| \lesssim \frac{\|\mathcal{B}_0\|^2}{\sqrt{\tau}}.$$

Hence

$$\begin{aligned} \|\nabla^2 U^1(\cdot, h)\|_{L^\infty} &\lesssim \|\mathcal{B}_0\|^2 \int_0^h \frac{1}{\sqrt{h - \tau}} \frac{1}{\sqrt{\tau}} d\tau \quad (\text{by Lemma 3 and (64)}) \\ &\lesssim \|\mathcal{B}_0\|^2 \end{aligned}$$

which is (71). □

**Lemma 6** (First derivative estimates for  $U$  and  $g$ ).

$$(\partial_{\bar{z}} U)^{-1}(p) = \left( \frac{1}{\sqrt{\pi h}} \exp\left(-\frac{g^2}{4h}\right) \right)^{-1} \left( 1 + \|\mathcal{B}_0\|^2 \mathcal{O}(h) \right), \quad (72)$$

$$|\nabla_{\bar{x}} g(0)| \lesssim \|\mathcal{B}_0\|^2 h. \quad (73)$$

*Proof.* First note that  $\partial_{\bar{z}} U(p) = \partial_{\bar{z}} U^0 + \partial_{\bar{z}} U^1$ . Then by (68) and (70), we have

$$\begin{aligned} \partial_{\bar{z}} U(p) &= \frac{1}{\sqrt{\pi h}} \exp\left(-\frac{g(0)^2}{4h}\right) + \|\mathcal{B}_0\|^2 \mathcal{O}(\sqrt{h}) \\ &= \frac{1}{\sqrt{\pi h}} \exp\left(-\frac{g(0)^2}{4h}\right) \left( 1 + \sqrt{h} \exp\left(\frac{g(0)^2}{4h}\right) \|\mathcal{B}_0\|^2 \mathcal{O}(\sqrt{h}) \right) \\ &= \frac{1}{\sqrt{\pi h}} \exp\left(-\frac{g(0)^2}{4h}\right) \left( 1 + \|\mathcal{B}_0\|^2 \mathcal{O}(h) \right) \end{aligned}$$

leading to (72). Note that we have used the estimate  $|g(0)| \lesssim \mathcal{O}(h^{\frac{3}{2}})$  from Theorem 2.

For (73), the statement follows from (62), (65), (72), and (70):

$$\begin{aligned} |\nabla_x g| &= \left| \frac{\nabla_{\bar{x}} U^0 + \nabla_{\bar{x}} U^1}{\partial_{\bar{z}} U} \right| = \left| \|\mathcal{B}_0\|^2 \mathcal{O}(\sqrt{h}) \left( \frac{1}{\sqrt{h}} \exp\left(-\frac{g(0)^2}{4h}\right) \right)^{-1} (1 + \|\mathcal{B}_0\|^2 \mathcal{O}(h)) \right| \\ &= \left| \|\mathcal{B}_0\|^2 \mathcal{O}(h) \exp\left(\frac{g(0)^2}{4h}\right) (1 + \|\mathcal{B}_0\|^2 h) \right| \\ &= \|\mathcal{B}_0\|^2 \mathcal{O}(h). \end{aligned}$$

□



### 4.3 Refined estimates for $\nabla_{\bar{x}}^2 U^1$

We now begin the most substantial computation in estimating  $A_h$ , which is directly related to  $\nabla_{\bar{x}}^2 U^1$ . The following is the key result of this section.

**Lemma 7.** *The following estimate holds,*

$$\|\nabla_{\bar{x}}^2 U^1\| \lesssim \|\mathcal{B}_0\|^3 \sqrt{h}. \quad (74)$$

Before proceeding with the proof of Lemma 7, we show how Lemma 7 along with the estimates of Section 4.2 can lead to Theorem 3. First, by substituting the estimates from Lemma 5 into (63), we have

$$\begin{aligned} & \nabla_{\bar{x}}^2 g(\bar{x}) \Big|_{(\bar{x}, g(\bar{x}), h)} \\ &= -\frac{1}{(\partial_{\bar{z}} U)} (\nabla_{\bar{x}}^2 U + \partial_{\bar{z}} \nabla_{\bar{x}} U \otimes \nabla_{\bar{x}} g) + \frac{1}{(\partial_{\bar{z}} U)^2} (\partial_{\bar{z}} \nabla_{\bar{x}} U \otimes \nabla_{\bar{x}} U + (\partial_{\bar{z}}^2 U) \nabla_{\bar{x}} g \otimes \nabla_{\bar{x}} U) \\ &= -\frac{1}{(\partial_{\bar{z}} U)} (\nabla_{\bar{x}}^2 U^0 + \nabla_{\bar{x}}^2 U^1 + \partial_{\bar{z}} \nabla_{\bar{x}} U^1 \otimes \nabla_{\bar{x}} g) \\ & \quad + \frac{1}{(\partial_{\bar{z}} U)^2} (\partial_{\bar{z}} \nabla_{\bar{x}} U^1 \otimes \nabla_{\bar{x}} U^1 + (\partial_{\bar{z}}^2 U^0 + \partial_{\bar{z}}^2 U^1) \nabla_{\bar{x}} g \otimes \nabla_{\bar{x}} U) \\ &= -\frac{1}{(\partial_{\bar{z}} U)} \left[ -\frac{1}{\sqrt{\pi h}} \exp\left(-\frac{g(0)^2}{4h}\right) \nabla_{\bar{x}}^2 r(p) + \nabla_{\bar{x}}^2 U^1 + \|\mathcal{B}_0\|^2 \|\mathcal{B}_0\|^2 h \right] \\ & \quad + \frac{1}{(\partial_{\bar{z}} U)^2} \left( \|\mathcal{B}_0\|^2 \|\mathcal{B}_0\|^2 \sqrt{h} + \|\mathcal{B}_0\|^2 \|\mathcal{B}_0\|^4 \mathcal{O}(h^{\frac{3}{2}}) \right) \\ &= -\left[ \frac{1}{\sqrt{\pi h}} \exp\left(-\frac{g(0)^2}{4h}\right) \right]^{-1} \left[ -\frac{1}{\sqrt{\pi h}} \exp\left(-\frac{g(0)^2}{4h}\right) \nabla_{\bar{x}}^2 r(p) + \nabla_{\bar{x}}^2 U^1 + \|\mathcal{B}_0\|^4 \mathcal{O}(h) \right] \\ & \quad + \left[ \frac{1}{\sqrt{\pi h}} \exp\left(-\frac{g(0)^2}{4h}\right) \right]^{-2} \left[ \|\mathcal{B}_0\|^4 \mathcal{O}(\sqrt{h}) + \|\mathcal{B}_0\|^6 \mathcal{O}(h^{\frac{3}{2}}) \right] \\ &= \nabla_{\bar{x}}^2 r(p) + \nabla_{\bar{x}}^2 U^1 \mathcal{O}(\sqrt{h}) + \|\mathcal{B}_0\|^4 \mathcal{O}(h^{\frac{3}{2}}) \\ &= \nabla_{\bar{x}}^2 r(p) + \|\mathcal{B}_0\|^3 \mathcal{O}(h) + \|\mathcal{B}_0\|^4 \mathcal{O}(h^{\frac{3}{2}}). \end{aligned}$$

Now apply the above to (61) which relates  $\mathcal{B}_h$  to  $\nabla_{\bar{x}}^2 g$ , we get

$$\begin{aligned} |\mathcal{B}_h| &= \frac{|\text{tr}(\nabla_{\bar{x}}^2 g)|}{\sqrt{1 + |\nabla_{\bar{x}} g|^2}} \left| I - \frac{\nabla_{\bar{x}} g \otimes \nabla_{\bar{x}} g}{\sqrt{1 + |\nabla_{\bar{x}} g|^2}} \right| \\ &\lesssim \left( |\Delta_{\bar{x}} r(p)| + \|\mathcal{B}_0\|^3 \mathcal{O}(h) + \|\mathcal{B}_0\|^4 \mathcal{O}(h^{\frac{3}{2}}) \right) \left( 1 + C \|\mathcal{B}_0\|^2 \mathcal{O}(h^2) \right) \\ &\lesssim \left( |H_{\{r=r(p)\}}(p)| + \|\mathcal{B}_0\|^3 \mathcal{O}(h) + \|\mathcal{B}_0\|^4 \mathcal{O}(h^{\frac{3}{2}}) \right) \left( 1 + C \|\mathcal{B}_0\|^2 \mathcal{O}(h^2) \right) \\ &\lesssim \left( \|A_h\| + \|\mathcal{B}_0\|^3 \mathcal{O}(h) + \|\mathcal{B}_0\|^4 \mathcal{O}(h^{\frac{3}{2}}) \right) \left( 1 + C \|\mathcal{B}_0\|^2 \mathcal{O}(h^2) \right) \\ &\lesssim \left( \|\mathcal{B}_0\| (1 + C \|\mathcal{B}_0\|^2 \mathcal{O}(h)) + \|\mathcal{B}_0\|^3 \mathcal{O}(h) + \|\mathcal{B}_0\|^4 \mathcal{O}(h^{\frac{3}{2}}) \right) \left( 1 + C \|\mathcal{B}_0\|^2 \mathcal{O}(h^2) \right) \\ &\lesssim \|\mathcal{B}_0\| \left( 1 + C \|\mathcal{B}_0\|^2 \mathcal{O}(h) \right) \end{aligned}$$

which is the statement of Theorem 3. In the above,  $\Delta_{\bar{x}} r(p) = H_{\{r=r(p)\}}(p)$  is the mean curvature of the manifold  $\{r = r(p, h)\}$  at  $p$ . It is estimated in terms of  $\|A_h\|$  by using (116):

$$|H_{\{r=r(p, h)\}}(p)| = \left| \sum_i \frac{\kappa_i(\Pi(p))}{1 - r(p, h)\kappa_i(\Pi(p))} \right| \leq \|A_h\| \left( 1 + \|A_h\| \mathcal{O}(h^{\frac{3}{2}}) \right)$$

and then in terms of  $\|\mathcal{B}_0\|$  by using Lemma 2.

The proof of **Lemma 7** is divided into several sub-sections.

### 4.3.1 Preparation for Estimating $\nabla_{\bar{x}}^2 U^1(p, h)$ , $p \in \Gamma_h$

Evaluating  $U^1(x, t)$  at  $x = p = (0, g(0))$  and  $t = h$ , we get,

$$U^1(p, h) = - \int_0^h \int_{T_\rho(\Gamma_0)} G(p - y, h - \tau) \psi(y, \tau) \frac{r(y, \tau)}{\sqrt{\pi\tau}} \exp\left(-\frac{r^2(y, \tau)}{4\tau}\right) dy d\tau.$$

For the integration in the  $y$  variable, we use the coordination system (56) (see also Figure 2):

$$\mathbb{R}^{n+1} \cong \mathbb{R}^n \times \mathbb{R} \cong \{(y_1, y_2, \dots, y_n) \in \mathbb{R}^n\} \times \{y_{n+1} \in \mathbb{R}\}. \quad (75)$$

Taking the second partial derivatives in tangential directions  $\bar{x}_i$  and  $\bar{x}_j$  for  $1 \leq i, j \leq n$  leads to

$$\begin{aligned} & \partial_{\bar{x}_i \bar{x}_j}^2 U^1(p, h) \\ &= - \int_0^h \int_{T_\rho(\Gamma_0)} \left( \partial_{y_j} G(p - y, h - \tau) \right) \partial_{y_i} \left[ \psi(y, \tau) \frac{r(y, \tau)}{\sqrt{\tau}} \exp\left(-\frac{r^2(y, \tau)}{4\tau}\right) \right] dy d\tau. \end{aligned} \quad (76)$$

In addition, for  $|x - y| \geq \sqrt{2(n+5)(h-\tau)|\log(h-\tau)|}$ , we have

$$|\nabla G(x - y, h - \tau)| \leq \frac{1}{2(4\pi)^{\frac{n+1}{2}}} (h - \tau)^{\frac{3}{2}} \sqrt{(h - \tau)|\log(h - \tau)|} = o(h).$$

Hence we can restrict our analysis of (76) to the following region

$$\left\{ y \in \mathbb{R}^{n+1} : |p - y| \leq \sqrt{2(n+5)h|\log h|} \right\}. \quad (77)$$

Now we compute:

$$\begin{aligned} & \partial_{y_i} \left[ \psi(y, \tau) \frac{r(y, \tau)}{\sqrt{\pi\tau}} \exp\left(-\frac{r^2(y, \tau)}{4\tau}\right) \right] \\ &= \psi(y, \tau) \left( 1 - \frac{r^2(y, \tau)}{2\tau} \right) \exp\left(-\frac{r^2(y, \tau)}{4\tau}\right) \frac{\partial_{y_j} r}{\sqrt{\pi\tau}} \end{aligned} \quad (78)$$

$$+ \frac{r(y, \tau)}{\sqrt{\pi\tau}} \exp\left(-\frac{r^2(y, \tau)}{4\tau}\right) \partial_{y_i} \psi(y, \tau) \quad (79)$$

Then we decompose (76) into the following two terms:

$$\partial_{x_i x_j}^2 U^1(0, g(0)) = -I - J$$

where

$$I = - \int_0^h \int_{T_\rho(\Gamma_0)} \partial_{y_j} \left( G(p - y, h - \tau) \right) \left( \text{expression (78)} \right) dy d\tau, \quad (80)$$

$$J = - \int_0^h \int_{T_\rho(\Gamma_0)} \partial_{y_j} \left( G(p - y, h - \tau) \right) \left( \text{expression (79)} \right) dy d\tau. \quad (81)$$

Observe that  $I$  does not involve any derivatives of the curvature term, while  $J$  does.

The main estimates we will arrive at are:

$$|I| \lesssim \|\mathcal{B}_0\|^3 \sqrt{h} \quad \text{and} \quad |J| \lesssim \|\mathcal{B}_0\|^3 \sqrt{h}$$

which together will give the result of **Lemma 7**.

### 4.3.2 Analysis of $I$ (80), term without derivatives of curvature

Note that

$$\partial_{y_i} r(y, \tau) = \langle \mathbf{N}(M_\tau, \Pi_{M_\tau} y), \mathbf{E}_i \rangle \quad (82)$$

where

$$\begin{aligned} \Pi_M(y) &= \text{closest point on } M \text{ to } y, \\ \mathbf{N}(M, q) &= \text{(outward) normal vector to } M \text{ at } q \in M. \\ \mathbf{E}_i &= \text{the } i\text{-th coordinate vector of } \mathbb{T}_{M_h}(p_0). \end{aligned}$$

We will also use the following abbreviation,

$$\Pi_\tau(y) = \Pi_{M_\tau}(y), \quad \mathbf{N}_\tau(y) = \mathbf{N}(M_\tau, \Pi_{M_\tau}(y)).$$

The main observation is that  $\mathbf{N}_\tau(y) \approx \mathbf{N}_h(p)$  which is orthogonal to  $\mathbf{E}_i$  for  $i = 1, 2, \dots, n$ . This is made precise as follows. First extend the normal vector to  $M_\tau$  as a vector field over  $T_\rho(\Gamma_0)$  by taking it to be constant in directions normal to  $M_\tau$ . Let  $\nabla$  be the gradient operator on  $\mathbb{R}^{n+1}$ . Then

$$\begin{aligned} \mathbf{N}_\tau(y) &= \mathbf{N}_h(p) + \int_0^1 \nabla \mathbf{N}_\tau(y + \lambda(p - y))(p - y) d\lambda - \int_\tau^h \frac{d}{ds} \mathbf{N}_s(y) ds \\ &= \mathbf{N}_h(p) + \mathcal{O}(1) \|\mathcal{B}_0\| \left( (h - \tau)^{\frac{1}{2}} + |p - y| \right) \quad (\text{by Lemma 3}) \end{aligned}$$

which leads to

$$\left| \langle \mathbf{N}_\tau(y), \mathbf{E}_i \rangle \right| \lesssim \|\mathcal{B}_0\| \left( (h - \tau)^{\frac{1}{2}} + |p - y| \right). \quad (83)$$

Substituting the above into  $I$  (80), we get

$$\begin{aligned} |I| &= \left| \int_0^h \int_{T_\rho(\Gamma_0)} \left( \partial_{y_j} G(p - y, h - \tau) \right) \psi(y, t) \times \right. \\ &\quad \left. \times \left( 1 - \frac{r^2(y, \tau)}{2\tau} \right) \exp \left( -\frac{r^2(y, \tau)}{4\tau} \right) \frac{\langle \mathbf{N}_\tau, \mathbf{E}_i \rangle}{\sqrt{\tau}} dy d\tau \right| \\ &\lesssim \left| \int_0^h \int_{T_\rho(\Gamma_0)} |\nabla_y G(p - y, h - \tau)| \|\psi\|_{L^\infty} \frac{\|\mathcal{B}_0\| \left( (h - \tau)^{\frac{1}{2}} + |p - y| \right)}{\sqrt{\tau}} dy d\tau \right| \\ &\lesssim \|\mathcal{B}_0\|^3 \left| \int_0^h \int_{T_\rho(\Gamma_0)} |\nabla_y G(p - y, h - \tau)| \frac{\left( (h - \tau)^{\frac{1}{2}} + |p - y| \right)}{\sqrt{\tau}} dy d\tau \right|. \end{aligned}$$

Now for

$$\begin{aligned} &\left| \int_0^h \int_{T_\rho(\Gamma_0)} |\nabla_y G(p - y, h - \tau)| \frac{(h - \tau)^{\frac{1}{2}}}{\sqrt{\tau}} dy d\tau \right| \\ &\lesssim \left| \int_0^h \|\nabla G(\cdot, h - \tau)\|_{L^1(\mathbb{R}^{n+1})} \frac{(h - \tau)^{\frac{1}{2}}}{\sqrt{\tau}} dy d\tau \right| \lesssim \int_0^h \frac{1}{\sqrt{\tau}} d\tau = \mathcal{O}(\sqrt{h}), \end{aligned}$$

while for

$$\begin{aligned} &\left| \int_0^h \int_{T_\rho(\Gamma_0)} |\nabla_y G(p - y, h - \tau)| \frac{|p - y|}{\sqrt{\tau}} dy d\tau \right| \\ &\lesssim \left| \int_0^h \int_{\mathbb{R}^{n+1}} \frac{1}{(h - \tau)^{\frac{n+1}{2}}} \exp \left( -\frac{|y|^2}{4(h - \tau)} \right) \frac{|y|}{h - \tau} \frac{|y|}{\sqrt{\tau}} dy d\tau \right| \\ &\lesssim \left| \int_0^h \int_{\mathbb{R}^{n+1}} \exp \left( -\frac{|y|^2}{4} \right) |y|^2 \frac{1}{\sqrt{\tau}} dy d\tau \right| \lesssim \left| \int_0^h \frac{1}{\sqrt{\tau}} d\tau \right| = \mathcal{O}(\sqrt{h}) \end{aligned}$$

which combined together gives the estimate for  $I$  stated in (80).

### 4.3.3 Analysis of $J$ (81), term involving derivatives of curvature

Recall the formula (45) for the function  $\psi$  where the curvatures are evaluated at  $\Pi_\tau(y)$ . Then we have

$$\begin{aligned}
\partial_{y_j} \psi &= \sum_l^n \partial_{y_j} \left( \frac{\kappa_l^2}{1 - r\kappa_l} \right) \\
&= 2 \sum_{l=1}^n \frac{\kappa_l \partial_{y_i} \kappa_l}{1 - r\kappa_l} + \sum_{l=1}^n \frac{\kappa_l^2}{(1 - r\kappa_l)^2} (r \partial_{y_i} \kappa_l + \kappa_l \partial_{y_j} r) \\
&= 2 \sum_{l=1}^n \kappa_l \partial_{y_i} \kappa_l + \sum_{l=1}^n \left[ r \left( \frac{2\kappa_l^2 \partial_{y_i} \kappa_l}{1 - r\kappa_l} + \kappa_l^2 \frac{\partial_{y_i} \kappa_l}{(1 - r\kappa_l)^2} \right) + \frac{\kappa_l^3 \langle \mathbf{N}_\tau(y), \mathbf{E}_i \rangle}{(1 - r\kappa_l)^2} \right] \\
&= 2 \sum_{l=1}^n \kappa_l \partial_{y_j} \kappa_l + \mathcal{O}(1) \|\mathcal{B}_0\|^3 \left[ \frac{r}{\sqrt{t}} + \|\mathcal{B}_0\| (|x_0 - y| + |h - t|^{\frac{1}{2}}) \right],
\end{aligned}$$

where (83) and Lemma 3 have been used. Substituting this back into (81), we obtain,

$$\begin{aligned}
-J &= 2 \int_0^h \int_{T_\rho(\Gamma_0)} \partial_{y_j} G(p - y, h - \tau) \left[ \frac{r(y, \tau)}{\sqrt{\tau}} \exp\left(-\frac{r^2(y, \tau)}{4\tau}\right) \sum_{l=1}^n \kappa_l \partial_{y_i} \kappa_l \right] dy d\tau \quad (84) \\
&\quad + \int_0^h \int_{T_\rho(\Gamma_0)} \partial_{y_j} G(p - y, h - \tau) \left[ \frac{r(y, \tau)}{\sqrt{\tau}} \exp\left(-\frac{r^2(y, \tau)}{4\tau}\right) \times \right. \\
&\quad \quad \left. \times \mathcal{O}(1) \|\mathcal{B}_0\|^3 \left( \frac{r(y, \tau)}{\sqrt{\tau}} + \|\mathcal{B}_0\| (|p - y| + |h - \tau|^{\frac{1}{2}}) \right) \right] dy d\tau \quad (85) \\
&:= J_1 + J_2.
\end{aligned}$$

Note that both  $J_1$  and  $J_2$  contains derivatives of  $\kappa_l$ 's. However,  $J_2$  is easier to deal with as it contains the pre-factor  $r$  which makes the integrand small. Hence it is analyzed first.

**Analysis of  $J_2$  (85).** Notice that,

$$\begin{aligned}
&\frac{r(y, \tau)}{\sqrt{\tau}} \exp\left(-\frac{r^2(y, \tau)}{4\tau}\right) \|\mathcal{B}_0\|^3 \left[ \frac{r(y, \tau)}{\sqrt{\tau}} + \|\mathcal{B}_0\| (|p - y| + |h - \tau|^{\frac{1}{2}}) \right] \\
&\lesssim \frac{r^2}{\tau} \exp\left(-\frac{r^2}{4\tau}\right) \|\mathcal{B}_0\|^3 + \|\mathcal{B}_0\|^4 \frac{r}{\sqrt{\tau}} \exp\left(-\frac{r^2}{\sqrt{4\tau}}\right) (|p - y| + |h - \tau|^{\frac{1}{2}}) \\
&\lesssim \|\mathcal{B}_0\|^3 + \|\mathcal{B}_0\|^4 (|p - y| + |h - \tau|^{\frac{1}{2}}).
\end{aligned}$$

Hence

$$\begin{aligned}
|J_2| &\lesssim \int_0^h \int_{\mathbb{R}^{n+1}} |\nabla G(\cdot, h - \tau)| \left[ \|\mathcal{B}_0\|^3 + \|\mathcal{B}_0\|^4 (|p - y| + |h - \tau|^{\frac{1}{2}}) \right] \\
&\lesssim \|\mathcal{B}_0\|^3 \sqrt{h} + \|\mathcal{B}_0\|^4 h \lesssim \|\mathcal{B}_0\|^3 \sqrt{h}.
\end{aligned}$$

**Analysis of  $J_1$  (84).** First note that the following simple bound for  $J_1$

$$\begin{aligned}
|J_1| &\lesssim \left| \int_0^h \int_{T_\rho(\Gamma_0)} \partial_{y_j} G(p - y, h - \tau) \left[ \frac{r(y, \tau)}{\sqrt{\tau}} \exp\left(-\frac{r^2(y, \tau)}{4\tau}\right) \sum_{l=1}^n \kappa_l \partial_{y_i} \kappa_l \right] dy d\tau \right| \\
&\lesssim \left| \int_0^h \|\nabla G(\cdot, h - \tau)\|_{L^1(\mathbb{R}^{n+1})} \frac{\|\mathcal{B}_0\|^2}{\sqrt{\tau}} dy d\tau \right| \\
&\lesssim \|\mathcal{B}_0\|^2 \int_0^h \frac{1}{\sqrt{h - \tau}} \frac{1}{\sqrt{\tau}} d\tau \lesssim \|\mathcal{B}_0\|^2
\end{aligned}$$

is too crude for our purpose. The following more refined computation takes into account the different scalings of the integrand along the tangential and normal directions to  $M_\tau$ .

For this purpose, we will use the co-area formula to perform the integration in  $J_1$  over  $\{r = r'\}$  for  $-\rho \leq r' \leq \rho$ :

$$J_1 = 2 \int_0^h \int_{-\rho}^\rho \int_{\{r=r'\}} \left( \partial_{y_j} G(p-y, h-\tau) \right) \left( \frac{r'(y, \tau)}{\sqrt{\tau}} \exp\left(-\frac{r'^2(y, \tau)}{4\tau}\right) \right) \left( \sum_{l=1}^n \kappa_l \partial_{x_i} \kappa_l \right)_{\Pi_\tau(y)} \times \\ \times dA_{r'}(y) dr' d\tau.$$

In the above, we recall the sign distance function  $r$  (16) to  $M_\tau$  and the notation (42)  $\{r = r'\} = \{y \in \mathbb{R}^{n+1} : r(y, \tau) = r'\}$ . Furthermore,  $dA_{r'}$  is the volume form of  $\{r = r'\}$ . Then we parametrize the collection of manifolds  $\{r = r'\} : -\rho \leq r' \leq \rho$  using  $M_\tau$  as follows:

$$\tilde{s} : M_\tau \times (-\rho, \rho) \longrightarrow \tilde{s}(s, r') := s + r' \mathbf{N}(s, \tau) \quad (86)$$

where  $\mathbf{N}(s, \tau)$  is the unit normal to  $M_\tau$  at  $s$ . Note that  $s = \Pi_\tau(\tilde{s})$ . Then

$$J_1 = 2 \int_0^h \int_{-\rho}^\rho \int_{M_\tau} \partial_{y_j} \left( G(p - \tilde{s}, h - \tau) \right) \left( \sum_{l=1}^n \kappa_l(s, \tau) \partial_{y_i} \kappa_l(s, \tau) \right) \frac{r'}{\sqrt{\tau}} \exp\left(-\frac{r'^2}{4\tau}\right) \times \\ \times \left| \frac{dA_{r'}(\tilde{s})}{dA_0(s)} \right| dA_0(s) dr' d\tau$$

where  $dA_0$  is the volume form of  $M_\tau$  and  $\left| \frac{dA_{r'}(\tilde{s})}{dA_0(s)} \right|$  is Jacobian for the change of variable from  $\tilde{s} \in \{r = r'\}$  to  $s \in M_\tau$ .

Next we decompose

$$J_1 = J_{11} + J_{12}$$

where

$$J_{11} = 2 \int_0^h \int_{-\rho}^\rho \int_{M_\tau} \partial_{y_j} \left( G(p - \tilde{s}, h - \tau) \right) \left( \sum_{l=1}^n \kappa_l(s, \tau) \partial_{y_i} \kappa_l(s, \tau) \right) \frac{r'}{\sqrt{\tau}} \exp\left(-\frac{r'^2}{4\tau}\right) \times \\ \times dA_0(s) dr' d\tau, \quad (87)$$

and

$$J_{12} = 2 \int_0^h \int_{-\rho}^\rho \int_{M_\tau} \partial_{y_j} \left( G(p - \tilde{s}, h - \tau) \right) \left( \sum_{l=1}^n \kappa_l(s, \tau) \partial_{y_i} \kappa_l(s, \tau) \right) \frac{r'}{\sqrt{\tau}} \exp\left(-\frac{r'^2}{4\tau}\right) \times \\ \times \left[ \left| \frac{dA_{r'}(\tilde{s})}{dA_0(s)} \right| - 1 \right] dA_0(s) dr' d\tau. \quad (88)$$

In a sense,  $J_{12}$  is the error term due to the deviation of the Jacobian from 1. Hence  $J_{11}$  is the dominating term. The remainder of this subsection is to show that both  $|J_{11}|$  and  $|J_{12}|$  can be bounded from above by  $C \|\mathcal{B}_0\| \sqrt{h}$ .

To proceed, we decompose  $\nabla G$  into two exponential kernels, corresponding to integrating in the tangential and normal directions to  $M_\tau$ . For this purpose, we introduce the operator  $\Sigma_{s, \tau}(p)$  which gives the orthogonal projection of  $p$  onto  $\text{Span}\{\mathbf{N}(s, \tau)\}$ , the normal line at  $s \in M_\tau$ . (See Figure 4.)

Then for any  $\tilde{s} = s + r' \mathbf{N}(s, \tau) \in \{r = r'\}$ , we write,

$$p - \tilde{s} = [p - \Sigma_{s, \tau}(p)] + [\Sigma_{s, \tau}(p) - \tilde{s}]. \quad (89)$$

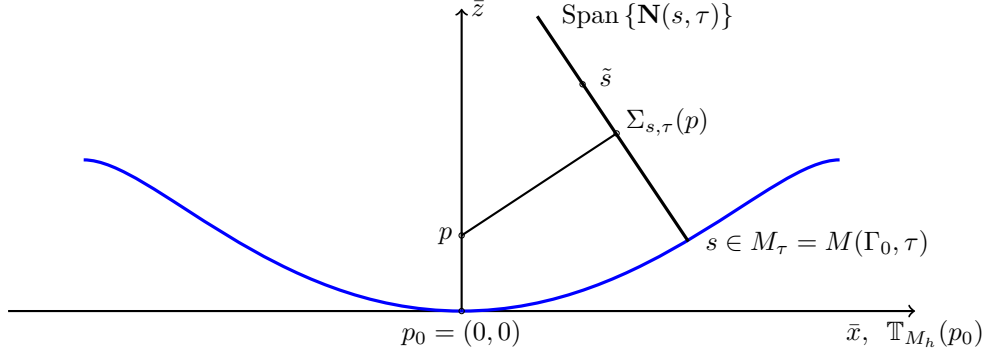


Figure 4: Illustration of  $\Sigma_{s,\tau}$ . Note that for any  $\tilde{s} \in T_p(\Gamma_0)$ , it can be decomposed as  $\tilde{s} = s + (\tilde{s} - s)$  where  $s = \Pi_\tau(\tilde{s}) \in M_\tau$  and  $\tilde{s} - s \in \text{Span}(\mathbf{N}(s, \tau))$ .

Note that

$$\partial_{y_j} G(p - y, h - \tau) = \frac{1}{(h - \tau)^{\frac{n+1}{2}}} \exp\left(-\frac{|p - y|^2}{4(h - \tau)}\right) \frac{\langle p - y, \mathbf{E}_j \rangle}{h - \tau}.$$

Hence

$$\begin{aligned} & \partial_{y_j} G(p - \tilde{s}, h - \tau) \\ &= \left[ \frac{\langle p - \Sigma_{s,\tau}(p), \mathbf{E}_j \rangle}{(h - \tau)^{\frac{n+1}{2}+1}} + \frac{\langle \Sigma_{s,\tau}(p) - \tilde{s}, \mathbf{E}_j \rangle}{(h - \tau)^{\frac{n+1}{2}+1}} \right] \exp\left(-\frac{|p - \Sigma_{s,\tau}(p)|^2}{4(h - \tau)}\right) \exp\left(-\frac{|\Sigma_{s,\tau}(p) - \tilde{s}|^2}{4(h - \tau)}\right). \end{aligned}$$

Substituting the above into the expression for  $J_{11}$ , we have

$$J_{11} = J_{111} + J_{112}$$

where

$$\begin{aligned} J_{111} &= \int_0^h \int_{M_\tau} \frac{\langle p - \Sigma_{s,\tau}(p), \mathbf{E}_j \rangle}{(h - \tau)^{\frac{n}{2}+1}} \exp\left(-\frac{|p - \Sigma_{s,\tau}(p)|^2}{4(h - \tau)}\right) \left( \sum_{l=1}^n \kappa_l \partial_{y_i} \kappa_l \right)_{(s,\tau)} \times \\ &\quad \times \left[ \frac{1}{(h - \tau)^{\frac{1}{2}}} \int_{r'} \exp\left(-\frac{|\Sigma_{s,\tau}(p) - \tilde{s}|^2}{4(h - \tau)}\right) \frac{r'}{\sqrt{\tau}} \exp\left(-\frac{r'^2}{4\tau}\right) dr' \right] dA_0(s) d\tau, \\ J_{112} &= \int_0^h \int_{M_\tau} \frac{1}{(h - \tau)^{\frac{n}{2}}} \exp\left(-\frac{|p - \Sigma_{s,t}(p)|^2}{4(h - \tau)}\right) \left( \sum_{l=1}^n \kappa_l \partial_{x_i} \kappa_l \right)_{(s,\tau)} \times \\ &\quad \times \left[ \frac{1}{(h - \tau)^{\frac{1}{2}}} \int_{r'} \frac{\langle \Sigma_{s,t}(p) - \tilde{s}, \mathbf{E}_j \rangle}{(h - \tau)} \exp\left(-\frac{|\Sigma_{s,t}(p) - \tilde{s}|^2}{4(h - \tau)}\right) \frac{r'}{\sqrt{\tau}} \exp\left(-\frac{r'^2}{4\tau}\right) dr' \right] dA_0(s) d\tau. \end{aligned}$$

which essentially correspond to integrations along the tangential and normal directions.

**Analysis of  $J_{111}$ .** We first evaluate the inner integration with respect to  $r'$ . From **Section A.2**, we have that,

$$\begin{aligned} & \frac{1}{(h - \tau)^{\frac{1}{2}}} \int_{-\infty}^{\infty} \exp\left(-\frac{|\Sigma_{s,\tau}(p) - \tilde{s}|^2}{4(h - \tau)}\right) \frac{r'}{\sqrt{\tau}} \exp\left(-\frac{r'^2}{4\tau}\right) dr' \\ &= 2\sqrt{\pi} \frac{\tau}{h^{\frac{3}{2}}} r(\Sigma_{s,t}(p), \tau) \exp\left(-\frac{r^2(\Sigma_{s,\tau}(p), \tau)}{4h}\right). \quad (90) \end{aligned}$$

Substituting this into  $J_{111}$ , we get

$$J_{111} = \frac{\sqrt{\pi}}{(h)^{\frac{3}{2}}} \int_0^h \tau \left[ \int_{M_\tau} \frac{\langle p - \Sigma_{s,\tau}(p), \mathbf{E}_j \rangle}{(h - \tau)^{\frac{n}{2}+1}} \exp\left(-\frac{|p - \Sigma_{s,\tau}(p)|^2}{4(h - \tau)}\right) \left( \sum_{l=1}^n \kappa_l \partial_{y_i} \kappa_l \right)_{(s,\tau)} \times \right. \\ \left. \times r(\Sigma_{s,\tau}(p), \tau) \exp\left(-\frac{r^2(\Sigma_{s,\tau}(p), \tau)}{4h}\right) dA_0(s) \right] d\tau. \quad (91)$$

Heuristically, the inner integration on  $M_\tau$  is essentially a convolution with the  $n$ -dim Green's function integrated on the tangent plane  $\mathbb{T}_{M_\tau}(p_0)$  to  $M_\tau$  at  $p_0 = \Pi_{M_\tau}(p)$ . Recall the coordinate system (75). Note also that by (76), we may restrict our attention to  $|y| \leq \sqrt{2(n+5)h} \log h \ll 1$ . Now we change the integration variable from  $s$  to  $y$ . The following estimates can be established,

$$\exp\left(-\frac{|p - \Sigma_{s,\tau}(p)|^2}{4(h - \tau)}\right) \lesssim \exp\left(-\frac{C|y|^2}{h - \tau}\right) \quad (92)$$

$$|p - \Sigma_{s,\tau}(p)| \lesssim |y|, \quad (93)$$

$$dA_0 \lesssim d^n y. \quad (94)$$

We further note the following two statements:

1. Let  $f$  be the graph of  $M_\tau$  over  $\mathbb{T}_{M_\tau}(p_0)$ . Then

$$\begin{aligned} |r(\Sigma_{s,\tau}(p), \tau)| &\lesssim |r(p, \tau)| + |\langle y, \nabla f \rangle| \lesssim |r(p, \tau)| + |y| |\nabla f| \\ &\lesssim \|\mathcal{B}_0\|^2 h^{\frac{3}{2}} + \|\mathcal{B}_0\|^2 h |y| \\ &\lesssim \|\mathcal{B}_0\|^2 h^{\frac{3}{2}} + \|\mathcal{B}_0\|^2 (h^2 + |y|^2). \end{aligned}$$

2. By the regularity estimate Lemma 3 for MMC, we have

$$|\nabla_{y_i} \kappa_l| \lesssim \frac{\|\mathcal{B}_\tau\|}{\sqrt{\tau}} \lesssim \frac{\|\mathcal{B}_0\|}{\sqrt{\tau}}. \quad (95)$$

Now substituting (92) - (95) into (91), we obtain,

$$|J_{111}| \lesssim \frac{1}{h^{\frac{3}{2}}} \int_0^h \tau \int_y \frac{|y|}{(h - \tau)^{\frac{n}{2}+1}} \exp\left(-\frac{3|y|^2}{4(h - \tau)}\right) \frac{\|\mathcal{B}_0\|^2}{\sqrt{\tau}} \times \\ \times \left\{ \|\mathcal{B}_0\|^2 (h)^{\frac{3}{2}} + \|\mathcal{B}_0\|^2 (h^2 + |y|^2) \right\} d^n y d\tau.$$

Using the fact that  $\int_{\mathbb{R}^n} \frac{|y|^K}{(h - \tau)^{n/2+1}} \exp\left(-\frac{3|y|^2}{4(h - \tau)}\right) d^n y \lesssim (h - \tau)^{\frac{K}{2}-1}$ , we have

$$\begin{aligned} &|J_{111}| \\ &\lesssim \frac{\|\mathcal{B}_0\|^4}{h^{\frac{3}{2}}} \int_0^h \frac{\sqrt{\tau} h^{\frac{3}{2}}}{\sqrt{h - \tau}} d\tau + \frac{\|\mathcal{B}_0\|^4}{h^{\frac{3}{2}}} \int_0^h \frac{\sqrt{\tau} h^2}{\sqrt{h - \tau}} d\tau + \frac{\|\mathcal{B}_0\|^4}{h^{\frac{3}{2}}} \int_0^h \sqrt{\tau} \sqrt{h - \tau} d\tau \\ &\lesssim \|\mathcal{B}_0\|^4 h + \|\mathcal{B}_0\|^4 h^{\frac{3}{2}} + \|\mathcal{B}_0\|^4 \sqrt{h} \\ &\lesssim \|\mathcal{B}_0\|^3 \sqrt{h}. \end{aligned}$$

**Analysis of  $J_{112}$ .** This is very similar to that for  $J_{111}$ , but with slightly different terms. The inner integral in  $r'$  is given by (see **Section A.2**):

$$\begin{aligned} &\frac{1}{\sqrt{h - t}} \int_{-\infty}^{\infty} \frac{\langle \Sigma_{s,\tau}(p) - \tilde{s}, \mathbf{E}_j \rangle}{(h - \tau)} \exp\left(-\frac{|\Sigma_{s,\tau}(p) - \tilde{s}|^2}{4(h - \tau)}\right) \frac{r'}{\sqrt{\tau}} \exp\left(-\frac{r'^2}{4\tau}\right) dr' \\ &= 2\sqrt{\pi} \langle \mathbf{N}(s, \tau), \mathbf{E}_j \rangle \exp\left(-\frac{r^2(\Sigma_{s,\tau}(p), \tau)}{4h}\right) \left( \frac{\tau}{h^{\frac{5}{2}}} r^2(\Sigma_{s,\tau}(p), \tau) - \frac{2\tau}{h^{\frac{3}{2}}} \right). \quad (96) \end{aligned}$$

As  $\frac{a^2}{b} \exp(-\frac{a^2}{b})$  is a bounded function of  $a$  and  $b$ , the above can be estimated as

$$\mathcal{O}(1) \frac{\tau}{(h)^{\frac{3}{2}}} \langle \mathbf{N}(s, \tau), \mathbf{E}_j \rangle.$$

Hence

$$|J_{112}| \lesssim \frac{\|\mathcal{B}_0\|^2}{(h)^{\frac{3}{2}}} \int_0^h \sqrt{\tau} \int \frac{1}{(h-\tau)^{\frac{n}{2}}} \exp\left(-\frac{|y|^2}{4(h-\tau)}\right) |\langle \mathbf{N}(s, \tau), \mathbf{E}_j \rangle| d^n y d\tau,$$

Similar to (83), we can infer that

$$\langle \mathbf{N}(s, \tau), \mathbf{E}_j \rangle = \mathcal{O}(1) \|\mathcal{B}_0\| \left( \sqrt{h} + |s| \right) = \mathcal{O}(1) \|\mathcal{B}_0\| \left( \sqrt{h} + |y| \right) \quad (\text{note } s = (y, h(y))).$$

Substituting this into the integral and using the  $L^1 - L^\infty$  estimate for the Green's function, we obtain,

$$\begin{aligned} |J_{112}| &\lesssim \frac{\|\mathcal{B}_0\|^3}{(h)^{\frac{3}{2}}} \int_0^h \sqrt{\tau} \int \frac{1}{(h-\tau)^{\frac{n}{2}}} \exp\left(-\frac{C|y|^2}{h-\tau}\right) (\sqrt{h} + |y|) d^n y d\tau \\ &\lesssim \frac{\|\mathcal{B}_0\|^3}{(h)^{\frac{3}{2}}} \int_0^h \sqrt{\tau} \left( \sqrt{h} + \sqrt{h-\tau} \right) d\tau \lesssim \|\mathcal{B}_0\|^3 \sqrt{h} \end{aligned}$$

completing the analysis for  $J_{11}$ .

**Analysis of  $J_{12}$ .** Note that  $\left| \frac{dA_{r'}(\bar{s})}{dA_0(s)} - 1 \right| \lesssim \|\mathcal{B}_0\| |r'|$ . Hence  $J_{12}$  is of smaller order. Using similar analysis as for  $J_{11}$ , we can then conclude that

$$|J_{12}| \lesssim \|\mathcal{B}_0\|^3 \sqrt{h}.$$

The above concludes the proof of Theorem 3.

**Remark 1.** Note that in the proof of Theorem 3, we establish point-wise estimates for  $\nabla^2 U^1$  of the form  $\|\mathcal{B}_0\|^3 \mathcal{O}(\sqrt{h})$ . Applying the same point-wise analysis to  $U^1$  and  $\nabla U^1$ , we could have established that  $U^1 = \|\mathcal{B}_0\|^3 \mathcal{O}((h)^{\frac{3}{2}})$  and  $\|\nabla U^1\| = \|\mathcal{B}_0\|^3 \mathcal{O}(h)$ . This will be improvements over the estimates (49) and (70) which follow from standard  $L^\infty$ -estimates given by convolving  $(\partial_t - \Delta)U^0$  with the Green's function  $G$ .

#### 4.4 Stability over successive iterations

In this section, we prove that the algorithm can be iterated over  $\lfloor \frac{T}{h} \rfloor$  steps, over which the numerical manifold stays embedded and has a uniform curvature bound. There are two tools we use toward this. The first is a ‘‘Ball Lemma’’ which can ensure the embeddedness of the numerically manifolds  $\Gamma_{nh}$ 's. The second is a discrete Gronwall-type inequality as given in [8]. We state the latter first in the following.

**Theorem 5.** ([8, Theorem 2.1]) Let  $D$  be a constant. Suppose  $\{x_k\}_{k \geq 0}$  is a sequence of numbers satisfying for all  $k \geq 1$  that

$$|x_k| \leq |x_{k-1}| (1 + D|x_{k-1}|^2 h).$$

Let further  $\Phi$  be the following monotone (and hence invertible) function,

$$\Phi(x) = \int_1^x \frac{ds}{s + Ds^3}.$$

Then for all  $k \geq 0$ , we have,

$$|x_k| \leq \Phi^{-1}\left(\Phi(|x_0|) + kh\right).$$



Next is the statement of the Ball Lemma.

**Lemma 8** (Ball Lemma). *(For one step.)* Suppose we have a constant  $C_0$  such that  $\|\mathcal{B}_0\|, \|\mathcal{B}_h\| \leq C_0$ . Then  $\Gamma_h$  satisfies the Ball Property (Definition 3) with radius  $r$  given by,

$$r = \min \left( \frac{1}{C_0 + 1}, m_0 - C(n, C_0)h \right), \quad (97)$$

where

$$m_0 = \min \left\{ |s_1 - s_2| : s_1, s_2 \in \Gamma_0 \subset \mathbb{R}^{n+1}, d_0(s_1, s_2) \geq \frac{C_*}{C_0} \right\}, \quad (98)$$

$d_0$  is the distance between points on  $\Gamma_0$  defined in (123),  $C_*$  is the universal constant in (36) and  $C(n, C_0)$  is some constant depending only on the spatial dimension  $n$  and the curvature bound  $C_0$ .

*(For multiple steps.)* Suppose we have iterated the algorithm  $N$  times and that  $\|\mathcal{B}_{kh}\| \leq C_0$  for  $0 \leq k \leq N$  for some  $N$ . Then  $\Gamma_{Nh}$  will satisfy the ball property with radius

$$r = \min \left( \frac{1}{C_0 + 1}, m_0 - C(n, C_0)Nh \right). \quad (99)$$

The proof is given in **Section A.5**.

**Remark 2.** *It is evident in the proof of the above lemma that  $M(\Gamma_0, t)$ , for  $0 \leq t \leq h$ , will also satisfy the Ball Property with the same radius  $r$  as in (97).*

We now show how Lemma 8 and Theorem 5 allow for the algorithm to be iterated repeatedly with uniform curvature estimate.

#### 4.4.1 Proof of Theorem 4

Let  $\tilde{T} = \Phi(C_0) - \Phi(\|\mathcal{B}_0\|)$  which is also the maximal  $t$  such that  $\Phi^{-1}(\Phi(\|\mathcal{B}_0\|) + t) \leq C_0$ . Now define

$$T := \min \left\{ \tilde{T}, \frac{m_0 - 2\rho}{C(n, C_0)} \right\} = \min \left\{ \Phi(C_0) - \Phi(\|\mathcal{B}_0\|), \frac{m_0 - 2\rho}{C(n, C_0)} \right\} \quad (100)$$

where  $C(n, C_0)$  is the constant coming from Lemma 8.

From Theorem 3, we have

$$\|\mathcal{B}_h\| \leq \|\mathcal{B}_0\| (1 + C\|\mathcal{B}_0\|^2 h).$$

Applying Theorem 5 we have that,

$$\|\mathcal{B}_h\| \leq \Phi^{-1}(\Phi(\|\mathcal{B}_0\|) + h) \leq C_0.$$

Consequently, the Ball Lemma implies that  $\Gamma_h$  is an embedded manifold, which satisfies the ball property with radius,

$$r = \min \left( \frac{1}{C_0 + 1}, m_0 - C(n, C_0)h \right).$$

Notice that  $\rho \leq \frac{1}{2} \min \left( \frac{1}{C_0 + 1}, m_0 - C(n, C_0)h \right)$ . As a consequence, we may repeat all the preceding analysis with  $\Gamma_0$  and  $T_\rho(\Gamma_0)$  replaced by  $\Gamma_h$  and  $T_\rho(\Gamma_h)$ . Iterating for  $k$  ( $0 \leq k \leq \lfloor \frac{T}{h} \rfloor$ ) steps, we arrive at the result.

**Remark 3.** *The above proof reveals the simultaneous preservation of uniform curvature estimates and the embeddedness of the numerical manifolds  $\Gamma_{kh}$ . The bigger the  $C_0$  and smaller the  $\rho$  are, the larger the maximum convergence time  $T$  can be chosen. The smaller  $\rho$  ensures that the Ball Property will hold for longer time interval. The only constraint is (37) which will impose a smaller maximum time-step size  $\delta$ .*

## 5 Convergence to MMC

In this section we prove that the algorithm converges to MMC. Recall that the convergence is phrased in the weak form using the framework bounded variation (BV-)functions or sets of finite perimeter (12)-(13) [27]. We first explain some key notations. We refer to [18, 19] for more detail exposition about the function space.

Let  $\chi : \mathbb{R}^{n+1} \rightarrow \mathbb{R}$  with  $\|\chi\|_{L^1(\mathbb{R}^{n+1})} < \infty$ . It is called a function with bounded variation, written as  $\chi \in BV(\mathbb{R}^{n+1}, \mathbb{R})$ , or simply  $\chi \in BV(\mathbb{R}^{n+1})$ , if its variational derivative is given by a (Radon) measure. Precisely, there is a finite constant  $C$  such that

$$\sup \left\{ \int \chi \operatorname{div} \varphi : \varphi \in C_0^1(\mathbb{R}^{n+1}, \mathbb{R}^{n+1}), |\varphi| \leq 1 \right\} \leq C. \quad (101)$$

In this case, there is a Borel measure  $\mu$  and a vector valued function  $\nu : \operatorname{supp}(\mu) \rightarrow \mathbb{S}^n$  such that for any  $\varphi \in C_0^1(\mathbb{R}^{n+1}, \mathbb{R}^{n+1})$ ,

$$\int \chi \operatorname{div} \varphi = - \int \varphi \cdot \nu \, d\mu. \quad (102)$$

It is also customary to denote  $\mu$  and  $\nu$  by  $|\nabla \chi|$  and  $\frac{\nabla \chi}{|\nabla \chi|}$ .

In the present paper,  $\chi$  takes its values from  $\{0, 1\}$  so that we write  $\chi \in BV(\mathbb{R}^{n+1}, \{0, 1\})$ . Then the set  $\Omega = \{x : \chi(x) = 1\}$  (so that  $\chi = \mathbb{1}_\Omega$ ) is called a *set of finite perimeter*. We state here two fundamental facts about such sets.

(i) There exists a notion of *reduced boundary*  $\Gamma = \partial^* \Omega$  which is  $\mathcal{H}^n$ -rectifiable such that  $\mu = \mathcal{H}^n \llcorner_\Gamma$  and  $\nu$  is a continuous function on  $\Gamma$   $\mu$  almost everywhere. Then  $|\nabla \chi|$  and  $\frac{\nabla \chi}{|\nabla \chi|}$  are essentially the area measure and outward unit normal vector function of  $\Gamma$ . The area (also commonly known as the *perimeter*) and the mean curvature of  $\Gamma$  are given by

$$\int_\Gamma d\mathcal{H}^n = \int |\nabla \chi| \quad \text{and} \quad H = \operatorname{div} \frac{\nabla \chi}{|\nabla \chi|}. \quad (103)$$

The mean curvature function  $H$  can also be defined via the following definition: for any  $\zeta \in C^1(\mathbb{R}^{n+1}, \mathbb{R}^{n+1})$ ,

$$\int H \zeta \cdot \nu \, |\nabla \chi| = \int (\operatorname{div} \zeta - \nu \cdot \nabla \zeta \nu) \, |\nabla \chi|.$$

(ii) The space of sets of finite perimeter (and more generally BV-functions) with uniform bounded perimeters is compact in  $L^1$ : if  $\sup_\alpha \int_{\mathbb{R}^{n+1}} |\nabla \chi_\alpha| + \int_{\mathbb{R}^{n+1}} |\chi_\alpha| < \infty$ , then there is a subsequence  $\alpha_i$  and a  $\chi^* \in BV(\mathbb{R}^{n+1})$  such that

$$\chi_{\alpha_i} \rightarrow \chi^* \quad \text{in } L_{loc}^1(\mathbb{R}^{n+1}).$$

In particular, if  $\chi_\alpha = \mathbb{1}_{\Omega_\alpha}$ , then  $\chi^* = \mathbb{1}_{\Omega^*}$  for some  $\Omega^* \subseteq \mathbb{R}^{n+1}$ . Furthermore, the area measure is lower-semicontinuous, i.e. for all Borel set  $B \subset \mathbb{R}^{n+1}$ , it holds that

$$\int_B |\nabla \chi^*| \leq \liminf_i \int_B |\nabla \chi_{\alpha_i}|.$$

In the following, for simplicity, we will simply use the terminology BV-function with the understanding that we are dealing exclusively with sets of finite perimeter.

Now we will make use of the above formulation to prove the convergence of  $\chi_h$  and identify the equation satisfied by its limit. We first recall the definition (9) and (10) of  $\chi_h$ ,  $\Omega_h$  and  $\Gamma_h$ . Note that now  $\chi_h : \mathbb{R}^{n+1} \times [0, T] \rightarrow \{0, 1\}$  is a function of both spatial and temporal variables. For

convenience, we denote  $\chi_h(t) = \chi_h(\cdot, t)$ . Since  $\chi_h(t)$  is a classical solution of MMC for  $t \in (kh, (k+1)h)$ , denoting  $H_h(t)$  to be the mean curvature of  $\Gamma_h(t)$ , we have for any  $\zeta \in C^\infty(\Lambda \times [0, T], \mathbb{R}^n)$ ,  $\zeta = 0$  and  $\xi \in C^\infty(\Lambda \times [0, T], \mathbb{R})$ ,  $\xi = 0$  on  $\partial\Lambda \times [0, T] \cup \Lambda \times \{0\}$  that

$$\int_{kh}^{(k+1)h} \int_{\Lambda} \left( \operatorname{div} \zeta - \frac{\nabla \chi_h(t)}{|\nabla \chi_h(t)|} \nabla \zeta \frac{\nabla \chi_h(t)}{|\nabla \chi_h(t)|} \right) |\nabla \chi_h(t)| = - \int_{kh}^{(k+1)h} \int_{\Lambda} H_h(t) \zeta \cdot \nabla \chi_h(t), \quad (104)$$

and

$$\int_{kh}^{(k+1)h} \int_{\Lambda} \chi_h(t) \partial_t \xi + \int_{\Lambda} \chi_h(kh^+) \xi(kh) - \int_{\Lambda} \chi_h((k+1)h^-) \xi((k+1)h^-) = - \int_{kh}^{(k+1)h} \int_{\Lambda} H_h(t) \xi |\nabla \chi_h(t)|. \quad (105)$$

Summing the above over  $k = 0, 1, \dots, \lfloor \frac{T}{h} \rfloor$ , we obtain:

$$\int_0^T \int_{\Lambda} \left( \operatorname{div} \zeta - \frac{\nabla \chi_h(t)}{|\nabla \chi_h(t)|} \nabla \zeta \frac{\nabla \chi_h(t)}{|\nabla \chi_h(t)|} \right) |\nabla \chi_h(t)| = - \int_0^T \int_{\Lambda} H_h(t) \zeta \cdot \nabla \chi_h(t), \quad (106)$$

and

$$\int_0^T \int_{\Lambda} \chi_h(t) \partial_t \xi + \int_{\Lambda} \chi_h(0) \xi(0) = - \int_0^T \int_{\Lambda} H_h(t) \xi |\nabla \chi_h(t)| + E(h), \quad (107)$$

where

$$E(h) = \sum_{k=1}^{\lfloor \frac{T}{h} \rfloor} \int_{\Lambda} \zeta(kh) \left( \chi_h(kh^+) - \chi_h(kh^-) \right) \quad (108)$$

is the sum of the ‘‘jump’’ errors made between iterations, precisely at the thresholding steps. The above are the discrete analog of (12)-(13).

By the consistency and stability estimates Theorem 2 and 4, as  $h \rightarrow 0$ , we have that

$$|E(h)| \lesssim \sum_{k=1}^{\lfloor \frac{T}{h} \rfloor} h^{\frac{3}{2}} \rightarrow 0.$$

Thus (106)-(107) is ‘‘almost’’ a solution to (12)-(13). The remaining step is to show that  $\chi_h$  and  $H_h$  exhibit appropriate compactness in  $h$ , so that we may pass to the limit  $h \rightarrow 0$  in (106)-(107).

With the consistency and stability estimates, we can already conclude that the sequence of manifolds  $\{\Gamma_{kh} : 0 \leq k \leq \lfloor \frac{T}{h} \rfloor\}$  converges to some limit in the Hausdorff distance  $d_{\mathcal{H}}$  (8). It remains to show that the limit satisfies the equation of MMC. We find the framework of BV-convergence as stated in Definition 2 to be the most convenient.

The outline of proof is as follows. We first show that  $\{\chi_h\}$  is compact in  $L^1(\mathbb{R}^{n+1} \times [0, T], \{0, 1\})$  and hence has a limit  $\chi^*$ . Next we show that the area measure converges in measure:  $|\nabla \chi_h| \rightharpoonup |\nabla \chi^*|$ . This enables us to prove that both the normal vectors  $\frac{\nabla \chi_h}{|\nabla \chi_h|}$  and the mean curvature  $H_h$  are also convergent.

For convenience, we use  $rca(\Lambda \times [0, T])$  to denote the space of regular Radon measures on  $\Lambda \times [0, T]$ .

## 5.1 $L^1$ -Compactness

The main conclusion in this section is that up to subsequence,  $\chi_h$  converges to some  $\chi^*$  in  $L^1(\mathbb{R}^{n+1} \times [0, T], \{0, 1\})$ . Furthermore, for each  $t$ ,  $\chi^*(t)$  is in  $BV(\mathbb{R}^{n+1}, \{0, 1\})$ . This is a consequence of the

Kolmogorov-Riesz-Frechet Theorem [4, Thm. 4.26] together with the compactness property of BV-functions. The following sequence of propositions facilitate the use of this theorem.

**Proposition 1.** *The perimeters of  $\chi_h(t)$  are uniformly bounded, i.e.*

$$\sup_{t \in [0, T]} \int_{\Lambda} |\nabla \chi_h(t)| < +\infty.$$

*Proof.* Using the implicit function theorem,  $\Gamma_h(kh)$  can be parametrized over  $X_h(kh^-)$  via a map,

$$q \in X_h(kh^-) \longrightarrow q + g(q)\mathbf{N}(q),$$

where  $\mathbf{N}(q)$  is the unit normal of  $X_h(kh^-)$  at  $q$ . By the consistency and stability estimates, we have that  $g = \mathcal{O}(h^{\frac{3}{2}})$  and  $\nabla g = \mathcal{O}(h)$ . (Here  $\nabla g$  is the gradient computed over  $\chi_h(kh^-)$ .) Hence,

$$\int |\nabla \chi_h(kh^+)| = \int |\nabla \chi_h(kh^-)|(1 + \mathcal{O}(h)).$$

Furthermore, as the area decreases through MMC, we have

$$\int |\nabla \chi_h(kh^-)| \leq \int |\nabla \chi_h((k-1)h^+)|.$$

By iterating, we obtain,

$$\int |\nabla \chi_h(kh^+)| \leq (1 + Ch) \int |\nabla \chi_h(0)|.$$

□

**Proposition 2.** *For all  $t \in [0, T]$ ,  $h > 0$  and  $w \in \mathbb{S}^n$ , the following spatial continuity statement holds,*

$$\lim_{s \rightarrow 0} \int_{\Lambda} |\chi_h(t, y + sw) - \chi_h(t, y)| dy = 0.$$

*Proof.* This follows from the estimate,

$$\int_{\Lambda} |\chi_h(t, y + sw) - \chi_h(t, y)| dy \lesssim s \int |\nabla \chi_h(t)|$$

and the uniform perimeter bound just proved. □

**Proposition 3.** *The collection  $\{\chi_h(t) : t \in [0, T]\}$  satisfies the following Lipschitz in time estimate,*

$$\int_{\Lambda} |\chi_h(t + \epsilon) - \chi_h(t)| \lesssim \epsilon,$$

whenever  $h \leq \epsilon \leq T - t$ .

*Proof.* We have the following estimates,

$$\int_{\Lambda} |\chi_h(kh^+) - \chi_h(kh^-)| \lesssim h^{\frac{3}{2}}, \quad \text{and} \quad \int_{\Lambda} |\chi_h(kh^-) - \chi_h((k-1)h^+)| \lesssim h.$$

The first follows by the consistency estimate while the second follows from the regularity of  $\Gamma_h(t)$  which solves the MMC for  $t \in [(k-1)h, kh)$ . The result follows by iterating these estimates. □

From the above, the Kolmogorov-Riesz-Frechet Theorem [4, Thm 4.26] implies that there is a subsequence such that  $\chi_h$  converges to  $\chi^*$  in  $L^1(\mathbb{R}^{n+1} \times [0, T], \{0, 1\})$ . By the uniform boundedness of the perimeters of  $\chi_h(t)$ , compactness of sets of finite perimeters implies that  $\chi^*(t) \in BV(\mathbb{R}^{n+1}, \{0, 1\})$  for almost every  $t$ . The Lipschitz continuity in time implies that this holds for every  $t$ . In particular, we have a fixed subsequence  $h_i \rightarrow 0$  such that for all  $t$ ,

$$\chi_{h_i}(t) \rightarrow \chi^*(t) \quad \text{in } L^1(\Lambda). \quad (109)$$

In the following, the notation  $\lim_h$  and  $h \rightarrow 0$  refer to  $h_i \rightarrow 0$ . In several occasions, this subsequence will be further refined. Hence for simplicity, the subscript  $i$  will be omitted.

## 5.2 Convergence of Area

In this section, we will prove that  $|\nabla\chi_h|$  converges weakly to  $|\nabla\chi^*|$  in measure. This is a stronger statement than just  $\nabla\chi_h \rightarrow \nabla\chi^*$ . It implies that the area converges:

$$\int_{\mathbb{R}^{n+1} \times [0, T]} |\nabla\chi_h| \rightarrow \int_{\mathbb{R}^{n+1} \times [0, T]} |\nabla\chi^*|.$$

By [27], this gives a sufficient condition for (106)-(107) to converge to (12)-(13). By the uniform boundedness of the perimeter (Proposition 1) and the Lebesgue Dominated Convergence Theorem, it suffices to prove that for each  $t \in [0, T]$ ,  $|\nabla\chi_h(t)| \rightarrow |\nabla\chi^*(t)|$ .

The first step toward this goal is the observation that the normal vectors  $\nu_h(t)$  to  $\Gamma_h(t)$  converges *strongly*. By the Lemma 8 (Ball Lemma), we may extend  $\nu_h(t)$  to be a smooth function defined on  $\mathbb{R}^{n+1}$ . By the uniform  $C^2$ -bound of the  $\Gamma_h(t)$ , we can invoke the Arzela-Ascoli theorem to pick a further subsequence of  $h$  such that  $\nu_h(t)$  converges to a  $\tilde{\nu}(t)$  in  $C^{1,\alpha}(\mathbb{R}^{n+1}; \mathbb{R}^{n+1})$ .

We are now ready to prove the main theorem of this section.

**Theorem 6.** *For every  $t \in [0, T]$ ,  $|\nabla\chi_h(t)| \rightarrow |\nabla\chi^*(t)|$  in rca( $\Lambda$ ) as  $h \rightarrow 0$ , i.e. for all open set  $B \subseteq \mathbb{R}^{n+1}$  such that  $|\nabla\chi_h(t)|(\partial B) = 0$ , it holds that*

$$\lim_h \int_B |\nabla\chi_h(t)| = \int_B |\nabla\chi^*(t)|. \quad (110)$$

We emphasize that the same sequence  $h \rightarrow 0$  works for every  $t$ .

*Proof.* We fix a  $t \in [0, T]$ . First, by the lower semi-continuity of area under  $L^1$ -convergence, we have,

$$\int_B |\nabla\chi^*(t)| \leq \liminf_{h \rightarrow 0} \int_B |\nabla\chi_h(t)|.$$

Next, let  $L(B)$  and a subsequence  $h_j$  of  $h$  be such that

$$L(B) = \limsup_h \int_B |\nabla\chi_h(t)| = \lim_{h_j} \int_B |\nabla\chi_{h_j}(t)|.$$

Note that the subsequence can depend on  $t$ . But this does not matter as  $t$  is fixed.

Let  $\nu^*(t)$  be the normal vector function of  $\chi^*(t)$ . By the weak convergence of  $\nabla\chi_{h_j}(t)$  to  $\nabla\chi^*$ , it holds that

$$\nu_{(h),j}(t) |\nabla\chi_{(h),j}(t)| = \nabla\chi_{(h),j}(t) \rightarrow \nabla\chi^*(t) = \nu^*(t) |\nabla\chi^*(t)|. \quad (111)$$

We now compute,

$$\begin{aligned}
L(B) &= \lim_{j \rightarrow \infty} \int_B |\nabla \chi_{(h)_j}(t)| \\
&= \lim_{j \rightarrow \infty} \int_B \langle \nu_{(h)_j}(t), \nu_{(h)_j}(t) \rangle |\nabla \chi_{(h)_j}(t)| = \lim_{j \rightarrow \infty} \int_B \langle \nu_{(h)_j}(t), \nabla \chi_{(h)_j}(t) \rangle \\
&= \lim_{j \rightarrow \infty} \int_B \langle \nu_{(h)_j}(t) - \tilde{\nu}(t), \nabla \chi_{h_j} \rangle + \int_B \langle \tilde{\nu}(t), \nabla \chi_{h_j} \rangle \\
&= \int_B \langle \tilde{\nu}(t), \nabla \chi^*(t) \rangle = \int_B \langle \tilde{\nu}(t), \nu^*(t) \rangle |\nabla \chi^*(t)| \\
&\quad (\text{since } \nu_h(t) \longrightarrow \tilde{\nu}(t) \text{ strongly in } C^{1,\alpha} \text{ and } \nabla \chi_{h_j} \rightharpoonup \nabla \chi^*) \\
&\leq \int_B |\nabla \chi^*(t)| \quad (\text{since } |\tilde{\nu}(t)|, |\nu^*(t)| \leq 1).
\end{aligned}$$

Hence

$$\limsup_h \int_B |\nabla \chi_h(t)| = L(B) \leq \int_B |\nabla \chi^*(t)| \leq \liminf_h \int_B |\nabla \chi_h(t)|$$

from which (110) follows.  $\square$

### 5.3 Convergence of $\chi_h$ to (12)-(13)

The procedure now largely follows the ideas of the proof of [27, Theorem 2.3]. Two additional technical but crucial results will be used.

The first is that we can control the normal vectors appropriately so that they can be passed to the limit. For this purpose, consider smooth functions  $\nu_\epsilon$  which approximate  $\nu^*$  in the sense that

$$\lim_{\epsilon \rightarrow 0} \int_0^T \int_\Lambda |\nu^* - \nu_\epsilon|^2 |\nabla \chi^*| dt = 0. \quad (112)$$

Then we claim that

$$\lim_{\epsilon \rightarrow 0} \lim_h \int_0^T \int_\Lambda |\nu_h - \nu_\epsilon|^2 |\nabla \chi_h| dt = 0 \quad (113)$$

which follows from

$$\begin{aligned}
&\lim_{\epsilon \rightarrow 0} \lim_h \int_0^T \int_\Lambda |\nu_h - \nu_\epsilon|^2 |\nabla \chi_h| dt \\
&= \lim_{\epsilon \rightarrow 0} \lim_h \int_0^T \int_\Lambda |\nabla \chi_h| dt + \int_0^T \int_\Lambda |\nu_\epsilon|^2 |\nabla \chi_h| dt - 2 \int_0^T \int_\Lambda \langle \nu_h, \nu_\epsilon \rangle |\nabla \chi_h| dt \\
&= \lim_{\epsilon \rightarrow 0} \int_0^T \int_\Lambda |\nabla \chi^*| dt + \int_0^T \int_\Lambda |\nu_\epsilon|^2 |\nabla \chi^*| dt - 2 \int_0^T \int_\Lambda \langle \nu^*, \nu_\epsilon \rangle |\nabla \chi^*| dt \\
&= \lim_{\epsilon \rightarrow 0} \int_0^T \int_\Lambda |\nu^* - \nu_\epsilon|^2 |\nabla \chi^*| dt \\
&= 0.
\end{aligned}$$

The second is the existence of a limiting mean curvature function. For this, notice that the measures  $H_h |\nabla \chi_h| dt$  are uniformly bounded in  $rca(\Lambda \times [0, T])$ . Therefore, we may pass to a further subsequence such that  $H_h |\nabla \chi_h| dt \rightharpoonup \sigma$ , for some  $\sigma \in rca(\Lambda \times [0, T])$ . We note the following two facts about  $\sigma$ :

(i) Since the sequence of mean curvatures  $H_h$  are uniformly bounded,  $\sigma$  is absolutely continuous with respect to  $|\nabla \chi^*(t)| dt$ . Indeed, let  $B$  be such that  $\int_B |\nabla \chi^*(t)| dt = 0$ . Then

$$\left| \int_B H_h |\nabla \chi_h(t)| dt \right| \lesssim \int_B |\nabla \chi_h(t)| dt \longrightarrow_{h \rightarrow 0} \int_B |\nabla \chi^*(t)| dt = 0.$$

Hence  $\int_B d\sigma = \lim_h \int H_h |\nabla \chi_h| = 0$ .

(ii) Let  $H^*$  to be the Radon-Nikodym derivative of  $\sigma$  with respect to  $|\nabla \chi^*|dt$ . Then  $H^*$  belongs to  $L^2(|\nabla \chi^*|dt)$ . Indeed, let  $\eta$  be some smooth test function. Then,

$$\left| \int_0^T \int_\Lambda H_h \eta |\nabla \chi_h| dt \right| \lesssim \left( \int_0^T \int_\Lambda |\eta|^2 |\nabla \chi_h| dt \right)^{\frac{1}{2}}.$$

Sending  $h$  to 0 we get,

$$\left| \int_0^T \int_\Lambda H^* \eta |\nabla \chi^*| dt \right| \lesssim \left( \int_0^T \int_\Lambda |\eta|^2 |\nabla \chi^*| dt \right)^{\frac{1}{2}}.$$

The conclusion follows as a consequence of the Riesz-representation theorem.

We now show the convergence of (106) to (12) as follows. Consider the right hand side. For any  $\zeta \in C^\infty(\mathbb{R}^{n+1} \times [0, T], \mathbb{R}^{n+1})$ ,

$$\begin{aligned} & \limsup_h \left| \int_0^T \int_\Lambda H_h \zeta \cdot \nabla \chi_h dt - \int_0^T \int_\Lambda H^* \zeta \cdot \nabla \chi^* dt \right| \\ &= \limsup_h \left| \int_0^T \int_\Lambda H_h \zeta \cdot \nu_h |\nabla \chi_h| dt - \int_0^T \int_\Lambda H^* \zeta \cdot \nu^* |\nabla \chi^*| dt \right| \\ &\leq \limsup_h \left| \int_0^T \int_\Lambda H_h \zeta \cdot \nu_\epsilon |\nabla \chi_h| dt - \int_0^T \int_\Lambda H^* \zeta \cdot \nu^* |\nabla \chi^*| dt \right| \\ &\quad + \limsup_h \left| \int_0^T \int_\Lambda H_h \zeta \cdot (\nu_h - \nu_\epsilon) |\nabla \chi_h| dt \right|. \end{aligned}$$

The first integral of the last line becomes

$$\left| \int_0^T \int_\Lambda H^* \zeta \cdot (\nu_\epsilon - \nu^*) |\nabla \chi^*| dt \right|$$

which by (112) tends to zero upon taking  $\epsilon \rightarrow 0$ . The second integral also tends to zero due to (113) and the  $(L^2)$ -boundedness of the  $H_h$ .

For the second term of the left hand side of (106), we similarly have

$$\begin{aligned} & \limsup_h \left| \int_0^T \int_\Lambda \nu_h \nabla \zeta \nu_h |\nabla \chi_h| dt - \int_0^T \int_\Lambda \nu^* \nabla \zeta \nu^* |\nabla \chi^*| dt \right| \\ &\leq \limsup_h \left| \int_0^T \int_\Lambda \nu_\epsilon \nabla \zeta \nu_h |\nabla \chi_h| dt - \int_0^T \int_\Lambda \nu^* \nabla \zeta \nu^* |\nabla \chi^*| dt \right| \\ &\quad + \limsup_h \left| \int_0^T \int_\Lambda (\nu_h - \nu_\epsilon) \nabla \zeta \nu_h |\nabla \chi_h| dt \right| \\ &\lesssim \left| \int_0^T \int_\Lambda (\nu_\epsilon - \nu^*) \nabla \zeta \nu^* |\nabla \chi^*| dt \right| + \limsup_h \left| \int_0^T \int_\Lambda |\nu_h - \nu_\epsilon|^2 |\nabla \chi_h| dt \right| \end{aligned}$$

which tends to zero as  $\epsilon \rightarrow 0$ , again by (112) and (113).

Finally, the convergence of the first integral of the left hand side of (106) and the whole of (107) follow from the weak-convergence of  $|\nabla \chi_h|$  to  $|\nabla \chi^*|$ , the  $L^1$ -convergence of  $\chi_h$  to  $\chi^*$  and the weak-convergence of  $H_h |\nabla \chi_h|$  to  $H^* |\nabla \chi^*|$ , respectively.

The above concludes that (12)-(13) hold as a limiting equation of (106)-(107) (with  $v$  replaced by  $H^*$ ).

# A Appendix

## A.1 Properties of the signed distance function

Here we give some basic properties of the signed distance function and prove **Lemma 1**, in particular formula (43) and (44).

Given an open set  $\Omega \in \mathbb{R}^{n+1}$  with a smooth boundary  $M = \partial\Omega$ , define  $\Pi = \Pi_M$  to be the projection operator which maps  $x \in \mathbb{R}^{n+1}$  to its closest point on  $M$ . By the smoothness assumption of  $M$ , the map  $\Pi$  is well-defined in a tubular neighborhood of  $M$ . In this tubular neighborhood, then the signed distance function  $r$  to  $M$  can be expressed as

$$r(x) = -(x - \Pi(x)) \cdot \mathbf{N}(\Pi(x)), \quad (114)$$

where  $\mathbf{N}(z) = \mathbf{N}_M(z)$  is the outward normal to  $M$  at  $z \in M$ .

Many geometrical aspects of  $M$  can be recovered from the signed distance function: (i) for any  $x \in \mathbb{R}^{n+1}$ ,  $\nabla r(x) = -\mathbf{N}(\Pi(x))$  and hence  $|\nabla r(x)| = 1$ ; (ii) for any  $p \in M$ ,  $-\nabla_M^2 r(p)$  is the Weingarten map of  $M$  at  $p$ ; (iii) the mean curvature  $H_M$  of  $M$  at  $p$  is given by  $-\Delta r(p)$ .

**Proof of (43).** We consider a general function  $f : \mathbb{R}^{n+1} \times \mathbb{R}_+ \rightarrow \mathbb{R}$  in the form

$$f(x, t) = \bar{f}(r(x, t), t)$$

where  $r$  is the signed-distance function to  $M_t = M(\Gamma_0, t)$ . We work in the regime that  $r$  and  $t$  are small enough so that  $r$  is a smooth function of its arguments. Let  $(x, t) \in \mathbb{R}^{n+1} \times \mathbb{R}_+$  be an arbitrary point. Then we compute,

$$\begin{aligned} & \partial_t f(x, t) - \Delta_x f(x, t) \\ &= \partial_t \bar{f}(r, t) + \partial_r \bar{f}(r, t) \partial_t r - \partial_{rr} \bar{f}(r, t) |\nabla_x r|^2 - \partial_r \bar{f}(r, t) \Delta_x r \\ &= \partial_r \bar{f}(r, t) (\partial_t r - \Delta_x r) + \partial_t \bar{f}(r, t) - \partial_{rr} \bar{f}(r, t) \end{aligned}$$

where all the arguments of  $r$  are evaluated at  $(x, t)$ . Note that: (i)  $\partial_t r = -H_{M_t}(\Pi(x))$  where  $H_{M_t}$  is the mean curvature function of  $M_t$  and  $\Pi(x)$  is the projection onto  $M_t$ ; and (ii)  $\Delta_x r = -H_{\{r=r(x,t)\}}(x)$ . Then we have

$$\partial_t f(x, t) - \Delta_x f(x, t) = \partial_r \bar{f}(r, t) (H_{\{r=r(x,t)\}}(x) - H_{M_t}(\Pi(x))) + \partial_t \bar{f}(r, t) - \partial_{rr} \bar{f}(r, t).$$

Finally, by choosing  $\bar{f} = U^0$ , the last two terms of the above vanish altogether as  $U^0$  solves the linear heat equation (19).

**Proof of (44).** The statement will follow from a formula for  $H_{\{r=r_0\}}$ .

Since the time variable  $t$  is irrelevant, we simply write  $M = M_t$ . We first fix a point  $x_0$  and let  $r_0 = r(x_0)$ . Then for  $|r_0|$  small enough, we can express the manifold  $\{r = r_0\}$  as a map over  $M$  via

$$F : M \rightarrow \mathbb{R}^{n+1} : F(y) = y + r_0 \mathbf{N}_M(y).$$

Note that  $x_0 \in \{r = r_0\}$ . Let  $\{\mathbf{E}_1, \dots, \mathbf{E}_n\}$  be an orthonormal basis for the tangent plane  $\mathbb{T}_M$  of  $M$ . Since  $M$  and  $\{r = r_0\}$  share the same normal vector, i.e.  $\mathbf{N}_{\{r=r_0\}}(x) = \mathbf{N}_M(\Pi(x))$  for all  $x \in \{r = r_0\}$ ,  $\mathbf{E}_1, \dots, \mathbf{E}_n$  is also an orthonormal basis for the tangent plane to  $\{r = r_0\}$ , in particular at  $x_0$ . Now let  $y_0 = \Pi_M(x_0)$  and we impose that at  $y_0$ ,  $\{\mathbf{E}_1, \dots, \mathbf{E}_n\}$  corresponds to the principal curvatures  $\{\kappa_1(y_0), \dots, \kappa_n(y_0)\}$  of  $M$  at  $y_0$ .

Next we compute the second fundamental form  $\tilde{A}_j^i$  and the Weingarten map  $\tilde{A}_{ij}$  of  $\{r = r_0\}$  in the following fashion.



- The metric  $\tilde{g}_{ij}$  of  $\{r = r_0\}$  is given in terms of  $F$  as follows,

$$\tilde{g}_{ij} = \langle \partial_i F, \partial_j F \rangle = \langle \mathbf{E}_i(1 - r_0 \kappa_i), \mathbf{E}_j(1 - r_0 \kappa_j) \rangle = \delta_{ij}(1 - r_0 \kappa_i)^2.$$

Its inverse  $\tilde{g}^{ij}$  is then,

$$(1 - r_0 \kappa_i)^{-2} \delta_{ij}.$$

- By definition, we have,

$$\tilde{A}_{ij} = \langle \partial_i N, \partial_j F \rangle = \langle \kappa_i e_i, e_j(1 - r_0 \kappa_j) \rangle = \delta_{ij} \kappa_i(1 - r_0 \kappa_i).$$

Hence the Weingarten map takes the form

$$\tilde{A}_j^i = \tilde{g}^{ij} \tilde{A}_{ij} = \delta_{ij} \frac{\kappa_i}{1 - r_0 \kappa_i}. \quad (115)$$

Upon taking the trace of  $\tilde{A}_j^i$ , the mean curvature of  $\{r = r_0\}$  at  $x_0$  is then given by

$$H_{\{r=r_0\}}(x_0) = \sum_{i=1}^n \frac{\kappa_i}{1 - r_0 \kappa_i} \Big|_{y_0} = H_M(y_0) + r_0 \sum_{i=1}^n \frac{\kappa_i^2}{1 - r_0 \kappa_i} \Big|_{y_0} \quad (116)$$

which is (44).

## A.2 Some Gaussian Integrations: Derivations of (90) and (96)

We first perform two computations regarding Gaussian integrals.

**Lemma 9.** *For  $d_0 > 0$ ,  $0 < t < h$ , we have*

$$\frac{1}{(h-t)^{\frac{1}{2}}} \int_{-\infty}^{\infty} \exp\left(-\frac{|d_0 - r'|^2}{4(h-t)}\right) \frac{r'}{\sqrt{t}} \exp\left(-\frac{r'^2}{4t}\right) dr' = 2\sqrt{\pi} \exp\left(-\frac{d_0^2}{4h}\right) \frac{d_0 t}{h^{\frac{3}{2}}} \quad (117)$$

and

$$\begin{aligned} \frac{1}{(h-t)^{\frac{1}{2}}} \int_{-\infty}^{\infty} \frac{d_0 - r'}{(h-t)} \exp\left(-\frac{|d_0 - r'|^2}{4(h-t)}\right) \frac{r'}{\sqrt{t}} \exp\left(-\frac{r'^2}{4t}\right) dr' \\ = 2\sqrt{\pi} \exp\left(-\frac{d_0^2}{4h}\right) \left(\frac{d_0^2 t}{h^{\frac{5}{2}}} - \frac{2t}{h^{\frac{3}{2}}}\right) \end{aligned} \quad (118)$$

*Proof.* We will note the following identities:

$$\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}, \quad \int_{-\infty}^{\infty} x e^{-x^2} dx = 0, \quad \int_{-\infty}^{\infty} x^2 e^{-x^2} dx = \frac{\sqrt{\pi}}{2},$$

and

$$\frac{(d_0 - r')^2}{4(h-t)} + \frac{r'^2}{4t} = \frac{h}{4t(h-t)} \left(r - \frac{d_0 t}{h}\right)^2 + \frac{d_0^2}{4h}$$

so that

$$\exp\left(-\frac{(d_0 - r')^2}{4(h-t)}\right) \exp\left(-\frac{r'^2}{4t}\right) = \exp\left(-\frac{d_0^2}{4h}\right) \exp\left(-\frac{h}{4t(h-t)} \left(r - \frac{d_0 t}{h}\right)^2\right). \quad (119)$$

**Proof of (117).** Using the (119), we have

$$\begin{aligned}
& \frac{e^{-\frac{d_0^2}{4h}}}{\sqrt{(h-t)}} \int_{-\infty}^{\infty} \frac{r'}{\sqrt{t}} \exp\left(-\frac{h}{4t(h-t)} \left(r' - \frac{d_0 t}{h}\right)^2\right) dr' \\
&= \frac{e^{-\frac{d_0^2}{4h}}}{\sqrt{t(h-t)}} \int_{-\infty}^{\infty} \frac{d_0 t}{h} \exp\left(-\frac{h}{4t(h-t)} \left(r' - \frac{d_0 t}{h}\right)^2\right) dr' \\
&= 2\sqrt{\pi} \frac{e^{-\frac{d_0^2}{4h}}}{\sqrt{t(h-t)}} \frac{d_0 t}{h} \frac{\sqrt{t(h-t)}}{\sqrt{h}} = 2\sqrt{\pi} \exp\left(-\frac{d_0^2}{4h}\right) \frac{d_0 t}{h^{\frac{3}{2}}}
\end{aligned}$$

which is (117).

**Proof of (118).** As before, using (119), we have

$$\begin{aligned}
& \frac{e^{-\frac{d_0^2}{4h}}}{(h-t)^{\frac{1}{2}}} \int_{-\infty}^{\infty} \frac{(d_0 - r')r'}{(h-t)\sqrt{t}} \exp\left(-\frac{h}{4t(h-t)} \left(r' - d_0 \frac{t}{h}\right)^2\right) dr' \\
&= -\frac{e^{-\frac{d_0^2}{4h}}}{(h-t)^{\frac{3}{2}}\sqrt{t}} \int_{-\infty}^{\infty} \left(r' - \frac{d_0 t}{h} + \frac{d_0 t}{h} - d_0\right) \left(r' - \frac{d_0 t}{h} + \frac{d_0 t}{h}\right) \exp\left(-\frac{h}{4t(h-t)} \left(r' - d_0 \frac{t}{h}\right)^2\right) dr' \\
&= -\frac{e^{-\frac{d_0^2}{4h}}}{(h-t)^{\frac{3}{2}}\sqrt{t}} \int_{-\infty}^{\infty} \left[\left(r' - \frac{d_0 t}{h}\right)^2 + \frac{d_0^2}{h^2}(t-h)t\right] \exp\left(-\frac{h}{4t(h-t)} \left(r' - d_0 \frac{t}{h}\right)^2\right) dr' \\
&= -\frac{e^{-\frac{d_0^2}{4h}}}{(h-t)^{\frac{3}{2}}\sqrt{t}} \int_{-\infty}^{\infty} \left[r'^2 + \frac{d_0^2}{h^2}(t-h)t\right] \exp\left(-\frac{h}{4t(h-t)} r'^2\right) dr' \\
&= -\frac{e^{-\frac{d_0^2}{4h}}}{(h-t)^{\frac{3}{2}}\sqrt{t}} \left[\left(\frac{(h-t)t}{h}\right)^{\frac{3}{2}} 4\sqrt{\pi} - \frac{d_0^2}{h^2}(t-h)t \left(\frac{(h-t)t}{h}\right)^{\frac{1}{2}} 2\sqrt{\pi}\right] \\
&= 2\sqrt{\pi} \exp\left(-\frac{d_0^2}{4h}\right) \left(\frac{d_0^2 t}{(h)^{\frac{5}{2}}} - \frac{2t}{(h)^{\frac{3}{2}}}\right)
\end{aligned}$$

which is (118).  $\square$

Now we proceed to derive (90) and (96), which follow almost immediately from the above computations.

**Proof of (90)** – the inner integral of  $J_{111}$ . Since

$$\Sigma_{s,\tau}(p) = s + r(\Sigma_{s,\tau}(p), \tau) \mathbf{N}(s, \tau), \quad \text{and} \quad \tilde{s} = s + r' \mathbf{N}(s, \tau), \quad (120)$$

we have that

$$|\Sigma_{s,\tau}(p) - \tilde{s}|^2 = |r(\Sigma_{s,\tau}(p), \tau) - r'|^2.$$

Thus (90) can be re-written as,

$$\frac{1}{(h-\tau)^{\frac{1}{2}}} \int_{-\infty}^{\infty} \exp\left(-\frac{|r(\Sigma_{s,\tau}(p), \tau) - r'|^2}{4(h-\tau)}\right) \frac{r'}{\sqrt{\tau}} \exp\left(-\frac{r'^2}{4\tau}\right) dr'.$$

Replacing  $r(\Sigma_{s,\tau}(p), \tau)$  by  $d_0$ , this integral is identical to (117).

**Proof of (96)** – the inner integral appearing in  $J_{112}$ . By (120) again, we have

$$\langle \Sigma_{s,\tau}(p) - \tilde{s}, \mathbf{E}_j \rangle = \langle \mathbf{N}(s, \tau), \mathbf{E}_j \rangle (r(\Sigma_{s,\tau}(p), \tau) - r').$$

Replacing once again  $r(\Sigma_{s,\tau}(p), \tau)$  by  $d_0$ , and substituting this back into (96), we get,

$$\frac{\langle \mathbf{N}(s, \tau), \mathbf{E}_j \rangle}{(h-\tau)^{\frac{1}{2}}} \int_{-\infty}^{\infty} \frac{d_0 - r'}{(h-\tau)} \exp\left(-\frac{|d_0 - r'|^2}{4(h-\tau)}\right) \frac{r'}{\sqrt{\tau}} \exp\left(-\frac{r'^2}{4\tau}\right) dr'$$

which is precisely  $\langle \mathbf{N}(s, \tau), \mathbf{E}_j \rangle$  times the integral (118).

### A.3 Proof of Lemma 2: Bound on the Curvature growth of MMC

What follow in this and next sections are by now classical computations about MMC. The readers can refer to [10] for general exposition. For simplicity, we will drop the subscript  $M_t$  and write  $A = A_{M_t}$ ,  $H = H_{M_t}$ ,  $\nabla = \nabla_{M_t}$ , and  $\Delta_{M_t} = \Delta$ .

By [20, Corollary 3.5], the quantity  $|A|^2$  (defined in (35)) satisfies the following equation,

$$\partial_t |A|^2 = \Delta |A|^2 - 2|\nabla A|^2 + 2|A|^4.$$

Let  $x_t$  be the point at which  $|A|^2$  attains its maximum. Then at  $(x_t, t)$ , we have

$$\partial_t |A|^2 \leq 2|A|^4.$$

Hence setting  $f(t) = \max_{M_t} |A|^2$ , we have the following differential inequality:

$$\frac{\partial_t f}{f^2} \leq 2.$$

Integrating both sides in  $t$  gives

$$f(t) \leq \frac{f(0)}{1 - 2f(0)t} \leq f(0)(1 + Cf(0)t).$$

Note that  $C$  may be chosen independent of  $\|A_0\|$  and  $t$  so long as  $\|A_0\|(h|\log h|)^{\frac{1}{4}} \leq 1$ . Taking square roots we then arrive at the desired result,

$$\|A_t\| \leq \|A_0\|(1 + C\|A_0\|^2 t).$$

### A.4 Proof of Lemma 3: Higher Order Regularity of MMC

We follow the arguments given in [11], where a similar bound was proven for the curvature. First note that the two estimates are equivalent, since by [20, Lemma 3.3], we have  $\partial_t \mathbf{N} = \nabla H$ .

First, we quote the following equations given in [11]

$$\begin{aligned} (\partial_t - \Delta)|A|^2 &= -2|\nabla A|^2 + 2|A|^4, \\ (\partial_t - \Delta)|\nabla A|^2 &\leq -2|\nabla^2 A|^2 + C|A|^2|\nabla A|^2. \end{aligned}$$

Introduce  $\psi(t) = \frac{R^2 t}{R^2 + t}$  for some  $R > 0$ . Then we compute:

$$\begin{aligned} (\partial_t - \Delta)(\psi|\nabla A|^2) &\leq -2\psi|\nabla^2 A|^2 + C\psi|A|^2|\nabla A|^2 + |\nabla A|^2 \frac{d}{dt}\psi \\ &\leq -2\psi|\nabla^2 A|^2 + |\nabla A|^2(1 + C\psi|A|^2). \end{aligned} \tag{121}$$

Next define  $f = \psi|\nabla A|^2(\Lambda_0 + |A|^2)$ , where  $\Lambda_0$  is a constant to be chosen later. Then we compute:

$$\begin{aligned} (\partial_t - \Delta)f &\leq (|A|^2 + \Lambda_0)\{|\nabla A|^2(1 + C|A|^2\psi) - 2\psi|\nabla^2 A|^2\} \\ &\quad + \psi|\nabla A|^2\{-2|\nabla A|^2 + 2|A|^4\} - 2\psi\nabla|\nabla A|^2 \cdot \nabla|A|^2. \end{aligned} \tag{122}$$

We estimate the last term as follows:

$$\begin{aligned} -2\psi\nabla|\nabla A|^2 \cdot \nabla|A|^2 &\leq 8\psi|\nabla A||A||\nabla|\nabla A|||\nabla|A|| \\ &\leq 8\psi|\nabla A|^2|A||\nabla^2 A| \quad (\text{by Kato's inequality}) \\ &\leq 2\psi|\nabla^2 A|^2(|A|^2 + \Lambda_0) + \frac{8\psi|A|^2|\nabla A|^4}{|A|^2 + \Lambda_0} \quad (ab \leq a^2\epsilon + \frac{b^2}{4\epsilon}). \end{aligned}$$

Substituting the above into (122), we get:

$$\begin{aligned}
(\partial_t - \Delta)f &\leq (|A|^2 + \Lambda_0)(1 + C|A|^4\psi)|\nabla A|^2 + 2\psi|\nabla A|^2|A|^4 - 2\psi|\nabla A|^2 + \frac{8\psi|A|^2|\nabla A|^4}{|A|^2 + \Lambda_0} \\
&= -\left\{2 - \frac{8|A|^2}{|A|^2 + \Lambda_0}\right\}\psi|\nabla A|^4 + \left\{(|A|^2 + \Lambda_0)(1 + C|A|^2\psi) + 2\psi|A|^4\right\}|\nabla A|^2 \\
&= -\left\{2 - \frac{8|A|^2}{|A|^2 + \Lambda_0}\right\}\frac{f^2}{\psi(|A|^2 + \Lambda_0)^2} + \psi^{-1}f\left\{(1 + C|A|^2\psi) + \frac{2\psi|A|^4}{\lambda_0 + |A|^2}\right\} \\
&= -\psi^{-1}(\delta f^2 - \bar{K}f),
\end{aligned}$$

where  $\delta = \left\{2 - \frac{8|A|^2}{|A|^2 + \Lambda_0}\right\}(|A|^2 + \Lambda_0)^{-2}$  and  $\bar{K} = (1 + C|A|^2\psi) + \frac{2\psi|A|^4}{\Lambda_0 + |A|^2}$ .

Next compute estimates for  $f\eta$ , where  $\eta = (R^2 - (|x - x_0|^2 + 2nt))^2 \equiv (R^2 - r(x, t))^2$  acts as a "localization" function:

$$\begin{aligned}
(\partial_t - \Delta)f\eta &\leq -\eta\psi^{-1}(\delta f^2 - \bar{K}f) + f(\partial_t - \Delta)\eta - 2\nabla f \cdot \nabla\eta \\
&\leq -\eta\psi^{-1}(\delta f^2 - \bar{K}f) + 4(R^2 - r(x, t))|x - x_0|^2 f - 2\nabla(f\eta) \cdot \frac{\nabla\eta}{\eta} \\
&\leq -\eta\psi^{-1}(\delta f^2 - \bar{K}f) + 4R^4 f - 2\nabla(f\eta) \cdot \frac{\nabla\eta}{\eta},
\end{aligned}$$

the last inequality holds over the set  $M_t \cap \{r < R^2\}$ .

Now consider  $m(h) = \sup_{0 \leq t \leq h} \left\{ \sup_{x \in \{M_t | r(x, t) \leq R^2\}} f\eta \right\}$ . Notice that  $\psi \equiv 0$  at  $t = 0$  and hence  $f\eta \equiv 0$  at  $t = 0$  also. Suppose  $m(h)$  is attained at some point  $(\tilde{x}, \tilde{t})$  with  $\tilde{t} \geq 0$ . At this point,  $(\partial_t - \Delta)f\eta \geq 0$ . Thus we can have the following sequence of implications:

$$\begin{aligned}
0 &\leq -\eta\psi^{-1}(\delta f^2 - \bar{K}f) + 4R^4 f \\
\psi^{-1}\delta f^2\eta &\leq \psi^{-1}\bar{K}f\eta + 4R^4 f \\
f^2\eta &\leq \frac{1}{\delta}(\bar{K}f\eta + 4R^4\psi f) \\
f^2\eta^2 &\leq \frac{1}{\delta}(\bar{K}\eta + 4R^4\psi)f\eta.
\end{aligned}$$

Applying Young's inequality to the last line of the above leads to that at  $(\tilde{x}, \tilde{t})$ ,

$$f^2\eta^2 \leq \frac{1}{\delta^2}(\bar{K}\eta + 4R^4\psi)^2.$$

Note that  $\eta \leq R^4$ . Substituting this estimate in above to obtain:

$$m(h)^2 \leq \frac{R^8}{\delta^2}(\bar{K} + 4t)^2, \quad \text{i.e.} \quad m(h) \leq \frac{R^4}{\delta}(\bar{K} + 4t).$$

This means that for all  $(x, t) \in \{M_t | r(x, t) \leq \theta R^2\}$  (for some  $0 < \theta < 1$ ) we have,

$$\psi|\nabla A|^2 \leq (1 - \theta)^{-2}\delta^{-1}(\Lambda_0 + |A|^2)^{-1}(\bar{K} + 4t).$$

Now set  $\Lambda_0 = 8c_0$ . We then have the following estimates/equalities:

- $\delta^{-1}(\Lambda_0 + |A|^2)^{-1} = (|A|^2 + \Lambda_0)\left(2 - \frac{8|A|^2}{|A|^2 + \Lambda_0}\right)^{-1} \leq 6c_0;$

2.  $\bar{K} = (1 + C|A|^2\psi) + \frac{2\psi|A|^4}{\Lambda_0 + |A|^2} \leq C + \frac{2}{9}\psi|A|^2 \leq C$  which is true by our assumption on the magnitude of the curvature in relation to the time step;
3.  $\psi^{-1} = \frac{R^2 + t}{R^2t} = \frac{1}{t} + \frac{1}{R^2}$ .

Utilizing the above estimates we obtain for all  $(x, t) \in \{M_t | r(x, t) \leq \theta R^2\}$ ,

$$|\nabla A|^2 \leq C \cdot c_0 \left( \frac{1}{t} + \frac{1}{R^2} \right).$$

Finally let  $R \rightarrow \infty$  to obtain the estimate for all of  $M_t$ . Taking square roots we get,

$$|\nabla A| \leq C \frac{c_0}{\sqrt{t}}.$$

## A.5 Proof of Lemma 8: the Ball Lemma

The Ball Lemma 8 states that a ball of uniform radius may be placed in a tangential manner in the interior and exterior of the numerical manifolds  $\Gamma_{kh}$  for  $0 \leq k \leq \lfloor \frac{T}{h} \rfloor$  – see Definition 3 to recall the ball property. This result is used to show that the numerical manifold converges to an *embedded* manifold.

The intuitive reason behind the Ball Lemma is quite simple: once a ball of radius  $r_k$  can be put inside and outside of  $\Gamma_{kh}$ , then by regularity of the motion, the same ball but with a smaller radius  $r_{k+1} = r_k - \mathcal{O}(h)$  can still be put inside and outside  $\Gamma_{(k+1)h}$ . Such a statement can then be iterated over multiple time steps. Though this also follows from comparison principle for MMC, we choose the following route of proof for the sake of its more applicability in other geometric flows.

The rigorous proof makes use of the intrinsic distance  $d_0$  between two points on the numerical manifold  $\Gamma_0 \subset \mathbb{R}^{n+1}$ . This is defined as the length of the shortest curve on  $\Gamma_0$  joining two points  $p, q \in \Gamma_0$ :

$$d_0(p, q) = \min \left\{ \int_0^1 |\dot{c}(s)| ds : c : [0, 1] \rightarrow \Gamma_0, c(0) = p, c(1) = q \right\}. \quad (123)$$

Analogously, we will use  $d_k$  to denote the intrinsic distance on the numerical manifold  $\Gamma_{kh}$ .

We first present a preliminary result that bounds the amount by which the intrinsic distance may change over a short time on a manifold moving by its mean curvature. Let  $M_0 \subset \mathbb{R}^{n+1}$  be a smooth, compact  $n$ -dimensional manifold. Recall the function  $F : M_0 \times \mathbb{R}^+ \rightarrow \mathbb{R}^{n+1}$  which solve (2) so that  $M_t = F(M_0, t)$  evolves according to MMC.

**Lemma 10.** *Let  $A_0$  be the Weingarten Map of  $M_0$  satisfying  $\|A_0\| (h |\log h|)^{\frac{1}{4}} \leq 1$ . Let  $c : [0, 1] \rightarrow M_0$  be curve on  $M_0$ . Then the following estimate holds.*

$$|\ell(F(c, t)) - \ell(c)| \lesssim \ell(c) \|A_0\|^2 t,$$

where  $\ell(c) = \int |\dot{c}(s)| ds$  and  $\ell(F(c, t)) = \int |DF(c(s), t)\dot{c}(s)| ds$  are the lengths of  $c$  and  $F(c, t)$  respectively.

*Proof.* Without loss of generality, suppose  $c$  is parametrized by arc length over the interval  $[0, \ell(c)] \subset \mathbb{R}$ . We quickly describe some notation. We use  $\mathbf{N}$  to denote the unit normal to  $M_t$ ,  $H$  its mean curvature,  $\nabla$  denotes the tangential gradient over  $M_t$ , and  $\dot{\cdot}$  denotes differentiation with respect to the arc length variable. Furthermore  $\langle \cdot, \cdot \rangle$  denotes the inner product in  $\mathbb{R}^{n+1}$ .

We first claim that,

$$\left| \frac{d}{dt} \left| \frac{d}{ds} F(c, t) \right|^2 \right| \lesssim \|A_0\|^2 \left| \frac{d}{ds} F(c, t) \right|. \quad (124)$$

Toward this end, consider:

$$\begin{aligned} \frac{d}{dt} \left| \frac{d}{ds} F(c, t) \right|^2 &= 2 \frac{d}{dt} \left\langle \nabla F(c(s), t) \dot{c}(s), \nabla F(c(s), t) \dot{c}(s) \right\rangle \\ &= 4 \left\langle \nabla (\partial_t F)(c(s), t) \dot{c}(s), \nabla F(c(s), t) \dot{c}(s) \right\rangle \\ &= -4 \left\langle \nabla (HN(c(s), t)) \dot{c}(s), \nabla F(c(s), t) \dot{c}(s) \right\rangle \\ &= -4 \left\{ \left\langle \mathbf{N} \otimes \nabla H \dot{c}, \nabla F \dot{c} \right\rangle + H \left\langle \nabla \mathbf{N} \dot{c}, \nabla F \dot{c} \right\rangle \right\}. \end{aligned}$$

For the first term on the right hand side, we compute: further compute:

$$\langle \mathbf{N} \otimes \nabla H \dot{c}, \nabla F \dot{c} \rangle = \sum_{j=1}^{n+1} \sum_{k=1}^{n+1} \partial_k F^{(j)} \dot{c}^{(k)} \mathbf{N}^{(j)} \partial_k H \dot{c}^{(k)} = \sum_{k=1}^{n+1} \partial_k H (\dot{c}^{(k)})^2 \langle \partial_k F, \mathbf{N} \rangle = 0.$$

while for the second term, we have:

$$|H \langle \nabla \mathbf{N} \dot{c}, \nabla F \dot{c} \rangle| \lesssim \|A_0\| \|\nabla F \dot{c}\| \|\nabla \mathbf{N} \dot{c}\| = \|A_0\| \left| \frac{d}{ds} F(c, t) \right| \left| \frac{d}{ds} \mathbf{N}(c, t) \right| \lesssim \|A_0\|^2 \left| \frac{d}{ds} F(c, t) \right|.$$

Combining the above computations leads to (124)

Finally,

$$\begin{aligned} |\ell(F(c, t) - \ell(c))| &= \left| \int_0^t \frac{d}{d\tau} \int_0^{\ell(c)} \left| \frac{d}{ds} F(c, \tau) \right| ds d\tau \right| \\ &= \left| \int_0^t \int_0^{\ell(c)} \frac{d}{d\tau} \left( \left| \frac{d}{ds} F(c, \tau) \right|^2 \right) \frac{1}{2} \left| \frac{d}{ds} F(c, \tau) \right|^{-1} ds d\tau \right| \\ &\lesssim \ell(c) \|A_0\|^2 t \end{aligned}$$

which concludes the proof of the Lemma.  $\square$

With the above, the Ball Lemma 8 then follows from the next two claims. We first let  $r_0 := \frac{1}{C_0+1}$ ,  $B_{r_0, p}^{\text{int}}$  and  $B_{r_0, p}^{\text{ext}}$  be the interior and exterior balls with radius  $r_0$  which are tangent to  $\Gamma_{kh}$  at  $p$  (recall Definition 3), and  $r_* := \frac{C_*}{C_0}$  where  $C_*$  is from (36).

**Claim 1.** Fix a  $p \in \Gamma_{kh}$ . Then

$$\left\{ q \in \Gamma_{kh} : d_k(p, q) \leq r_* \right\} \cap \left( B_{r_0, p}^{\text{int}} \cup B_{r_0, p}^{\text{ext}} \right) = \{p\}. \quad (125)$$

*Proof.* It follows from the remark after (36) that the connected component of  $\Gamma_{kh} \cap B_{r_*}(p)$  containing  $p$  can be written as the graph of a function  $f$  over the tangent plane  $T_p(\Gamma_{kh})$ . By the curvature bound of  $\Gamma_{kh}$ , we infer that

$$\left\{ (x, f(x)) : x \in T_p(\Gamma_{kh}), |x| \leq r_* \right\} \cap \left( B_{r_0, p}^{\text{int}} \cup B_{r_0, p}^{\text{ext}} \right) = \{p\}.$$

Hence any curve in  $\Gamma_{kh}$  that joins  $p$  and any other  $q \in B_{r_0, p}^{\text{int}} \cup B_{r_0, p}^{\text{ext}}$  must have length at least  $r_*$ . Hence the claim follows.  $\square$

**Claim 2.** Define  $m_k := \min \{|p - q| : p, q \in \Gamma_{kh}, d_k(p, q) \geq r_*\}$ . Then

$$m_k \geq m_{k-1} - C(n, C_0)h, \quad (126)$$

for some fixed constant  $C(n, C_0)$  depending only on the spatial dimension  $n$  and the curvature bound  $C_0$ .

*Proof.* We will compare the intrinsic distance between  $\Gamma_{kh}$  and  $\Gamma_{(k-1)h}$ . Recall that  $\Gamma_{kh}$  is obtained as a graph over  $M_{kh} := M(\Gamma_{(k-1)h}, h)$ , the solution at time  $h$  of MMC with initial data  $\Gamma_{(k-1)h}$ . More precisely,

1. using the function  $F$  in (2), we have  $M_{kh} = F(\Gamma_{(k-1)h}, h)$  and,
2. there is a function  $h : M_{kh} \rightarrow \mathbb{R}$  such that  $\Gamma_{kh} = M_{kh} + h\mathbf{N}_{M_{kh}}$ .

Note that both the transformations from  $\Gamma_{(k-1)h}$  to  $M_{kh}$  and from  $M_{kh}$  to  $\Gamma_{kh}$  are diffeomorphisms. Furthermore, by the consistency and stability Theorems 2 and 4, we have  $\|h\|_{L^\infty} \lesssim \|\mathcal{B}_0\|^2 h^{\frac{3}{2}}$ .

Now let  $p, q \in \Gamma_{kh}$  be such that  $d_k(p, q) \geq r_*$ . We can find  $p_1, q_1 \in \Gamma_{(k-1)h}$  and  $p_2, q_2 \in M_{kh}$  satisfying

$$p_2 = F(p_1, h), \quad p = p_2 + h(p_2)\mathbf{N}_{M_{kh}}(p_2), \quad q_2 = F(q_1, h), \quad q = q_2 + h(q_2)\mathbf{N}_{M_{kh}}(q_2).$$

Then we have

$$|p - q| \geq |p_2 - q_2| - C(n)C_0^2(h)^{\frac{3}{2}} \geq |p_1 - q_1| - C_1(n, C_0)h \quad (127)$$

where the first inequality is due to the consistency of the scheme while the second is due to the regularity of MMC.

If  $d_{k-1}(p_1, q_1) \geq r_*$ , then we automatically have  $|p_1 - q_1| \geq m_{k-1}$  and hence we are done.

Now suppose  $d_{k-1}(p_1, q_1) < r_*$ . Making use of Lemma 10, we have

$$\left| d_k(p, q) - d_{(k-1)}(p_1, q_1) \right| \lesssim d_{(k-1)}(p_1, q_1) \|\mathcal{B}_0\|^2 h.$$

from which it follows that

$$d_{k-1}(p_1, q_1) \geq \frac{d_k(p, q)}{1 + \|\mathcal{B}_0\|^2 h} \geq \frac{r_*}{1 + \|\mathcal{B}_0\|^2 h} \geq r_* - C_2(n, C_0)h$$

where  $C_2(n, C_0)$  is another constant depending on the spatial dimension  $n$  and the curvature bound  $C_0$ . Next, by extending the geodesic curve which joins  $p_1$  and  $q_1$ , we can find  $\hat{p}_1, \hat{q}_1 \in \Gamma_{(k-1)h}$  such that  $|p_1 - \hat{p}_1|, |q_1 - \hat{q}_1| \leq C_2(n, C_0)h$  and  $d_{k-1}(\hat{p}_1, \hat{q}_1) \geq r_*$ . Hence

$$|p_1 - q_1| \geq |\hat{p}_1 - \hat{q}_1| - C_2(n, C_0)h \geq m_{k-1} - C_2(n, C_0)h. \quad (128)$$

Now, going back to (127), we have

$$|p - q| \geq |p_1 - q_1| - C_1(n, C_0)h \geq m_{k-1} - (C_1(n, C_0) + C_2(n, C_0))h$$

from which (126) follows.  $\square$

**Acknowledgment.** The authors would like to thank Bob Jerrard for pointing us to this direction, Selim Esedoglu for useful discussion, Tim Laux for reading and suggestions concerning the manuscript, and the Purdue Research Foundation for partial support.

## References

- [1] Guy Barles and Christine Georgelin. A simple proof of convergence for an approximation scheme for computing motions by mean curvature. *SIAM Journal on Numerical Analysis*, 32(2):484–500, 1995.
- [2] Guy Barles and Panagiotis E Souganidis. A new approach to front propagation problems: theory and applications. *Archive for rational mechanics and analysis*, 141(3):237–296, 1998.
- [3] Eric Bonnetier, Elie Bretin, and Antonin Chambolle. Consistency result for a non monotone scheme for anisotropic mean curvature flow. *Interfaces and Free Boundaries*, 14(1):1–35, 2012.
- [4] Haim Brezis. *Functional analysis, Sobolev spaces and partial differential equations*. Springer Science & Business Media, 2010.
- [5] Lia Bronsard and Robert V Kohn. Motion by mean curvature as the singular limit of ginzburg-landau dynamics. *Journal of differential equations*, 90(2):211–237, 1991.
- [6] Xinfu Chen. Generation and propagation of interfaces for reaction-diffusion equations. *Journal of Differential equations*, 96(1):116–141, 1992.
- [7] Yun Gang Chen, Yoshikazu Giga, Shunichi Goto, et al. Uniqueness and existence of viscosity solutions of generalized mean curvature flow equations. *J. Differential Geom*, 33(3):749–786, 1991.
- [8] Wing-Sum Cheung. Some discrete nonlinear inequalities and applications to boundary value problems for difference equations. *Journal of Difference Equations and Applications*, 10(2):213–223, 2004.
- [9] Piero De Mottoni and Michelle Schatzman. Geometrical evolution of developed interfaces. *Transactions of the American Mathematical Society*, 347(5):1533–1589, 1995.
- [10] Klaus Ecker. *Regularity theory for mean curvature flow*. Springer, 2004.
- [11] Klaus Ecker and Gerhard Huisken. Interior estimates for hypersurfaces moving by mean curvature. *Inventiones mathematicae*, 105(1):547–569, 1991.
- [12] Matt Elsey and Selim Esedoglu. Threshold dynamics for anisotropic surface energies. Technical report, preprint, 2016, 2016.
- [13] Selim Esedoglu and Felix Otto. Threshold dynamics for networks with arbitrary surface tensions. *Communications on Pure and Applied Mathematics*, 68(5):808–864, 2015.
- [14] Selim Esedoglu, Steven Ruuth, and Richard Tsai. Threshold dynamics for high order geometric motions. *Interfaces and Free Boundaries*, 10(3):263–282, 2008.
- [15] Lawrence C Evans. Convergence of an algorithm for mean curvature motion. *Indiana University Mathematics Journal*, 42(2):533–557, 1993.
- [16] Lawrence C Evans, H Mete Soner, and Panagiotis E Souganidis. Phase transitions and generalized motion by mean curvature. *Communications on Pure and Applied Mathematics*, 45(9):1097–1123, 1992.
- [17] Lawrence C Evans, Joel Spruck, et al. Motion of level sets by mean curvature i. *J. Diff. Geom*, 33(3):635–681, 1991.



- [18] Lawrence Craig Evans and Ronald F Gariepy. *Measure theory and fine properties of functions*. CRC press, 2015.
- [19] E Giusti. *Minimal surfaces and functions of bounded variation*. Springer Science & Business Media, 1984.
- [20] Gerhard Huisken et al. Flow by mean curvature of convex surfaces into spheres. *Journal of Differential Geometry*, 20(1):237–266, 1984.
- [21] Tom Ilmanen et al. Convergence of the allen-cahn equation to brakkes motion by mean curvature. *J. Differential Geom*, 38(2):417–461, 1993.
- [22] Hitoshi Ishii. A generalization of the bence, merriman and osher algorithm for motion by mean curvature. *Curvature flows and related topics (Levico, 1994)*, 5:111–127, 1995.
- [23] Hitoshi Ishii, Gabriel E Pires, and Panagiotis E Souganidis. Threshold dynamics type approximation schemes for propagating fronts. *Journal of the Mathematical Society of Japan*, 51(2):267–308, 1999.
- [24] Jurgen Jost. *Partial Differential Equations*. Graduate texts in mathematics. Springer, 2002.
- [25] Tim Laux and Felix Otto. Convergence of the thresholding scheme for multi-phase mean-curvature flow. *Calculus of Variations and Partial Differential Equations*, 55(5):129, 2016.
- [26] Fabiana Leoni. Convergence of an approximation scheme for curvature-dependent motions of sets. *SIAM journal on numerical analysis*, 39(4):1115–1131, 2001.
- [27] Stephan Luckhaus and Thomas Sturzenhecker. Implicit time discretization for the mean curvature flow equation. *Calculus of variations and partial differential equations*, 3(2):253–271, 1995.
- [28] Barry Merriman, James Kenyard Bence, and Stanley Osher. *Diffusion generated motion by mean curvature*. CAM Report, 92-18, Department of Mathematics, University of California, Los Angeles, 1992.
- [29] Steven J Ruuth, Barry Merriman, Jack Xin, and Stanley Osher. Diffusion-generated motion by mean curvature for filaments. *Journal of Nonlinear Science*, 11(6):473–493, 2001.
- [30] Halil Mete Soner. Ginzburg-landau equation and motion by mean curvature, i: convergence. *The Journal of Geometric Analysis*, 7(3):437–475, 1997.
- [31] Halil Mete Soner. Ginzburg-landau equation and motion by mean curvature, ii: Development of the initial interface. *The Journal of Geometric Analysis*, 7(3):477–491, 1997.