Journal of Machine Learning Research 24 (2023) 1-49 Submitted 9/22; Revised 6/23; Published 11/23

Large data limit of the MBO scheme for data clustering: convergence of the dynamics

Tim Laux* TIM.LAUX@HCM.UNI-BONN.DE
Hausdorff Center for Mathematics and Institute for Applied Mathematics
University of Bonn
Villa Maria, Endenicher Allee 62, 53115 Bonn, Germany

Jona Lelmi† LELMI@HCM.UNI-BONN.DE
Hausdorff Center for Mathematics and Institute for Applied Mathematics
University of Bonn
Villa Maria, Endenicher Allee 62, 53115 Bonn, Germany

Editor: Marco Cuturi

Abstract

We prove that the dynamics of the MBO scheme for data clustering converge to a viscosity solution to mean curvature flow. The main ingredients are (i) a new abstract convergence result based on quantitative estimates for heat operators and (ii) the derivation of these estimates in the setting of random geometric graphs.

To implement the scheme in practice, two important parameters are the number of eigenvalues for computing the heat operator and the step size of the scheme. The results of the current paper give a theoretical justification for the choice of these parameters in relation to sample size and interaction width.

Keywords: Graph MBO, clustering, semi-supervised learning, continuum limits, viscosity solutions.

1. Introduction

The MBO scheme was originally introduced by Merriman et al. (1992, 1994) as a numerical method to approximate evolution by mean curvature flow. More recently, van Gennip et al. (2014); Merkurjev et al. (2013, 2014) adapted the scheme to problems in data science such as data clustering. The algorithm is a graph based learning method that produces successive partitions of a data set by alternating between two operations: (i) diffusion through the graph heat operator; and (ii) pointwise thresholding. Due to its conceptual simplicity, the MBO scheme is an efficient and robust algorithm. In its original form, the MBO scheme is used for data clustering, and Merkurjev et al. (2013, 2014); Garcia-Cardona et al. (2014) proved that it has comparable or better accuracy than several algorithms commonly adopted in practice. In recent years, graph based learning methods have gained attention for semi-supervised learning tasks (i.e. classification problems with low labeling rate), and several authors have considered modifications of the MBO scheme that retain its conceptual simplicity while improving its accuracy in classification tasks. For instance, VolumeMBO,

*. Now at Universität Regensburg TIM.LAUX@UR.DE
†. Now at University of California, Los Angeles. JONA.LELMI@MATH.UCLA.EDU

introduced by Jacobs et al. (2018a) and PoissonMBO, introduced by Calder et al. (2020), have both showed better accuracy compared to several more well-known methods. The present paper builds a solid mathematical foundation to the MBO scheme and therefore gives yet another motivation to use this algorithm in applications.

Let us recall how the scheme works in the simple case of two classes, i.e. when the goal is to split a dataset $V = \{x_1, ..., x_n\}$ into two subsets. Let $G = (V, W)$ be a weighted graph with vertex set V and weight matrix W. Let Δ be a suitable graph Laplacian. Assume that $\chi^0 : V \to \{0, 1\}$ encodes an initial guess for the clustering. After choosing a step-size $h > 0$ and the number of iterations $N \in \mathbf{N}$ that we want to run, for $0 \le l \le N - 1$ define inductively a new clustering $\chi^{l+1} : V \to \{0, 1\}$ by performing the following two steps:

1. **Diffusion**. For $t > 0$ define
$$u^l(t) := e^{-t\Delta}\chi^l.$$

2. **Thresholding**. Update the clustering by setting
$$\left\{\chi^{l+1} = 1\right\} = \left\{u^l(h) \ge \frac{1}{2}\right\}.$$

By a result of Esedoḡlu and Otto (2015), χ^{l+1} solves

$$\chi^{l+1} \in \operatorname*{argmin}_{u:V\to[0,1]} \left\{E^h_G(u) - E^h_G(u - \chi^l)\right\},$$

where E^h_G is the thresholding energy on G and is defined for $v : V \to [0, 1]$ as

$$E^h_G(v) := \frac{1}{\sqrt{h}}\langle(1 - v), e^{-h\Delta}v\rangle_V,$$

with $\langle\cdot, \cdot\rangle_V$ denoting an inner product on V defined so that Δ becomes self-adjoint. In Laux and Lelmi (2021) we presented the first rigorous study of the large-data limit of the MBO scheme in data clustering. More precisely, given a sequence of random geometric graphs $G_n = (V_n, W_n)$ – i.e. such that $V_n = \{X_1, ..., X_n\}$ for a family $\{X_i\}_{i=1}^{+\infty}$ of i.i.d. random points $X_i \in M$, for a k-dimensional closed Riemannian submanifold $M \subset \mathbf{R}^d$ – we studied the Γ-convergence of the family $\{E^h_{G_n}\}_{n\in\mathbf{N},h>0}$. When the number of iterations of the MBO scheme is very large, its outcome can be thought of as a local minimizer of the thresholding energy, and thus our Γ-convergence result says that this will be qualitatively close to a local minimizer of a suitable variational problem in the continuum. As the selection of the local minimizer strongly depends on the dynamics of the *gradient descent* followed by the algorithm, the next question is to study the convergence properties of said dynamics. This is the content of the present paper: we study the convergence of the dynamics of the MBO scheme in the two-class setting. In general – i.e. when the number of classes to cluster into is greater than two – this is a much harder problem. In the two-class setting the task is easier because one can use the comparison principle for mean curvature flow, and thus the viscosity solutions setting. After the first works on viscosity solutions of Crandall et al. (1992), the machinery has proven to be a solid way to develop a theory of weak solutions for many problems satisfying a maximum principle – and its use is the base for many fundamental contributions in geometric PDEs (Chen et al. (1991); Evans and Spruck (1991)) numerical analysis (Barles and Georgelin (1995); Ishii et al. (1999)) and, more recently, for new results in theoretical data science (Calder (2019a,b); Bungert et al. (2021)).

We will always work with a sequence of weighted geometric graphs $G_n = (V_n, W_n)$, where the vertex sets V_n are defined by $V_n := \{x_i\}_{i=1}^n$, where $\{x_i\}_{i=1}^{+\infty}$ is a sequence of points on $M \subset \mathbf{R}^d$ – a k-dimensional closed Riemannian submanifold of $\mathbf{R}^d$ – and the weight matrix W_n is obtained in the by-now-standard way of weighting the edge between two distinct points with a suitable non-increasing function of the Euclidean distance between them, properly rescaled by a localization parameter $\epsilon_n > 0$, see Section 3 for the precise construction. In this setting, we study the convergence of the sequence of dynamics of the MBO scheme on these graphs as the data size n goes to infinity.

This paper can be conceptually thought of as divided into two main results: in the first one, Theorem 4, we work in an abstract setting. First, in the MBO scheme, we replace the heat operators on the graphs with abstract operators $S_n : (0, +\infty) \times \mathcal{V}_n \to \mathcal{V}_n$ which are linear in the second variable (here $\mathcal{V}_n$ is the space of real-valued functions defined on the vertex set V_n) and we show that if the sequence $\{S_n\}_{n\in\mathbf{N}}$ approximates well-enough the heat kernel corresponding to a weighted Laplace–Beltrami operator on the manifold, then we have convergence of the dynamics of the MBO scheme on the graphs to the viscosity solution of mean curvature flow on the manifold. The conditions that the operators $\{S_n\}$ have to satisfy are three: (i) they should satisfy an approximate maximum principle, (ii) they should approximate the action of the heat kernel on smooth functions in a uniform sense, and (iii) their action on the constant function $\mathbf{1}$ should be close enough to the constant $\mathbf{1}$. All these properties are made quantitatively precise in Theorem 4.

The second main result is Theorem 6 and Corollary 8, where we check that (i), (ii) and (iii) are satisfied with high probability on random geometric graphs – i.e. when the points $\{x_i\}_{i=1}^{+\infty}$ are sampled independently from a probability measure $\nu = \rho \operatorname{Vol}_M \in \mathcal{P}(M)$, absolutely continuous with respect to the volume form – and when S_n are chosen to be the heat operators on the graphs or the operators obtained by cutting off frequencies higher than a threshold K_n defined precisely in Item (iv) in Theorem 6. Let us stress that the latter result is crucial for applications. Indeed, when one implements the MBO scheme on a large dataset, computing the full heat kernel is intractable, and thus one usually works with an approximate version of it obtained by cutting off high frequencies in precisely the way described above. Our result gives a solid mathematical justification for this procedure, proving that the scheme converges in the large data limit to the viscosity solution to mean curvature flow provided the frequency cut-off is chosen according to $K_n \geq (\log(n))^q$ where q is a suitable positive real number and n is the number of data points. We also notice that Theorem 6 gives sufficient conditions on how to choose the length scale ϵ_n and the time-step size h_n in order to ensure convergence of the scheme. In particular, the choice of h_n is not anymore based solely on rules-of-thumb but has theoretical foundations. Previously, only a negative result ensuring pinning of the scheme was known (van Gennip et al., 2014, Theorem 4.2). However, we point out that the conditions on ϵ_n and h_n are only sufficient, but not sharp. Indeed, we expect that the convergence of the scheme should hold true whenever $\epsilon_n = o(h_n)$, while our conditions imply that $\epsilon_n = o(h_n^{3/2})$. The sharp rate $\epsilon_n = o(h_n)$ was verified in the simple setting of the deterministic two-dimensional regular grid $\mathbf{Z}^2$ by Misiats and Yip (2016), and is based on the explicit expression for the heat kernel on regular grids.

But an extension to the general setting in which we are working requires a different strategy, see also the discussion in Remark 5 to better understand how our result compares to the one in the simple setting of Misiats and Yip (2016).

Let us spend a few words on the strategy of the proofs of Theorem 4 and Theorem 6. For Theorem 4 we follow the general scheme of proof of Barles and Georgelin (1995), also used in Misiats and Yip (2016). The authors prove convergence of the classical MBO scheme to a viscosity solution to mean curvature flow in the Euclidean space. Given a smooth open set $\Omega \subset M$, the idea is to prove that the upper semicontinuous envelope u^* and the lower semicontinuous envelope u_* of the piecewise constant in time interpolations of outcomes of the MBO scheme (with initial values $\Omega \cap G_n$) as defined in (2) and (3) are, respectively, a viscosity subsolution and a viscosity supersolution to mean curvature flow on the manifold. After doing that, one has to use the comparison principle in Theorem 16 to compare u^* and u_* with the unique viscosity solution u to mean curvature flow with initial value Ω to show that $\operatorname{sign}_*(u) \leq u_*$ and $\operatorname{sign}^*(u) \geq u^*$. In order to check that u^* and u_* are, respectively, a viscosity subsolution and a viscosity supersolution to mean curvaturue flow we have to adapt the strategy in Barles and Georgelin (1995) to our setting: we need to carefully identify admissible error terms for the argument of Barles and Georgelin (1995). The estimate in Item (ii) in Theorem 4 plays a central role in this, as well as the extension of the consistency step to weighted manifolds (Theorem 22). Finally, to apply the comparison principle in Theorem 16, it is crucial to show an ordering of the initial values in the sense that $\operatorname{sign}_*(u(0,\cdot)) \leq u_*(0,\cdot)$ and $\operatorname{sign}^*(u(0,\cdot)) \geq u^*(0,\cdot)$. We verify this in the general case of a weighted manifold by carefully checking that one iteration of the MBO scheme with step size h produces a set whose normal distance from the previous one is of order h (Theorem 20). This issue seems to have been overlooked in the literature and we believe that our proof fills an important gap in the previous works, even in the Euclidean setting.

For Theorem 6 we draw inspiration from Dunson et al. (2021). There, the authors work on a fixed graph with points sampled independently from a weighted manifold and consider the error in a uniform sense between the restriction of the manifold heat kernel to the graph and the operator obtained by considering the first K frequencies of the graph heat kernel. Their estimate, however, cannot be applied in our setting because, since we want to take the number of data points to infinity, we have to be able to take the frequency cut-off K to infinity together with them. For this reason, a careful interplay between the chosen rates of convergence for K, the step size h and the localization parameter ϵ is needed. In Lemma 9 we obtain a new estimate giving precise conditions on the relation between the frequency cut-off and the number of data points. To get this, we make use of recent results on convergence of spectra of graph Laplacians (García Trillos et al. (2020); Calder and García Trillos (2022); Calder et al. (2022)).

After its introduction in this setting, several authors have developed variants of the MBO scheme. For instance, volume-preserving MBO scheme (Jacobs et al. (2018b); Jacobs (2017))– a version of the algorithm developed by Jacobs, Merkurjev, and Esedoḡlu, where the number of points belonging to each class is invariant through iterations – and poissonMBO – a variant of the scheme for semi-supervised learning at low labeling rates introduced by Calder et al. (2020). When there are just two classes to split the dataset into, we believe that the techniques developed in the present work may be suitably modified to extend the results to these variants of the scheme. One may need to combine our ideas with the techniques developed by Kim and Kwon (2020), where the authors develop a viscosity solution approach for volume preserving mean curvature flow in Euclidean space.

The rest of the paper is organized as follows: in Section 2 we introduce some notation and the two versions of the MBO scheme that we study – the classical one by Merkurjev et al. (2013, 2014); van Gennip et al. (2014), and a more practical one in which the heat operator in the diffusion step is modified by cutting off high frequencies. In Section 3 we state the main results of the current paper: Theorem 4 gives sufficient conditions for the abstract MBO scheme in Algorithm 3 to converge to a viscosity solution to mean curvature flow on a weighted manifold. In Theorem 6 and its corollary, Corollary 8, we show that these conditions are satisfied for the two versions of the algorithm that we study. In Section 4 we introduce the notion of viscosity solution to mean curvature flow on a weighted manifold by simply extending well-known ideas and results in the literature for mean curvature flow on compact manifolds (Ilmanen (1992)) and Euclidean spaces (Chen et al. (1991); Evans and Spruck (1991); Ambrosio and Dancer (2000)). In Appendix A we introduce the MBO scheme on a weighted manifold and we state Theorem 20, which says that one iteration of MBO produces a set whose normal distance from the previous one is of order h, the chosen step-size. In this appendix, we also give an extension to weighted manifolds of the consistency step in the work of Barles and Georgelin (1995). In Appendix B we present the proofs of the results of the paper. In Appendix C, we collect some results about the behavior of the heat kernel on weighted manifolds and on the asymptotics of the spectra for graph Laplacians which are needed in the proofs.

Notation. In the present work, we make extensive use of the Landau symbols o, O. To explain these, we let $\{a_\omega\}_{\omega\in\Omega}, \{b_\omega\}_{\omega\in\Omega}$ be two families of real numbers, with $b_\omega > 0$, indexed by $\omega \in \Omega \subset \mathbf{R}$. Let $\omega_0 \in \mathbf{R} \cup \{-\infty, +\infty\}$ be a limit point for the set Ω, which will be clear from the context. We say that $a_\omega = O(b_\omega)$ if

$$\limsup_{\omega\to\omega_0} \frac{a_\omega}{b_\omega} < +\infty.$$

We say that $a_\omega = o(b_\omega)$ if

$$\lim_{\omega\to\omega_0} \frac{a_\omega}{b_\omega} = 0.$$

We also alternatively write $a_\omega \lesssim b_\omega$ for $a_\omega = O(b_\omega)$ and $a_\omega \ll b_\omega$ for $a_\omega = o(b_\omega)$. In the following, usually (Ω, ω_0) will be $(\mathbf{N}, +\infty)$ or $(\mathbf{R}^+, 0)$, and this will be clear from the context.

2. The MBO scheme on graphs

In this section, we describe the MBO algorithm on graphs originally given by Merkurjev et al. (2013); van Gennip et al. (2014); Merkurjev et al. (2014). We refer to Laux and Lelmi (2021) for more information about its use in data clustering. We consider a weighted connected graph $G = (V, W)$ with n vertices, with $w_{ii} = 0$ for every $i = 1, ..., n$. For each vertex $x_i \in V, i \in \{1, ..., n\}$, we can define

$$d(x_i) = \frac{1}{n} \sum_{j=1}^{n} w_{ij}.$$

We define $D := \operatorname{diag}(d(x_1), ..., d(x_n))$. We let $\mathcal{V} := \{u | u : V \to \mathbf{R}\}$, the set of functions defined on V, which we endow with the inner product

$$\langle u, v\rangle_{\mathcal{V}} := \frac{1}{n}\sum_{i=1}^{n} d(x_i)u(x_i)v(x_i). \tag{1}$$

We define the random walk Laplacian $\Delta : \mathcal{V} \to \mathcal{V}$ as the operator induced by the matrix

$$\Delta := \left(I - \frac{1}{n}D^{-1}W\right).$$

One can check that Δ is non-negative and self-adjoint with respect to $\langle \cdot, \cdot\rangle_{\mathcal{V}}$, in particular, it has n eigenvalues (counted with multiplicity) which we order in the following way

$$0 = \lambda^1 \leq ... \leq \lambda^n.$$

We denote by $\{v^l\}_{1\leq l\leq n}$ a basis of corresponding eigenvectors, orthonormal with respect to $\langle \cdot, \cdot\rangle_{\mathcal{V}}$. For $0 < K \leq n$ we define a kernel $H^K : (0, +\infty) \times V \times V \to \mathbf{R}$ via

$$H^K(t, x, y) := \sum_{l=1}^{K} e^{-t\lambda^l} v^l(x)v^l(y)\frac{d(y)}{n}.$$

The choice $K = n$ corresponds to the heat kernel associated to Δ, which is the unique function $H : (0, +\infty) \times V \times V \to \mathbf{R}$ with the property that for every $u_0 \in \mathcal{V}$, the function

$$u(t, x) := e^{-t\Delta}u_0(x) := \sum_{y\in V} H(t, x, y)u_0(y),\ \ x \in V,\ \ t > 0$$

satisfies

$$\begin{cases} \partial_t u = -\Delta u & \text{on } (0, +\infty) \times V, \\ \lim_{t\downarrow 0} u(t, x) = u_0(x) & \text{on } V. \end{cases}$$

We are now ready to introduce the MBO scheme on graphs.

Algorithm 1 (MBO scheme) *Fix a time-step size $h > 0$ and initial conditions $\chi^0 : V \to \{0, 1\}$. For each $l \in \mathbf{N}$ define inductively $\chi^{l+1} : V \to \{0, 1\}$ as follows:*

1. ***Diffusion**. Define*
$$u^l := e^{-h\Delta}\chi^l.$$

2. ***Thresholding**. Define χ^{l+1} by*
$$\left\{\chi^{l+1} = 1\right\} = \left\{u^l \geq \frac{1}{2}\right\}.$$

We then define the piecewise constant in time, right-continuous interpolation

$$\chi^{h,V}(t, x) = \chi^l(x) \text{ for } t \in [lh, (l + 1)h) \text{ and } x \in V.$$

We are interested in understanding whether this approximation is consistent at the level of the evolution by mean curvature flow on the manifold.

In practice, computing the exact diffusion in the first step of the algorithm may be computationally intractable. For this reason, one usually implements the MBO scheme by considering only a smaller number of eigenvectors of the Laplacian, say K. In other words, one uses the following more efficient variant of MBO.

Algorithm 2 (Approximate MBO scheme) *Fix a time-step size $h > 0$ and initial conditions $\chi^0 : V \to \{0, 1\}$. For each $l \in \mathbf{N}$ define inductively $\chi^{l+1} : V \to \{0, 1\}$ as follows:*

1. ***Diffusion**. Define*
$$u^l(x) := \sum_{y \in V} H^K(h, x, y)\chi^l(y).$$
2. ***Thresholding**. Define χ^{l+1} by*
$$\left\{\chi^{l+1} = 1\right\} = \left\{u^l \geq \frac{1}{2}\right\}.$$

Again, we then define the piecewise constant in time, right-continuous interpolation

$$\chi^{h,V,K}(t, x) = \chi^l(x) \text{ for } t \in [lh, (l + 1)h) \text{ and } x \in V.$$

At present, the choice of h and the exact value of K to pick in order to get a good approximation of the MBO scheme is obtained by trial and error. In this work, under the standard *manifold assumption*, we rigorously justify that an admissible regime to get a consistent result in the large-data limit is $K \geq (\log(n))^q$, $h \gg (\log(n))^{-\alpha}$ for some $q, \alpha > 0$ (see Theorem 6 for the precise choices of q, α). The lower bound on K is consistent with the common choice used by practitioners, and it is reminiscent of the scaling in Johnson–Lindenstrauss' Lemma. The requirement on h is not sharp. However, to avoid pinning (and also to speed up convergence of the scheme towards local minimizers), practitioners tend to use relatively large step sizes.

3. Main results

Our convergence analysis is divided into two distinct steps. Our first main result is an abstract convergence theorem providing general criteria for the convergence of an MBO scheme over a discrete structure to mean curvature flow. The second main result verifies the conditions of the abstract statement in the setting of random geometric graphs.

3.1 Abstract convergence result

Hereafter $M \subset \mathbf{R}^d$ is a k-dimensional closed Riemannian submanifold. We denote by g its metric and by d_M the induced distance. Let $\{x_i\}_{i=1}^{+\infty}$ be a sequence of points on M. Let $V_n := \{x_1, \ldots, x_n\}$ be the set of the first n points, and let $\mathcal{V}_n$ be the set of real-valued functions defined on V_n. Assume that we are given a sequence of operators $S_n : (0, +\infty) \times \mathcal{V}_n \to \mathcal{V}_n$ which are linear in the second variable, we then consider the following abstract version of the MBO scheme on V_n.

Algorithm 3 (Abstract MBO scheme) *Fix a time-step size $h_n > 0$ and initial conditions $\chi^{0,V_n} : V_n \to \{0,1\}$. For each $l \in \mathbf{N}$ define inductively $\chi^{l+1,V_n} : V_n \to \{0,1\}$ as follows:*

1. ***Diffusion**. Define*
$$u_n^l := S_n(h_n, \chi^{l,V_n}).$$
2. ***Thresholding**. Define χ^{l+1,V_n} by*
$$\left\{\chi^{l+1,V_n} = 1\right\} = \left\{u_n^l \geq \frac{1}{2}\right\}.$$

We then define $\chi^{h_n,V_n} : [0,+\infty) \times V_n \to \{0,1\}$ by

$$\chi^{h_n,V_n}(t,x) := \chi^{l,V_n}(x),\ x \in V_n,\ t \in [lh_n,(l+1)h_n).$$

For convenience, we will mostly work with the $\{-1,1\}$-valued functions

$$u^{h_n,V_n}(t,x) := 2\chi^{h_n,V_n}(t,x) - 1.$$

We also define the upper and lower limits of the family $\{u^{h_n,V_n}\}_{n\in\mathbf{N}}$ as

$$u^*(t,x) := \sup\left\{\limsup_{n\to+\infty} u^{h_n,V_n}(t_n,x_n) \,\middle|\, t_n > 0,\ \lim_{n\to+\infty} t_n = t,\ x_n \in V_n,\ \lim_{n\to+\infty} x_n = x\right\}, \tag{2}$$

$$u_*(t,x) := \inf\left\{\liminf_{n\to+\infty} u^{h_n,V_n}(t_n,x_n) \,\middle|\, t_n > 0,\ \lim_{n\to+\infty} t_n = t,\ x_n \in V_n,\ \lim_{n\to+\infty} x_n = x\right\}. \tag{3}$$

Let $\xi > 0$ be a smooth function on the manifold M, observe that since we assume the manifold M to be compact, ξ is actually bounded away from zero. Let $\Omega_0 \subset M$ be an open set with smooth boundary Γ_0. We let $u : [0,+\infty) \times M \to \mathbf{R}$ be the unique viscosity solution of the level set formulation of the mean curvature flow with density ξ (see Section 4 for the details) with initial value $sd(\cdot,\Gamma_0) = d_M(x,\Omega_0^c) - d_M(x,\Omega_0)$, the signed distance function from Γ_0. For any $t > 0$ we also define

$$\Omega_t := \{x \in M \mid u(t,x) > 0\},\ \Gamma_t = \{x \in M \mid u(t,x) = 0\}. \tag{4}$$

Note that by definition, the viscosity solution u is continuous and thus the sets we just introduced are well-defined. Let us denote by Δ_ξ the weighted Laplacian on M with weight $\mu := \xi \operatorname{Vol}_M$, i.e.,

$$\Delta_\xi f = -\frac{1}{\xi}\operatorname{div}(\xi\nabla f) \quad \text{for } f \in C^\infty(M).$$

We denote by $e^{-t\Delta_\xi}$ the corresponding heat semigroup and by $H : (0,+\infty) \times M \times M \to \mathbf{R}$ the corresponding heat kernel.

Our first main result is the following conditional convergence of the abstract formulation of the MBO scheme when the initial data on V_n is chosen to be the set of points contained in Ω_0.

Theorem 4 *Let $\chi^{0,G_n} := \mathbf{1}_{V_n\cap\Omega_0}$ for each $n \in \mathbf{N}$. Assume that:*

(i) The operators S_n satisfy the maximum principle up to errors $h_n^{3/2}$, i.e., for n large enough and for each $u, v \in \mathcal{V}_n$ it holds

$$u \le v \Rightarrow S_n(h_n, u) \le S_n(h_n, v) + \left(\max_{V_n} |u| + \max_{V_n} |v|\right) O(h_n^{3/2}).$$

(ii) The operators S_n approximate the heat operator on the manifold, i.e., there exists a constant $\kappa > 0$ such that for every function $f \in C^\infty(M)$ we have

$$\max_{x\in V_n} \left|S_n(h_n, f)(x) - e^{-h_n\kappa\Delta_\xi} f(x)\right| = (\sup|f|)\, o(\sqrt{h_n}) + \mathrm{Lip}(f)\, O(h_n^{3/2}). \tag{5}$$

where the functions $o(\sqrt{h_n}), O(h_n^{3/2})$ are independent of f.

(iii) The operators S_n almost preserve the total mass in the sense that

$$\max_{x\in V_n} |S_n(h_n, \mathbf{1}_{V_n})(x) - 1| = O(h_n^{3/2}).$$

Then u^* and u_* defined in (2) *and, respectively,* (3) *satisfy*

$$u_*(x,t) = 1 \quad \text{if } x \in \Omega_t, \tag{6}$$
$$u^*(x,t) = -1 \text{ if } x \in (\Omega_t \cup \Gamma_t)^c. \tag{7}$$

Here Ω_t and Γ_t are defined as in (4).

Remark 5 *(i) We will be interested in applying Theorem 4 when the set V_n is endowed with a weighted graph structure, and the operator S_n is chosen to be, for instance, the heat semigroup associated to the random walk Laplacian Δ_n. In this context, Item* (i) *in Theorem 4 holds true without the error term. Indeed, the comparison principle for the heat operator precisely says that for any pair of functions $u, v \in \mathcal{V}_n$ it holds that*

$$u \le v \ \Rightarrow e^{-t\Delta_n} u \le e^{-t\Delta_n} v \quad \forall t \ge 0. \tag{8}$$

To clarify, we point out that by $u \le v$ we mean that the inequality $u(x) \le v(x)$ holds pointwise for all $x \in V_n$. For readers who are not familiar with these PDE methods, we present a short argument for (8). It is easy to see that, by linearity, it suffices to prove the claim for $u = 0$. For this, we show the stronger statement:

$$\begin{cases} \frac{d}{dt} v \ge -\Delta_n v, \\ v(0,\cdot) \ge 0. \end{cases} \quad \Rightarrow v(t,x) \ge 0, \ \forall x \in V_n, t \ge 0. \tag{9}$$

Claim (9) *is proved as follows: First, observe that by adding $t\delta$ and letting $\delta \downarrow 0$ we can reduce to the case where $\frac{d}{dt} v > -\Delta_n v$. In this case, assume the claim does not hold. Then there exist $t_0 > 0$ and $j_0 \in \{1, \dots, n\}$ such that*

$$v(t_0, x_{j_0}) = \min_{x\in V_n} v(t_0, x) < 0, \quad \frac{d}{dt} v(t_0, x_{j_0}) \le 0.$$

In particular we have, due to the minimality of x_{j_0}*,*

$$\Delta_n v(t_0, x_{j_0}) = \sum_{i \neq j_0} \frac{w_{j_0,i}}{n d_n(x_i)} (v(t_0, x_{j_0}) - v(t_0, x_i)) \leq 0.$$

These two inequalities imply that $(\frac{d}{dt}v + \Delta_n v)(t_0, x_{j_0}) \leq 0$*, which is a contradiction to* $\frac{d}{dt}v > -\Delta_n v$*.*

(ii) *Let us compare Theorem 4 with Misiats and Yip (2016), where the authors prove convergence of the dynamics of the graph MBO scheme to a viscosity solution to mean curvature flow in the case of regular, two-dimensional grids. More precisely, they work in the following setting: the manifold* M *is the standard Euclidean plane* $\mathbf{R}^2$*,and the sets* V_n *are given by* $V_n := \epsilon_n \mathbf{Z}^2$ *for a sequence of localization parameters* $\epsilon_n \downarrow 0$*. To put ourselves in a setting that is precisely the one we are working in we could actually work with* $M = \mathbf{T}^2$*, the 2-dimensional torus, and the sequence* $V_n \cap \mathbf{T}^2$*, but to keep the presentation simple we prefer to continue this discussion in the precise setting of Misiats and Yip (2016). Let* $v : \epsilon_n \mathbf{Z}^2 \to \mathbf{R}$ *be a function which is zero outside a compact subset of* $\mathbf{R}^2$*. We denote by* $S_n(t, v) : [0, +\infty) \times \epsilon_n \mathbf{Z}^2 \to \mathbf{R}$ *the solution to the discrete heat equation with initial value* v*, i.e.,* $u := S_n(t, v)$ *solves*

$$\begin{cases} \frac{d}{dt}u(t,(i,j)) = \frac{1}{\epsilon_n^2}\Big[u(t,(i+1,j)) + u(t,(i-1,j)) & \\ \qquad + u(t,(i,j+1)) + u(t,(i,j-1)) & \text{for } (i,j) \in \epsilon_n \mathbf{Z}^2, \\ \qquad - 4u(t,(i,j))\Big] & \\ u(0,(i,j)) = v((i,j)) & \text{for } (i,j) \in \epsilon_n \mathbf{Z}^2. \end{cases}$$

By using Fourier analysis methods, it can be shown that for every $h > 0$ *and every* $(x_1, x_2) \in \epsilon_n \mathbf{Z}^2$

$$S_n(h, v)((x_1, x_2)) = \sum_{(i,j) \in \epsilon_n \mathbf{Z}^2} Q_{i-x_1}\left(\frac{2h}{\epsilon_n^2}\right) Q_{j-x_2}\left(\frac{2h}{\epsilon_n^2}\right) v((i,j)),$$

where

$$Q_l(\alpha) := \frac{1}{2\pi} \int_{-\pi}^{\pi} \cos(l\xi) e^{\alpha(\cos(\xi)-1)} d\xi. \tag{10}$$

Using the asymptotic expansions (Misiats and Yip, 2016, Proposition 3) for (10) *it is not hard to prove that for any smooth, compactly supported function* $f \in C_c^\infty(\mathbf{R}^2)$

$$\begin{aligned} \sup_{(i,j) \in \epsilon_n \mathbf{Z}^2} \left| S_n(h,f)((i,j)) - G_h^{\mathbf{R}^2} * f((i,j)) \right| &= \mathrm{Lip}(f) o(\epsilon_n) + \sup |f| O\left(\frac{\epsilon_n^2}{h}\right) \\ &\quad + \sup |f| O\left(\frac{\epsilon_n}{\sqrt{h}} \log\left(\frac{\epsilon_n}{\sqrt{h}}\right)\right), \end{aligned} \tag{11}$$

where $G_h^{\mathbf{R}^2}$ *denotes the heat kernel in the Euclidean plane at time* h*. In particular, when* $\epsilon_n = h_n^\alpha$ *for* $\alpha \geq \frac{3}{2}$*, we see that* (11) *implies* (5)*. This allows us to use Theorem 4* *to recover the results of Misiats and Yip (2016) when $\alpha \geq \frac{3}{2}$. Actually, an inspection of the proof of Theorem 4 shows that to check that u^* and u_* are, respectively, a viscosity subsolution and a viscosity supersolution to mean curvature flow, estimate* (5) *can be replaced by*

$$\max_{x \in V_n} \left| S(h_n, f)(x) - e^{-h\kappa\Delta_\xi} f(x) \right| = (\sup|f|)\, o(\sqrt{h_n}) + \mathrm{Lip}(f)\, O(h_n^\gamma),$$

for some $\gamma > 1$. In particular, we see that in the setting of the two-dimensional regular grid this is satisfied whenever $\epsilon_n = h_n^\gamma$. This allows to recover the full parameter range $\gamma > 1$ of Misiats and Yip (2016). We need the slightly stronger assumption $\gamma = \frac{3}{2}$ for checking the initial conditions for u^* and u_*, i.e. to verify that

$$\begin{cases} u^*(0,x) \leq \mathrm{sign}^*(u(0,x)) & x \in M, \\ u_*(0,x) \geq \mathrm{sign}_*(u(0,x)) & x \in M. \end{cases} \tag{12}$$

Inequalities (12) *are needed to be able to apply the comparison principle for viscosity solutions to mean curvature flow, which is a crucial ingredient in our proof, as well as in Barles and Georgelin (1995); Misiats and Yip (2016). Again, we point out that this step seems to have been overlooked in the latter works.*

3.2 Results on the MBO scheme and on the approximate MBO scheme

Here, we focus on the MBO scheme on weighted graphs, where the operator in the abstract scheme is replaced by the heat operator or a suitable spectral approximation of it. In this case, when the points $\{x_i\}_{i=1}^{+\infty}$ are sampled independently from a suitable probability distrubution on a k-dimensional closed Riemannian submanifold $M \subset \mathbf{R}^d$, we will verify that Items (i), (ii), (iii) in Theorem 4 hold true with high probability, and we will thus conclude the almost sure convergence of the MBO scheme to mean curvature flow. We will work in the following setting. For each $n \in \mathbf{N}$ we define weighted graphs $G_n = (V_n, W_n)$ where the vertex set V_n is given by $\{x_1, ..., x_n\}$ and the adjacency matrix $W_n = (w_{ij}^{(n,\epsilon_n)})_{1 \leq i,j \leq n}$ is given by

$$\begin{aligned} w_{ii}^{(n,\epsilon_n)} &= 0 \text{ for } 1 \leq i \leq n, \\ w_{ij}^{(n,\epsilon_n)} &= \frac{1}{\epsilon_n^k} \eta\left(\frac{\|x_i - x_j\|_d}{\epsilon_n}\right) \text{ for } 1 \leq i,j \leq n,\ i \neq j. \end{aligned}$$

Here $\epsilon_n > 0$ are given length scales and $\eta : [0, +\infty) \to [0, +\infty)$ is a non-increasing function with support on the interval $[0,1]$, whose restriction to the interval $[0,1]$ is Lipschitz continuous. We define

$$C_1 := \int_{\mathbf{R}^k} \eta(\|y\|_k)dy, \quad C_2 := \int_{\mathbf{R}^k} \eta(\|y\|_k) y_1^2 dy, \quad \kappa(\eta) := \frac{C_2}{2C_1}.$$

We also define, for every $x \in M$ and every $n \in \mathbf{N}$,

$$d_n(x) := \frac{1}{n} \sum_{j=1}^{n} \frac{1}{\epsilon_n^k} \eta\left(\frac{\|x - x_j\|_d}{\epsilon_n}\right) \mathbf{1}_{\{x \neq x_j\}}.$$

Note that, when $x = x_i$ for some $1 \leq i \leq n$, then $d_n(x)$ is the degree of the i-th node. We denote by $D_n := \mathrm{diag}(d_n(x_1), ..., d_n(x_n))$ the diagonal matrix of the degrees. The random walk Laplacian Δ_n is the linear operator induced by the $(n \times n)$-matrix given by

$$\Delta_n := \frac{1}{\epsilon_n^2}\left(I - \frac{1}{n}D_n^{-1}W_n\right).$$

We denote by $\{v_n^l\}_{1\leq l\leq n}$ an orthonormal basis (with respect to the inner product $\langle\cdot,\cdot\rangle_{\mathcal{V}_n}$, see (1)) made of eigenvectors for the Laplacian Δ_n corresponding to the eigenvalues $\{\lambda_n^l\}_{1\leq l\leq n}$, which are ordered in the following way

$$0 = \lambda_n^1 \leq \lambda_n^2 \leq ... \leq \lambda_n^n.$$

Like in Section 2, for every $0 < K \leq n$ we define

$$H_n^K(t,x,y) = \sum_{l=1}^{K} e^{-t\lambda_n^l} v_n^l(x) v_n^l(y) \frac{d_n(y)}{n},$$

and we set $H_n = H_n^n$ when $K = n$. The MBO scheme as stated in Algorithm 1 corresponds to the choices $S_n(t,\cdot) = e^{-t\Delta_n}(\cdot)$, the heat semigroup on the n-th graph, which acts on functions $u \in \mathcal{V}_n$ by

$$e^{-t\Delta_n}(u)(x) = \sum_{y\in V_n} H_n(t,x,y)u(y).$$

Let $0 < K_n \leq n$ be a sequence of numbers converging to $+\infty$, then the approximate MBO scheme as stated in Algorithm 2 corresponds to the choices $S_n = P_n$, where the operators P_n act on functions $u \in \mathcal{V}_n$ by

$$P_n(t,u)(x) := \sum_{y\in V_n} H_n^{K_n}(t,x,y)u(y). \tag{13}$$

Our second main result states that on random geometric graphs the operators $e^{-t\Delta_n}(\cdot)$ and P_n satisfy the assumptions of Theorem 4 with high probability.

Theorem 6 *Let us assume that $\nu := \rho\,\mathrm{Vol}_M$ is a probability measure with a smooth and positive density ρ, and denote by $\{\lambda_i\}_{i\in\mathbf{N}}$ the spectrum of the ρ^2-weighted Laplacian on M in non-decreasing order. Assume that the points $\{x_i\}_{i=1}^{+\infty}$ in the above construction are i.i.d. random points sampled from M, distributed according to ν. Assume that $q > 0, \frac{2}{k} > s > 0$ are such that:*

(i) $q > \frac{1}{\frac{2}{k}-s}$,

(ii) It holds that $\inf_{i\in\mathbf{N}}(\lambda_{i+1} - \lambda_i) > 0$.

(iii) $K_n \geq (\log(n))^q$,

(iv) $h_n \gg (\log(n))^{-\alpha}$, *with* $\alpha = -1 + \frac{2q}{k} - sq \geq 0$,

(v) $\epsilon_n \ll (\log(n))^{-\beta}$, *with* $\beta = -\frac{1}{2} + 4q + \frac{13q}{k} - \frac{sq}{2} \geq 0$,

(vi) It holds that

$$\epsilon_n \gtrsim \begin{cases} \left(\frac{\log(n)}{n}\right)^{\frac{1}{k}} & \text{if } k \geq 3, \\ \left(\frac{\log(n)}{n}\right)^{\frac{1}{8}} & \text{if } k = 2. \end{cases}$$

Then the operators $e^{-t\Delta_n}(\cdot)$ *and* P_n *satisfy conditions* (i), (ii) *and* (iii) *in Theorem 4 (with* $\xi = \rho^2$ *and* $\kappa = \kappa(\eta)$*) on* G_n *with probability greater than*

$$1 - C\epsilon_n^{-6k} \exp(-\frac{n\epsilon_n^{k+4}}{C}) - Cn \exp(-\frac{n}{CK_n^2}).$$

Remark 7 *Let us comment on this second result.*

(i) For each $k \geq 2$*, the space of admissible parameters* (s, q) *in Theorem 6 is quite* large. *To see this, we plot the space of admissible parameters. The shaded region represents the space of admissible pairs* (s, q)*.*

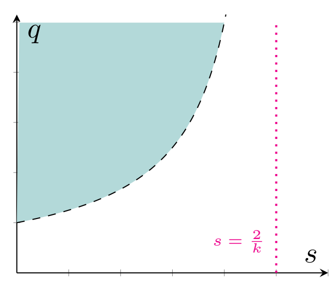

Figure 1: Parameter space.

(ii) Condition (ii) *in Theorem 6 contains two implicit assumptions: the first one is that the eigenvalues of the manifold Laplacian are simple, and the second one is that the gap between the eigenvalues counted without multiplicity is bounded away from zero. The first assumption is of technical nature, and we included it for simplicity of exposition, although we believe that the result can be proved even if the eigenvalues are not simple by working with eigenprojections instead of eigenfunctions. The second assumption concerns the geometry of the manifold* M *and it is crucial for the proof. It is for example satisfied by the* k*-torus and by the* k*-sphere with standard unit density, see (Chavel, 1984, Chapter II, Section 2) and (Chavel, 1984, Chapter II, Section 4).*

(iii) Let us observe that conditions (v) *and* (vi) *in Theorem 6 are compatible, indeed the right-hand side of* (v) *in Theorem 6 is a rational function of* $\log(n)$*, while the lower bound in condition* (vi) *in Theorem 6 converges to zero as a power of* n*, up to a logarithmic factor. We also remark that items* (iv) *and* (v) *of Theorem 6 imply in particular*

$$\epsilon_n \lesssim h_n^{3/2},$$

an estimate that will also be crucial later in the proof. In analogy to the regular grid setting of Misiats and Yip (2016), we expect that the convergence of the scheme should be true up to the critical scaling

$$\epsilon_n \ll h_n.$$

Observe furthermore that condition (iv) *in Theorem 6 gives a lower bound for* h_n *of the form*

$$h_n \gg (\log(\delta_n))^\alpha,$$

where $\delta_n = (\frac{1}{n})^{1/k}$ *is the characteristic distance between the nodes of the graph. This is perhaps not too surprising because the diffusion needs some time to smear out the fine details in the graph that appear at its characteristic length scale. A similar condition already appeared in Dunson et al. (2021).*

Corollary 8 *In the setting of Theorem 6, if we additionally assume that*

$$\epsilon_n \gg \left(\frac{\log(n)}{n}\right)^{\frac{1}{k+4}}, \tag{14}$$

then the conclusion of Theorem 4 holds almost surely both for the MBO scheme, Algorithm 1, and the approximate MBO scheme, Algorithm 2.

An important ingredient for the proof of Theorem 6 is the following lemma, which gives an estimate of the distance between the approximate heat kernel on the graph and the heat kernel on the manifold in a uniform sense. Such heat kernel estimates are of independent interest, for example, one should compare with (Dunson et al., 2021, Theorem 3), where the authors obtain a similar estimate when the frequency cut-off K_n and the time-scale h_n are fixed. In Lemma 9 we improve their result by showing how to choose K_n in terms of n as $n \to +\infty$.

Lemma 9 *In the setting of Theorem 6, there exist constants* $a_1, a_2, a_3, a_4 > 0$ *such that if* n *is large enough, then, with probability greater than* $1 - a_1\epsilon_n^{-6k}\exp(-a_2 n\epsilon_n^{k+4}) - a_3 n\exp(-a_4 n\,(\lambda_{K_n}+1)^{-k})$*, we have*

$$\max_{x,y\in V_n}\left|H_{\epsilon_n}^{K_n}(h_n,x,y) - \frac{\rho(y)}{n}H(\kappa(\eta)h_n,x,y)\right| = o\left(\frac{\sqrt{h_n}}{n}\right). \tag{15}$$

4. The level set equation for MCF on a weighted manifold

In this section, we provide the basic framework for viscosity solutions to mean curvature flow in weighted Riemannian manifolds.

Hereafter (M,g) is a k-dimensional closed Riemannian manifold, and $\xi > 0$ is a smooth function on M. Recall that the evolution of a smooth open set Ω_0 by mean curvature follows the trajectory of steepest descent of the area functional (see, for instance, Proposition 1.2.1 in Mantegazza (2011)), which is defined as

$$\Omega \mapsto \int_{\partial\Omega} dS,$$

where Ω ranges over all open sets in M with a smooth boundary. When we consider a weight ξ on the manifold, the correct functional to consider is the weighted-area functional, defined as

$$\Omega \mapsto \int_{\partial\Omega} \xi dS,$$

where Ω ranges over all open sets in M with smooth boundary. We define the evolution of mean curvature flow with density ξ - hereafter denoted as MCF_ξ - as the trajectory of steepest descent of this functional. To derive an equation for MCF_ξ we consider a family $\{\Omega(t)\}_{0\leq t<T}$ of smooth open sets evolving smoothly in time with normal velocity vector V. Denote by $\nu(t)$ a smooth extension of the outer unit normal such that $g(\nu(t),\nu(t)) = 1$ in a neighborhood of $\partial\Omega(t)$. We then have by Gauss' Theorem

$$\begin{aligned}
\frac{d}{dt}\int_{\partial\Omega(t)} \xi dS &= \frac{d}{dt}\int_{\Omega(t)} \frac{1}{\xi}\operatorname{div}(\xi\nu(t))\xi d\operatorname{Vol}_M \\
&= \int_{\partial\Omega(t)} \frac{1}{\xi}\operatorname{div}(\xi\nu(t))g(V(t),\nu(t))\xi dS + \int_{\Omega(t)} \operatorname{div}(\xi\partial_t\nu(t))d\operatorname{Vol}_M \\
&= \int_{\partial\Omega(t)} \frac{1}{\xi}\operatorname{div}(\xi\nu(t))g(V(t),\nu(t))\xi dS + \int_{\partial\Omega(t)} \xi g(\partial_t\nu(t),\nu(t))dS \\
&= \int_{\partial\Omega(t)} \frac{1}{\xi}\operatorname{div}(\xi\nu(t))g(V(t),\nu(t))\xi dS,
\end{aligned}$$

where in the last line we used that $g(\partial_t\nu(t),\nu(t)) = \frac{1}{2}\frac{d}{dt}g(\nu(t),\nu(t)) = 0$ because $g(\nu(t),\nu(t)) = 1$ in a neighborhood of $\partial\Omega(t)$. We thus see that the trajectory of steepest descent is given by

$$g(V,\nu) = -\frac{1}{\xi}\operatorname{div}(\xi\nu).$$

We are thus led to the following definition.

Definition 10 *Let (M,g) be a smooth k-dimensional closed Riemannian manifold. Let $\xi > 0$ be a smooth function on M. A family $\{\Omega_t\}_{t\geq 0}$ of smooth open subsets of M is said to evolve by MCF_ξ if*

$$g(V,\nu) = -\frac{1}{\xi}\operatorname{div}(\xi\nu). \tag{16}$$

where V is the velocity vector field of the evolution and ν is the outer unit normal field.

Remark 11 *Let $N \subset M$ be a $(k-1)$-dimensional Riemannian submanifold of M. Recall that the scalar mean curvature of N is the map $H : N \to \mathbf{R}$ defined for $x \in N$ as $H(x) = tr(s_x)$, where $tr(s_x)$ denotes the trace of the second fundamental form of N in M at x. Let $\{\Omega_t\}_{t\geq 0}$ be as in Definition 10, it can be shown that the mean curvature $H(t)$ of $\partial\Omega(t)$ satisfies $\operatorname{div}(\nu(t)) = H(t)$, equation (16) can thus be rewritten as*

$$g(V,\nu) = -H - g\left(\frac{\nabla\xi}{\xi},\nu\right), \tag{17}$$

which yields the following interpretation for (16): the evolution by MCF_ξ as defined in Definition 10 is driven by the minimization of two quantities, area and density. The first *term on the right-hand side of* (17) *forces the evolution to follow a trajectory which decreases as much as possible the area of* $\partial\Omega(t)$*, whereas the second term on the right-hand side forces the evolution to move towards regions where the density* ξ *is low.*

We now derive the corresponding level set formulation for the above evolution in the spirit of Evans and Spruck (1991); Chen et al. (1991). Let $u : [0, +\infty) \times M \to \mathbf{R}$ be a smooth function, assume for this heuristic discussion that $Du \neq 0$ everywhere. For any $s \in \mathbf{R}$ define $\Omega_t^s := \{x \in M : \ u(t,x) > s\}$ and assume that $\{\Omega_t^s\}_{t\geq 0}$ evolves by MCF_ξ defined in Definition 10. Let $s \in \mathbf{R}$ and let $x : (0,T) \to M$ a smooth curve such that $x(t) \in \partial\Omega_t^s$ for every time $0 < t < T$. Then

$$\begin{aligned}0 &= \frac{d}{dt}u(t,x(t)) \\ &= (\partial_t u)(t,x(t)) + g(\nabla u(t,x(t)), \dot{x}(t)).\end{aligned}$$

Using the fact that the outer normal to the super level set Ω_t^s is given by $\nu(t,x) = -\frac{\nabla u(t,x)}{|\nabla u(t,x)|}$ and plugging in (16) we obtain

$$\begin{aligned}(\partial_t u)(t,x(t)) &= |\nabla u(t,x(t))| g(\nu(t,x(t)), V(t,x(t))) \\ &= -|\nabla u(t,x(t))| \frac{1}{\xi(x(t))} \operatorname{div}(\xi\nu)(t,x(t)).\end{aligned}$$

Using the product rule for the divergence and recalling that $\nu = -\frac{\nabla u}{|\nabla u|}$ we observe that u solves

$$\partial_t u = \left\langle g - \frac{Du \otimes Du}{|Du|^2}, D^2 u \right\rangle + g\left(\frac{\nabla \xi}{\xi}, \nabla u\right), \tag{18}$$

where we denoted by $\langle\cdot,\cdot\rangle$ the extension of g to the linear bundle of $T^*M \otimes T^*M$, i.e. for A, B sections of $T^*M \otimes T^*M$ we have in local coordinates

$$\langle A, B\rangle := \sum_{i,j,k,l=1}^{k} A_{ij} g^{jk} g^{kl} B_{li}.$$

From (18) we are led to the following definition.

Definition 12 *Let* $u : (0,T) \times M \to \mathbf{R}$ *be a smooth function with* $Du \neq 0$ *everywhere. Then* u *is said to solve the level set formulation of* MCF_ξ *if* (18) *holds on* $(0,T) \times M$.

Remark 13 *Another way of deriving directly equation* (18) *without relying on* (16) *is by computing the steepest descent of the total variation functional* $\int_M |\nabla u| \xi d\operatorname{Vol}_M$ *with respect to the metric*

$$(v, w) = \int_M \left(\frac{v}{|\nabla u|}\right)\left(\frac{w}{|\nabla u|}\right)|\nabla u|\xi d\operatorname{Vol}_M,$$

which is precisely the metric obtained by integrating the standard $L^2(\xi dS)$ *metric on normal velocities over all level sets of* u*. This can be made rigorous by using the co-area formula. Indeed, consider a smooth function* $u : (0,T) \times M \to \mathbf{R}$ *with* $Du \neq 0$*, we then compute*

$$\frac{d}{dt}\int_M |\nabla u(t,x)|\xi(x) d\operatorname{Vol}_M = \int_M g\left(\frac{\nabla u(x,t)}{|\nabla u(t,x)|}, \nabla\partial_t u(t,x)\right)\xi(x) d\operatorname{Vol}_M$$

$$= -\int_M \operatorname{div}\left(\xi \frac{\nabla u}{|\nabla u|}\right)(t,x)\partial_t u(t,x) d\operatorname{Vol}_M .$$

Thus the steepest descent of the total variation functional with respect to the metric defined above is given by requiring

$$\partial_t u = |\nabla u| \frac{1}{\xi} \operatorname{div}\left(\xi \frac{\nabla u}{|\nabla u|}\right),$$

which is equivalent to (18)*.*

We are now ready to introduce a weak solution concept for (18) based on the notion of a viscosity solution. This is a way of making sense of equation (18) when u is just continuous. The equation satisfies a comparison principle like the one we saw in the simple case of the heat equation (see Item (i) in Remark 5), and it thus allows to give an interpretation for (18) in a weak sense. To this aim, let us first observe that a classical solution to (18) is characterized by being both a subsolution and a supersolution (i.e. a smooth function satisfying equation (18) with $\leq$ and, respectively, $\geq$). Thus we need to interpret the notion of subsolution (respectively, supersolution) in a weak sense. We focus on the former. To do this, we want to be able to make sense of the pointwise inequality

$$\partial_t u \leq \left\langle g - \frac{Du \otimes Du}{|Du|^2}, D^2 u \right\rangle + g\left(\frac{\nabla \xi}{\xi}, \nabla u\right)$$

for an upper semi-continuous function $u : (0,T) \times M \to \mathbf{R}$, for which the required differential operators are not well defined. For this, we consider any point $(t_0, x_0) \in (0,T) \times M$ and observe that if $\varphi \in C^\infty((0,T) \times M)$ is such that φ touches the graph of u from above at (t_0, x_0), then $u - \varphi$ has there a local maximum (see Figure 2). Thus, if u was smooth, at (t_0, x_0) its gradient and its time derivative would coincide with the ones of φ and we would have an ordering for the Hessians. In other words, at (t_0, x_0), using also that u is a subsolution

$$\begin{aligned}\partial_t \varphi = \partial_t u &\leq \left\langle g - \frac{Du \otimes Du}{|Du|^2}, D^2 u \right\rangle + g\left(\frac{\nabla \xi}{\xi}, \nabla u\right) \\ &\leq \left\langle g - \frac{D\varphi \otimes D\varphi}{|D\varphi|^2}, D^2 \varphi \right\rangle + g\left(\frac{\nabla \xi}{\xi}, \nabla \varphi\right).\end{aligned}$$

Observe that to write down the inequality for φ we do not need smoothness of u, and thus this will become the basis for the definition of viscosity subsolution.

In the context of mean curvature flow with constant density $\xi = 1$ viscosity solutions were introduced in Evans and Spruck (1991) and Chen et al. (1991) in the Euclidean case, and in Ilmanen (1992) on curved manifolds. We now provide the rigorous definition we will be using. If $U \subset (0,T) \times M$ is an open set, $(t_0, x_0) \in U$ and if $u : (0,T) \times M \to \mathbf{R}$ is an upper (lower) semi-continuous function, a smooth function $\varphi : U \to \mathbf{R}$ is said to be tangent to u at (t_0, x_0) from above (below), if $u - \varphi$ has a local maximum (minimum) at (t_0, x_0).

Definition 14 *An upper (lower) semi-continuous function* $u : (0,T) \times M \to \mathbf{R}$ *is said to be a viscosity subsolution (supersolution) for* (18) *if for every* $(t_0, x_0) \in (0,T) \times M$ *and every smooth function* φ *tanget to* u *from above (below):*

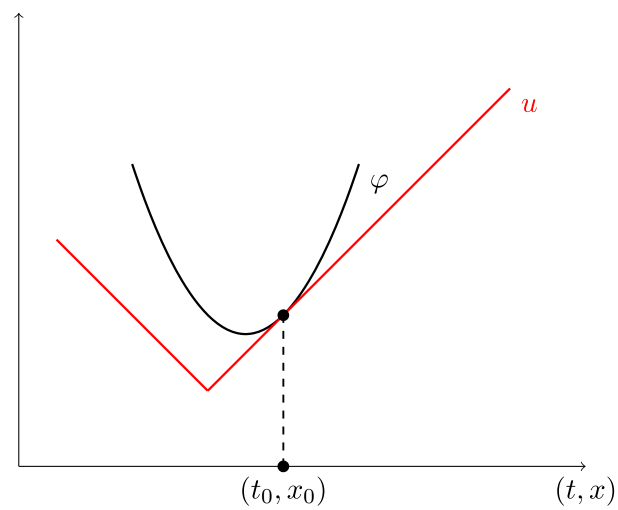

Figure 2: Illustration of the concept of viscosity subsolution. The graph of φ is touching the graph of u from above at (t_0, x_0).

(i) If $D\varphi(t_0, x_0) \neq 0$ then

$$\partial_t \varphi \leq \left\langle g - \frac{D\varphi \otimes D\varphi}{|D\varphi|^2}, D^2\varphi \right\rangle + g\left(\frac{\nabla \xi}{\xi}, \nabla \varphi\right) \ (\geq) \quad \text{at } (t_0, x_0)$$

(ii) Otherwise there exists $\nu \in T^*_{x_0}M$ with $|\nu| \leq 1$ such that

$$\partial_t \varphi \leq \langle g - \nu \otimes \nu, D^2\varphi \rangle \ (\geq) \quad \text{at } (t_0, x_0)$$

We say that u is a viscosity solution if it is both a subsolution and a supersolution.

In Ilmanen (1992) the author introduces the notion of viscosity subsolution/supersolution to mean curvature flow on a manifold (which corresponds to choosing the constant density $\xi = 1$) requiring continuity of the function u. We need to work with this slightly more general definition because the functions u_* and u^* in Theorem 4 are not continuous. We recall the following useful characterization of Definition 14, which says that we need to check condition (ii) only when also $D^2\varphi(t_0, x_0) = 0$.

Proposition 15 *Let $u : (0, T) \times M \to \mathbf{R}$ be an upper (lower) semicontinuous function. Then u is a viscosity subsolution (supersolution) of the level set formulation of MCF$_\xi$ if and only if whenever φ is tangent to u at (t_0, x_0) from above (below),* (i) *is satisfied and if $D\varphi(t_0, x_0) = 0$ and $D^2\varphi(t_0, x_0) = 0$, then*

$$\partial_t \varphi(t_0, x_0) \leq 0 \ (\geq).$$

Proposition 15 is proved in the Euclidean case in (Barles and Georgelin, 1995, Proposition 2.2). On a manifold, the proof is analogous and is therefore omitted. We recall the following comparison principle.

Theorem 16 *Let M be a closed k-dimensional Riemannian manifold. Let $\xi > 0$ be a smooth function on M. Let u be a viscosity subsolution of* (18) *on* $(0,T] \times M$ *and let v be a viscosity supersolution of* (18) *on* $(0,T] \times M$. *Define*

$$\overline{u}(x) := \limsup_{y \to x,\, t \to 0} u(t,y), \ \ \underline{v}(x) := \liminf_{y \to x,\, t \to 0} v(t,y).$$

Assume that $\overline{u} \leq \underline{v}$ and that either $\overline{u}$ or $\underline{v}$ is continuous. Then for every $t \in (0,T]$

$$u(t,\cdot) \leq v(t,\cdot).$$

Theorem 16 is proved when $\xi = 1$ is the constant density and the functions u, v are assumed to be continuous in Ilmanen (1992). A careful look at the proof of Ilmanen (1992) reveals that the same argument goes trough with the above assumptions. When $M = \mathbf{R}^k$ is the flat Euclidean space, an even more general version of Theorem 16 can be found in (Ambrosio and Dancer, 2000, Theorem 18). We also recall the following result concerning the existence of viscosity solutions, which can be again found in Ilmanen (1992) for the case of a constant density $\xi = 1$, and for which we thus omit the proof.

Theorem 17 *Let M be a k-dimensional closed Riemannian manifold, and let $\xi > 0$ be a smooth function on M. Let $u_0 : M \to \mathbf{R}$ be continuous. Then there exists a unique viscosity solution $u : [0,T) \times M \to \mathbf{R}$ to* (18) *such that $u(0) = u_0$.*

Finally, we recall the following *relabeling property*.

Lemma 18 *Let M be a k-dimensional closed Riemannian manifold, and let $\xi > 0$ be a smooth function on M. Let $u : [0,T) \times M \to \mathbf{R}$ be a viscosity solution to* (18). *Then for every continuous map $\Psi : \mathbf{R} \to \mathbf{R}$, the function $v := \Psi \circ u$ is a viscosity solution to* (18).

Lemma 18 is proved in Ilmanen (1992) in the case of a constant density $\xi = 1$, and we thus skip the proof.

Acknowledgments

This project has received funding from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC-2047/1 – 390685813.

References

L. Ambrosio and N. Dancer. *Calculus of variations and partial differential equations.* Springer-Verlag, Berlin, 2000. ISBN 3-540-64803-8. doi: 10.1007/978-3-642-57186-2. Topics on geometrical evolution problems and degree theory, Papers from the Summer School held in Pisa, September 1996, Edited by G. Buttazzo, A. Marino and M. K. V. Murthy.

G. Barles and C. Georgelin. A simple proof of convergence for an approximation scheme for computing motions by mean curvature. *SIAM J. Numer. Anal.*, 32(2):484–500, 1995. ISSN 0036-1429. doi: 10.1137/0732020.

P. Bérard, G. Besson, and S. Gallot. Embedding Riemannian manifolds by their heat kernel. *Geom. Funct. Anal.*, 4(4):373–398, 1994. ISSN 1016-443X. doi: 10.1007/BF01896401.

L. Bungert, J. Calder, and T. Roith. Uniform convergence rates for Lipschitz learning on graphs. *ArXiv preprint*, 2021. ArXiv:2111.12370.

J. Calder. Consistency of Lipschitz learning with infinite unlabeled data and finite labeled data. *SIAM J. Math. Data Sci.*, 1(4):780–812, 2019a. doi: 10.1137/18M1199241.

J. Calder. The game theoretic p-Laplacian and semi-supervised learning with few labels. *Nonlinearity*, 32(1):301–330, 2019b. ISSN 0951-7715. doi: 10.1088/1361-6544/aae949.

J. Calder and N. García Trillos. Improved spectral convergence rates for graph Laplacians on ε-graphs and k-NN graphs. *Appl. Comput. Harmon. Anal.*, 60:123–175, 2022. ISSN 1063-5203. doi: 10.1016/j.acha.2022.02.004.

J. Calder, B. Cook, M. Thorpe, and D. Slepcev. Poisson learning: Graph based semi-supervised learning at very low label rates. *Proceedings of the 37th International Conference on Machine Learning*, 119:1306–1316, 13–18 Jul 2020.

J. Calder, N. García Trillos, and M. Lewicka. Lipschitz regularity of graph Laplacians on random data clouds. *SIAM J. Math. Anal.*, 54(1):1169–1222, 2022. ISSN 0036-1410. doi: 10.1137/20M1356610.

I. Chavel. *Eigenvalues in Riemannian geometry*, volume 115 of *Pure and Applied Mathematics*. Academic Press, Inc., Orlando, FL, 1984. ISBN 0-12-170640-0.

Y. G. Chen, Y. Giga, and S. Goto. Uniqueness and existence of viscosity solutions of generalized mean curvature flow equations. *J. Differential Geom.*, 33(3):749–786, 1991. ISSN 0022-040X.

M. G. Crandall, H. Ishii, and P.-L. Lions. User's guide to viscosity solutions of second order partial differential equations. *Bull. Amer. Math. Soc. (N.S.)*, 27(1):1–67, 1992. ISSN 0273-0979. doi: 10.1090/S0273-0979-1992-00266-5.

D. B. Dunson, H.-T. Wu, and N. Wu. Spectral convergence of graph Laplacian and heat kernel reconstruction in L^∞ from random samples. *Appl. Comput. Harmon. Anal.*, 55: 282–336, 2021. ISSN 1063-5203. doi: 10.1016/j.acha.2021.06.002.

S. Esedoḡlu and F. Otto. Threshold dynamics for networks with arbitrary surface tensions. *Comm. Pure Appl. Math.*, 68(5):808–864, 2015. ISSN 0010-3640. doi: 10.1002/cpa.21527.

L. C. Evans and J. Spruck. Motion of level sets by mean curvature. I. *J. Differential Geom.*, 33(3):635–681, 1991. ISSN 0022-040X.

J. Fuchs and T. Laux. Strong convergence of the thresholding scheme for the mean curvature flow of mean convex sets. *ArXiv Preprint*, 2022. ArXiv:2201.00413.

C. Garcia-Cardona, E. Merkurjev, A. L. Bertozzi, A. Flenner, and A. G. Percus. Multiclass data segmentation using diffuse interface methods on graphs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 36(8):1600–1613, 2014. doi: 10.1109/TPAMI.2014.2300478.

N. García Trillos, M. Gerlach, M. Hein, and D. Slepčev. Error estimates for spectral convergence of the graph Laplacian on random geometric graphs toward the Laplace-Beltrami operator. *Found. Comput. Math.*, 20(4):827–887, 2020. ISSN 1615-3375. doi: 10.1007/s10208-019-09436-w.

T. Ilmanen. Generalized flow of sets by mean curvature on a manifold. *Indiana Univ. Math. J.*, 41(3):671–705, 1992. ISSN 0022-2518. doi: 10.1512/iumj.1992.41.41036.

H. Ishii, G. E. Pires, and P. E. Souganidis. Threshold dynamics type approximation schemes for propagating fronts. *J. Math. Soc. Japan*, 51(2):267–308, 1999. ISSN 0025-5645. doi: 10.2969/jmsj/05120267.

M. Jacobs. A fast MBO scheme for multiclass data classification. In *Scale Space and Variational Methods in Computer Vision*, pages 335–347. Springer International Publishing, 2017. ISBN 978-3-319-58771-4.

M. Jacobs, E. Merkurjev, and S. Esedoḡlu. Auction dynamics: a volume constrained MBO scheme. *J. Comput. Phys.*, 354:288–310, 2018a. ISSN 0021-9991. doi: 10.1016/j.jcp.2017.10.036.

M. Jacobs, E. Merkurjev, and S. Esedoḡlu. Auction dynamics: a volume constrained MBO scheme. *J. Comput. Phys.*, 354:288–310, 2018b. ISSN 0021-9991. doi: 10.1016/j.jcp.2017.10.036. URL `https://doi.org/10.1016/j.jcp.2017.10.036`.

I. Kim and D. Kwon. Volume preserving mean curvature flow for star-shaped sets. *Calc. Var. Partial Differential Equations*, 59(2):Paper No. 81, 40, 2020. ISSN 0944-2669. doi: 10.1007/s00526-020-01738-0.

T. Laux and J. Lelmi. Large data limit of the MBO scheme for data clustering: Γ-convergence of the thresholding energies. *Arxiv preprint*, 2021. ArXiv:2112.06737.

C. Mantegazza. *Definition and Short Time Existence*, pages 1–23. Springer Basel, Basel, 2011. ISBN 978-3-0348-0145-4. doi: 10.1007/978-3-0348-0145-4_1. URL `https://doi.org/10.1007/978-3-0348-0145-4_1`.

P. Mascarenhas. Diffusion generated motion by mean curvature. *CAM Reports, Department of Mathematics, University of California, Los Angeles*, 1992.

E. Merkurjev, T. Kostić, and A. L. Bertozzi. An MBO scheme on graphs for classification and image processing. *SIAM J. on Imaging Sci.*, 6(4):1903–1930, 2013. doi: 10.1137/120886935.

E. Merkurjev, C. Garcia-Cardona, A. L. Bertozzi, A. Flenner, and A. G. Percus. Diffuse interface methods for multiclass segmentation of high-dimensional data. *Appl. Math. Lett.*, 33:29–34, 2014. ISSN 0893-9659. doi: 10.1016/j.aml.2014.02.008.

B. Merriman, J. K. Bence, and S. J. Osher. Diffusion generated motion by mean curvature. *CAM Reports, Department of Mathematics, University of California, Los Angeles*, 1992.

B. Merriman, J. K. Bence, and S. J. Osher. Motion of multiple functions: a level set approach. *J. Comput. Phys.*, 112(2):334–363, 1994. ISSN 0021-9991. doi: 10.1006/jcph.1994.1105.

O. Misiats and N. K. Yip. Convergence of space-time discrete threshold dynamics to anisotropic motion by mean curvature. *Discrete Contin. Dyn. Syst.*, 36(11):6379–6411, 2016. ISSN 1078-0947. doi: 10.3934/dcds.2016076.

S. Rosenberg. *The Laplacian on a Riemannian manifold*, volume 31 of *London Math. Soc. Stud. Texts*. Cambridge University Press, Cambridge, 1997. ISBN 0-521-46300-9; 0-521-46831-0. doi: 10.1017/CBO9780511623783.

A. G. Setti. Gaussian estimates for the heat kernel of the weighted Laplacian and fractal measures. *Canad. J. Math.*, 44(5):1061–1078, 1992. ISSN 0008-414X. doi: 10.4153/CJM-1992-065-4.

Y. van Gennip, N. Guillen, B. Osting, and A. L. Bertozzi. Mean curvature, threshold dynamics, and phase field theory on finite graphs. *Milan J. Math.*, 82(1):3–65, 2014. ISSN 1424-9286. doi: 10.1007/s00032-014-0216-8.

Appendix A. MBO scheme on manifolds

As in the previous section, M will denote a k-dimensional closed Riemannian manifold and $\xi > 0$ will denote a smooth function on M. The following algorithm can be used to approximate the evolution of an open set $\Omega_0 \subset M$ with smooth boundary by MCF_ξ. The results contained in the present section, i.e. Theorem 20, Corollary 21, and Theorem 22 are proved in Section B.3.

Algorithm 19 (MBO scheme on manifolds) *Fix a time-step size $h > 0$, a diffusion coefficient $\kappa > 0$ and a (bounded) drift $f : M \to \mathbf{R}$. Let $\Omega_0 \subset M$ be an open set with a smooth boundary. For each $n \in \mathbf{N}$ define inductively Ω_{l+1} as follows.*

1. ***Diffusion**. Define*
$$u_l := e^{-h\kappa\Delta_\xi}\mathbf{1}_{\Omega_l}.$$
2. ***Thresholding**. Define Ω_{n+1} by*
$$\Omega_{l+1} = \left\{u_l \geq \frac{1}{2} + f\sqrt{h}\right\}.$$

Before stating the next result, let us introduce some notation. Given a set $E \subset M$, we denote by $\mathrm{diam}(E)$ its diameter, i.e. the supremum of $d_M(x, y)$ taken over pair of points $x, y \in E$. We denote by $\mathrm{inj}(M)$ the injectivity radius of the manifold M, which is defined as the quantity such that for any $r < \mathrm{inj}(M)$ and any $x \in M$, the exponential map is a diffeomorphism onto the open ball $B_r(x)$. We then have the following result for one step of MBO.

Theorem 20 *Let M, ξ be as above. Let Ω_0 be a smooth open set such that $\mathrm{diam}(\Omega_0) < \frac{\mathrm{inj}(M)}{2}$. Let Ω_1 be obtained by applying one step of MBO with a bounded drift $f : M \to \mathbf{R}$ to Ω_0 with a given step size $h > 0$ and a given diffusion coefficient $\kappa > 0$. Let $x \in \partial\Omega_0$. Let $\nu(x) \in T_xM$ be the outer unit normal to $\partial\Omega_0$ at x and define*

$$z(x) := \begin{cases} \sup\{s \in \mathbf{R}^- |\ \exp_x(s\nu(x)) \in \Omega_1\} & \text{if } x \notin \Omega_1, \\ \inf\ \{s \in \mathbf{R}^+ |\ \exp_x(s\nu(x)) \notin \Omega_1\} & \text{if } x \in \Omega_1. \end{cases}$$

Then we have
$$|z(x)| \leq Vh,$$
where the constant V depends only on κ, the L^∞-norm of f, the ambient manifold M, and the C^0-norm of the second fundamental form of $\partial\Omega_0$.

We will apply Theorem 20 in the special case where $\Omega_0 = B_r(x_0)$ is a geodesic ball in the Riemannian manifold M. In this case, Theorem 20 specialized to the following corollary.

Corollary 21 *Let $x_0 \in M$ and $R < \frac{\mathrm{inj}(M)}{4}$ be fixed. Then there is a constant $C_R < +\infty$ such that if $\frac{R}{2} < r \leq R$ and, in the above theorem, $\Omega_0 = B_r(x_0)$, then*
$$|z(x)| \leq C_R h$$
for every $x \in \partial B_r(x_0)$.

Finally, we have the following consistency result, which will be crucial in proving Theorem 4. Hereafter, $C^{1,2}((0,+\infty)\times M)$ is defined as the space of functions on $(0,+\infty)\times M$ which are C^1 in the first variable, and C^2 in the second one. The space $C^{1,2}((0,+\infty)\times M)$ is endowed with a norm, and thus a notion of convergence, given by

$$\begin{aligned}\|f\|_{C^{1,2}((0,+\infty)\times M)} =& \|f\|_{C^0((0,+\infty)\times M)} + \|\partial_t f\|_{C^0((0,+\infty)\times M)} \\ &+ \|Df\|_{C^0((0,+\infty)\times M)} + \|D^2 f\|_{C^0((0,+\infty)\times M)}.\end{aligned}$$

Theorem 22 *Let h_n be a sequence of positive real numbers converging to zero. Assume that $\varphi_{h_n}:(0,+\infty)\times M\to\mathbf{R}$ are $C^{1,2}((0,+\infty)\times M)$ functions converging in $C^{1,2}((0,+\infty)\times M)$ to a function $\varphi:(0,+\infty)\times M\to\mathbf{R}$. Assume that $(s_{h_n},z_{h_n})\in(0,+\infty)\times M$ are converging to a point $(s,z)\in[0,+\infty)\times M$. Assume also that $\delta_n:=\varphi_{h_n}(s_{h_n},z_{h_n})$ are such that*

$$\lim_{n\to+\infty}\frac{\delta_n}{\sqrt{h_n}}=0. \tag{19}$$

Then we have that:

(i) If $D\varphi(s,z)\neq 0$ then

$$\begin{aligned}&\liminf_{n\to+\infty}\frac{1}{\sqrt{\kappa h_n}}\left(\frac12-\int_{\{\varphi_{h_n}(s_{h_n}-h_n,\cdot)\geq 0\}}H(\kappa h_n,z_{h_n},y)\xi(y)d\operatorname{Vol}_M\right)\\ &\quad\geq\frac{1}{2\sqrt{\pi}|D\varphi(s,z)|}\left(\partial_t\varphi-\left\langle g-\frac{D\varphi\otimes D\varphi}{|D\varphi|^2},D^2\varphi\right\rangle-g\left(\frac{\nabla\xi}{\xi},\nabla\varphi\right)\right)(s,z).\end{aligned} \tag{20}$$

(ii) Otherwise if $D\varphi(s,z)=0, D^2\varphi(s,z)=0$ and

$$\frac12-\int_{\{\varphi_{h_n}(s_{h_n}-h_n,\cdot)\geq 0\}}H(\kappa h_n,z_{h_n},y)\xi(y)d\operatorname{Vol}_M\leq o(\sqrt{h_n}),$$

then

$$\partial_t\varphi(s,z)\leq 0.$$

Appendix B. Proofs

B.1 Conditional convergence: Proof of Theorem 4

The purpose of this section is the proof of Theorem 4, which is inspired by the works Barles and Georgelin (1995) and Misiats and Yip (2016).

Proof of Theorem 4. Let u be the unique viscosity solution to MCF_ξ from Theorem 17 with $\xi=\rho^2$, starting from $u(0,\cdot)=sd(\cdot,\Gamma_0):=d_M(x,\Omega_0^c)-d_M(x,\Omega_0)$. We will show later that u^* and u_* are, respectively, a viscosity subsolution and a viscosity supersolution of the level set formulation of MCF_ξ according to Definition 14. We furthermore claim that for every $x\in M$,

$$u^*(0,x)\leq\operatorname{sign}^*(u(0,x)), \tag{21}$$

$$u_*(0,x) \geq \operatorname{sign}_*(u(0,x)), \tag{22}$$

where sign^* and sign_* are, respectively, the upper semi-continuous envelope and the lower semi-continuous envelope of the sign function.

Once these facts are proved, it follows from Theorem 16 that for every $x \in M$ and every $t \geq 0$,

$$u^*(t,x) \leq \operatorname{sign}^*(u(t,x)), \tag{23}$$
$$u_*(t,x) \geq \operatorname{sign}_*(u(t,x)). \tag{24}$$

To see this, we observe that if $\Psi : \mathbf{R} \to \mathbf{R}$ is a continuous function such that $\Psi \geq \operatorname{sign}^*$, then the relabeling property in Lemma 18 implies that $\Psi \circ u$ is a continuous solution to (18) with $u^*(0,x) \leq \operatorname{sign}^*(u(0,x)) \leq \Psi(u(0,x))$ for every $x \in M$, thus Theorem 16 implies that for every $0 \leq t \leq T$ and every $x \in M$

$$u^*(t,x) \leq \inf_{\Psi \text{ continuous}, \Psi \geq \operatorname{sign}^*} \Psi(u(t,x)) = \operatorname{sign}^*(u(t,x)).$$

A similar argument gives (24). Let us now conclude the proof of the theorem assuming that (23) and (24) hold. If $x \in \Omega_t$, then $u(t,x) > 0$, thus (24) yields $u_*(t,x) = 1$. In a similar way (23) implies that $u^*(t,x) = -1$ on $(\Omega_t \cup \Gamma_t)^c$. We are thus left with proving that u^* is a subsolution, that u_* is a supersolution and with verifying the initial conditions (21) and (22).

We now show that indeed u^* is a viscosity subsolution. Pick a test functions φ tangent to u^* at $(t_0, x_0) \in (0,+\infty) \times M$ from above. We may assume without loss of generality that

$$\lim_{t \to +\infty} \max_M \varphi(t,\cdot) = +\infty, \tag{25}$$

and that $u^* - \varphi$ has a strict global maximum at (t_0, x_0). Thanks to Proposition 15, we only need to check that

1. Either $D\varphi(t_0,x_0) \neq 0$ and
$$\partial_t \varphi \leq \left\langle g - \frac{D\varphi \otimes D\varphi}{|D\varphi|^2}, D^2\varphi \right\rangle + g\left(\frac{\nabla \xi}{\xi}, \nabla \varphi\right) \text{ at } (t_0,x_0).$$

2. Or $D\varphi(t_0,x_0) = 0$, $D^2\varphi(t_0,x_0) = 0$ and
$$\partial_t \varphi(t_0,x_0) \leq 0.$$

If $(t_0,x_0) \in \{u^* = -1\}$ or $(t_0,x_0) \in \operatorname{Int}\{u^* = 1\}$ the claim is trivial, because in that case u^* is constant in a neighborhood of (t_0,x_0). We thus assume that $(t_0,x_0) \in \partial\{u^* = 1\}$. By definition, there exists a sequence (t_{n_j}, z_{n_j}) such that $z_{n_j} \in G_{n_j}$ for every $j \in \mathbf{N}$ and, as $j \to +\infty$,

$$n_j \to +\infty,$$
$$z_{n_j} \to x_0,$$

$$
\begin{aligned}
&t_{n_j} \to t_0,\\
&u^{n_j,G_{n_j}}(t_{n_j},z_{n_j}) \to u^*(t_0,x_0).
\end{aligned}
$$

For every $j \in \mathbf{N}$, pick

$$(s_j,x_j) \in \operatorname{argmax}_{x\in G_{n_j}, s\in(0,+\infty)} \left\{ u^{n_j,G_{n_j}}(s,x) - \varphi(s,x) \right\}. \tag{26}$$

We observe that, up to extracting a subsequence, $(s_j,x_j) \to (t_0,x_0)$ as $j \to +\infty$. Indeed by the compactness of M and the assumption (25), we may assume that the sequence (s_j,x_j) converges to some limit point $(\underline{s},\underline{x})$. Then by definition of u^*, by the choice (26) and by the properties of the points (t_{n_j},z_{n_j}) we must have

$$
\begin{aligned}
(u^*-\varphi)(\underline{s},\underline{x}) &\geq \limsup_{j\to+\infty} \, (u^{n_j,G_{n_j}} - \varphi)(s_j,x_j)\\
&\geq \limsup_{j\to+\infty} \, (u^{n_j,G_{n_j}} - \varphi)(t_{n_j},z_{n_j})\\
&= (u^*-\varphi)(t_0,x_0).
\end{aligned}
$$

This forces $(t_0,x_0) = (\underline{s},\underline{x})$, because (t_0,x_0) is a strict global maximum for $u^*-\varphi$. It is also easy to check that $u^{n_j,G_{n_j}}(s_j,x_j) = 1$ for j large enough. We now pick a sequence $\delta_j \downarrow 0$ to be determined later, and we define $\theta_j : \mathbf{R} \to [-1,1]$ to be a smooth function such that

$$
\begin{aligned}
&\theta_j(t) = \operatorname{sign}(t) \text{ for } |t| \geq \delta_j,\\
&\|\theta_j'\|_\infty \leq \frac{2}{\delta_j}.
\end{aligned}
$$

We claim that

$$u^{n_j,G_{n_j}}(s,z) \leq \theta_j(\varphi(s,z) - \varphi(s_j,x_j) + \delta_j) \tag{27}$$

for every j large enough, $z \in G_{n_j}$ and $s \in (0,+\infty)$. Indeed, inequality (27) holds trivially if $u^{n_j,G_{n_j}}(s,z) = -1$. If instead $u^{n_j,G_{n_j}}(s,z) = 1$, probing (26) with (s,z), we have

$$
\begin{aligned}
1 = u^{n_j,G_{n_j}}(s,z) &\leq u^{n_j,G_{n_j}}(s_j,x_j) - \varphi(s_j,x_j) + \varphi(s,z)\\
&= 1 - \varphi(s_j,x_j) + \varphi(s,z),
\end{aligned}
$$

where we used that $u^{n_j,G_{n_j}}(s_j,x_j) = 1$ for j large enough. In particular

$$0 \leq -\varphi(s_j,x_j) + \varphi(s,z),$$

which, by definition of θ_j, yields (27).

We now choose $s = s_j - h_{n_j}$ in (27), we apply $S_{n_j}(h_{n_j},\cdot)$ to both sides of the inequality and we evaluate at x_j. Recalling assumption (i) of Theorem 4 we get

$$
\begin{aligned}
&S_{n_j}(h_{n_j}, u^{n_j,G_{n_j}}(s_j - h_{n_j},\cdot))(x_j)\\
&\leq S_{n_j}(h_{n_j}, \theta_j(\varphi(s_j - h_{n_j},\cdot) - \varphi(s_j,x_j) + \delta_j))(x_j) + O\left(h_{n_j}^{3/2}\right).
\end{aligned}
$$

We now apply sign* to both sides of the inequality to get

$$1 = u^{n_j,G_{n_j}}(s_j,x_j) \leq \operatorname{sign}^*\left(S_{n_j}(h_{n_j},\theta_j(\varphi(s_j-h_{n_j},\cdot)-\varphi(s_j,x_j)+\delta_j))(x_j)+O\left(h_{n_j}^{3/2}\right)\right),$$

which, by definition of the function sign*, implies

$$0 \leq S_{n_j}(h_{n_j},\theta_j(\varphi(s_j-h_{n_j},\cdot)-\varphi(s_j,x_j)+\delta_j))(x_j)+O\left(h_{n_j}^{3/2}\right).$$

We now divide both sides of the previous inequality by 2 and we add $1/2$ to both sides of the inequality. Using assumption (iii) of Theorem 4 and the linearity of S_n in the second variable yields

$$\frac{1}{2} \leq S_{n_j}\left(h_{n_j},\left(\frac{1+\theta_j}{2}\right)\left(\varphi(s_j-h_{n_j},\cdot)-\varphi(s_j,x_j)+\delta_j\right)\right)(x_j)+O\left(h_{n_j}^{3/2}\right).$$

Define

$$f_j(z) := \left(\frac{1+\theta_j}{2}\right)\left(\varphi(s_j-h_{n_j},z)-\varphi(s_j,x_j)+\delta_j\right).$$

Then by applying the estimate (5) in assumption (ii) in Theorem 4 we obtain

$$\frac{1}{2} \leq (e^{-h_{n_j}\kappa\Delta_\xi}f_j)(x_j)+o(h_{n_j}^{1/2})+\frac{2}{\delta_j}O(h_{n_j}^{3/2}).$$

In other words, we have

$$\begin{aligned} o\left(h_{n_j}^{1/2}\right)+\frac{2}{\delta_j}O\left(h_{n_j}^{3/2}\right) &\geq \frac{1}{2}-\int_M H(h_{n_j}\kappa,x_j,y)f_j(y)\xi(y)d\operatorname{Vol}_M(y) \\ &\geq \frac{1}{2}-\int_{\{\varphi(s_j-h_{n_j},\cdot)-\varphi(s_j,x_j)+\delta_j\geq 0\}} H(h_{n_j}\kappa,x_j,y)\xi(y)d\operatorname{Vol}_M(y). \end{aligned}$$

We divide the previous inequality by $\sqrt{h_{n_j}\kappa}$, and we choose $\delta_j = h_{n_j}^{2/3}$ so that on the one hand $\frac{h_{n_j}}{\delta_j} \to 0$ and on the other hand we can apply Theorem 22. If $D\varphi(t_0,x_0) \neq 0$, then by (i) in Theorem 22,

$$0 \geq \frac{1}{2\sqrt{\pi}|D\varphi(s,z)|}\left(\partial_t\varphi-\left\langle g-\frac{D\varphi\otimes D\varphi}{|D\varphi|^2},D^2\varphi\right\rangle-g\left(\frac{D\xi}{\xi},D\varphi\right)\right)(t_0,x_0),$$

which gives (i) in Definition 14. If $D\varphi(t_0,x_0)=0$ and $D^2\varphi(t_0,x_0)=0$ then we can apply (ii) in Theorem 22 to get the second item in the equivalent description of viscosity subsolution in Proposition 15. Thus u^* is a viscosity subsolution. In a similar way one can prove that u_* is a supersolution.

We are left with checking the initial conditions for u^* and u_*. Again, we focus on the inequality (21) for u^*, since the argument for u_* is similar. Observe that

$$\operatorname{sign}^*(u(0,x)) = \begin{cases} 1 & \text{if } x \in \overline{\Omega_0} \\ -1 & \text{if } x \in M \setminus \overline{\Omega_0} \end{cases}$$

and since $u^* \in \{-1,1\}$, we just have to show that $u^*(0,x)=-1$ for $x \in M\setminus\overline{\Omega_0}$. To this aim, pick a sequence $(t_n,z_n)\in(0,+\infty)\times G_n$ such that $t_n\to 0$ and $z_n\to x$ as $n\to+\infty$. We have to show that $u^{n,G_n}(t_n,z_n)=-1$ for n large enough. For $q\in\mathbf{R}$, denote by $T^{q,G_n}(h_n)(\Omega_0)$ the outcome of the abstract thresholding scheme with thresholding value given by q and step size h_n on the graph G_n with initial value $\Omega_0\cap V_n$. For $m\in\mathbf{N}$ we also write $(T^{q,G_n}(h_n))^m$ for $T^{q,G_n}(h_n)\circ ... \circ T^{q,G_n}(h_n)$. Since $x\in M\setminus\overline{\Omega_0}$ there exists $R>0$ such that $B_R(x)\subset M\setminus\overline{\Omega_0}$. We denote by $w_n : V_n\to[0,+\infty)$ a sequence of non-negative functions which, for n large enough and for every $u,v\in\mathcal{V}_n, |u|\le 1, |v|\le 1$, satisfy

$$u\le v \Rightarrow S(h_n,u)\le S(h_n,v)+w_n, \tag{28}$$

$$\begin{aligned} &a_n := \|w_n\|_{L^\infty(G_n)} = O(h_n^{3/2}),\\ &\max_{x\in V_n}|S(h_n,\mathbf{1}_{G_n})(x)-1| < a_n. \end{aligned} \tag{29}$$

Such functions exist by assumptions (i) and (iii) in Theorem 4. We now check that

$$V_n\setminus\left(T^{1/2,G_n}(h_n)\right)^m(\Omega_0)\supset\left(T^{1/2+2ma_n,G_n}(h_n)\right)^m(B_R(x)). \tag{30}$$

To see this, we proceed by induction over m. We treat just the base case $m=1$, the inductive step being analogous. To prove (30) for $m=1$, we show

$$V_n\setminus T^{1/2,G_n}(h_n)(\Omega_0)\supset T^{1/2+a_n,G_n}(h_n)(M\setminus\Omega_0)\supset T^{1/2+2a_n,G_n}(h_n)(B_R(x)). \tag{31}$$

To see this, let $y\in T^{1/2+a_n,G_n}(h_n)(M\setminus\Omega_0)$, observe that by (29) we have

$$S(h_n,\mathbf{1}_{\Omega_0})(y)+\frac{1}{2}+a_n\le S(h_n,\mathbf{1}_{\Omega_0})(y)+S_n(h_n,\mathbf{1}_{M\setminus\Omega_0})(y)<1+a_n,$$

in particular, we have that $y\in V_n\setminus T^{1/2,G_n}(h_n)(\Omega_0)$. Thus $V_n\setminus T^{1/2,G_n}(h_n)(\Omega_0)\supset T^{1/2+a_n,G_n}(M\setminus\Omega_0)$. We now observe that since $\mathbf{1}_{B_R(x)}\le\mathbf{1}_{M\setminus\Omega_0}$, (28) yields that for $y\in T^{1/2+2a_n,G_n}(h_n)(B_R(x))$

$$\frac{1}{2}+2a_n\le S(h_n,\mathbf{1}_{B_R(x)})(y)\le S(h_n,\mathbf{1}_{M\setminus\Omega_0})(y)+a_n,$$

which yields (31).

We will show that there is a constant $C<+\infty$ such that

$$\left(T^{1/2+2\left[\frac{t_n}{h_n}\right]a_n,G_n}(h_n)\right)^{\left[\frac{t_n}{h_n}\right]}(B_R(x))\supset B_{R-Ct_n}(x)\cap V_n. \tag{32}$$

Once this is proved, we have that using also (30), since $t_n\downarrow 0$,

$$M\setminus\left(T^{1/2,G_n}(h_n)\right)^{\left[\frac{t_n}{h_n}\right]}(\Omega_0)\supset B_{\frac{R}{2}}(x)$$

when n is large enough. In particular, since z_n is converging to x, we must have that $u^{n,G_n}(t_n,z_n)=-1$ for n large enough. Finally, to show (32) we argue as follows. Let C_R

be the constant in Corollary 21. Let $f \in C_c^\infty(B_R(x))$ such that $\mathbf{1}_{B_{R-C_R h_n}(x)} \leq f \leq \mathbf{1}_{B_R(x)}$ with $\mathrm{Lip}(f) \leq c/h_n$, using assumptions (i) and (ii) in Theorem 4 we have for $y \in M \cap V_n$

$$\begin{aligned} S_n(h_n, \mathbf{1}_{B_R(x)})(y) &\geq S_n(h_n, f)(y) + O(h_n^{3/2}) \\ &\geq e^{-h_n \kappa \Delta_\xi} f(y) + O(h_n^{1/2}) \\ &\geq e^{-h_n \kappa \Delta_\xi} \mathbf{1}_{B_{R-C_R h_n}(x)}(y) + O(h_n^{1/2}). \end{aligned}$$

Observe that $\frac{1}{2} + 2\left[\frac{t_n}{h_n}\right] a_n = \frac{1}{2} + O(h_n^{1/2})$, in particular, we can apply Corollary 21 to obtain, for n large enough, whenever $y \in B_{R-2C_R h_n}(x) \cap V_n$

$$e^{-h_n \kappa \Delta_\xi} \mathbf{1}_{B_{R-C_R h_n}(x)}(y) + O(h_n^{1/2}) \geq \frac{1}{2} + 2\left[\frac{t_n}{h_n}\right] a_n.$$

By an induction argument we get (32). ■

B.2 Heat kernel estimate in random geometric graphs

The main purpose of this subsection is the proof of Theorem 6. We first introduce some notation. We denote by $\{\lambda_l\}_{l=1}^{+\infty}$ the eigenvalues of the weighted Laplacian Δ_{ρ^2} on the manifold (M, g), which are ordered in the following way (recall that we are assuming that the eigenvalues are simple)

$$0 = \lambda_1 < \lambda_2 < \lambda_3 < \ldots$$

We denote by $\{f_l\}_{l=1}^{+\infty}$ an orthonormal basis (with respect to the $L^2(\rho^2 \,\mathrm{Vol}_M)$-inner product on M) made of the corresponding eigenvectors. Then, for $x, y \in M$, the heat kernel on M can be written as

$$H(t, x, y) = \sum_{l=1}^{+\infty} e^{-t\lambda_l} f_l(x) f_l(y). \tag{33}$$

In the proof of Theorem 6 we will assume, for simplicity, that $K_n = \log(n)^q \in \mathbf{N}$. In this setting we will use condition (v) of Theorem 6 in the form

$$\epsilon_n \ll \frac{\sqrt{\log(n)}}{K_n^{1+\frac{1}{k}-\frac{s}{2}} \left(\lambda_{K_n}^{\frac{k}{2}+1} + 1\right)^2 \left(\lambda_{K_n}^{4+\frac{k}{2}} + 1\right)}. \tag{34}$$

Observe that condition (v) of Theorem 6 implies (34) because by Weyl's law we have $\lambda_{K_n} \sim K_n^{2/k}$.

Proof of Theorem 6. As we pointed out in Remark 7, in the present proof we will for simplicity assume that $K_n = \log(n)^q \in \mathbf{N}$. We will indicate by γ the quantity $\gamma := \inf_{i \in \mathbf{N}}(\lambda_{i+1} - \lambda_i)$, which is positive by Item (ii) in Theorem 6.

Observe that items (i) and (iii) in Theorem 4 hold exactly (i.e. without error) for the choice $S_n(t, \cdot) = e^{-t\Delta_n}(\cdot)$. To show that these hold true with high probability also for the choice $S_n = P_n$ defined in (13) we take $w \in \mathcal{V}_n$ and we consider, for $x \in V_n$, the difference

$$\left| e^{-h_n \Delta_n} w(x) - P_n(h_n, w)(x) \right| = \left| \sum_{y \in V_n} \sum_{l=K_n+1}^{n} e^{-h_n \lambda_n^l} v_n^l(x) v_n^l(y) \frac{d_n(y)}{n} w(y) \right|$$

$$\leq n \max_{z\in V_n} |w(z)| \max_{z\in V_n} |d_n(z)| \frac{1}{n} \max_{z\in V_n} \sum_{l=K_n+1}^{n} e^{-h_n\lambda_n^l}(v_n^l(z))^2,$$

where in the last line we used the Cauchy–Schwarz inequality. To get items (i) and (iii) in Theorem 6 for P_n, it thus suffices to show that

$$\mathcal{R}_n := \max_{z\in V_n} d_n(z) \max_{z\in V_n} \frac{1}{n} \sum_{l=K_n+1}^{n} e^{-h_n\lambda_n^l}(v_n^l(z))^2 = O\left(\frac{h_n^{3/2}}{n}\right).$$

To show this, we start by observing that for every $n \in \mathbf{N}$, every $z \in V_n$ and $1 \leq l \leq n$

$$1 = \langle v_n^l, v_n^l \rangle_{\mathcal{V}_n} \geq \frac{d_n(z)}{n}(v_n^l(z))^2. \tag{35}$$

By applying Theorem 28 we can also choose n so large that, with probability greater than $1 - Q_6\epsilon_n^{-k}\exp(-Q_7 n\epsilon_n^{k+2})$, we have

$$\max_{z\in V_n} |d_n(z) - C_1\rho(z)| \leq Q_8\epsilon_n,$$

and we can clearly assume that n is so large that

$$C_1\frac{\min\rho}{2} \leq d_n \leq 2C_1 \max\rho.$$

Using (35) and the ordering $\lambda_n^l \geq \lambda_n^{K_n}$ for $n \geq l \geq K_n$ we get

$$\begin{aligned}\mathcal{R}_n &\leq \frac{C}{n}\left(n^2 e^{-\lambda_n^{K_n}h_n}\right)\\ &= \frac{C}{n}\left(n^2 e^{-\kappa(\eta)\lambda_{K_n}h_n} e^{-\left(\lambda_n^{K_n}-\kappa(\eta)\lambda_{K_n}\right)h_n}\right).\end{aligned}$$

We now use Theorem 27 and Theorem 24 to infer that with probability greater than $1 - Q_1\epsilon_n^{-6k}\exp(-Q_2 n\epsilon_n^{k+4}) - Q_3 n\exp(-Q_4 n\,(\lambda_{K_n}+1)^{-k})$ we have

$$\mathcal{R}_n \leq \frac{C}{n}\left(n^2 e^{-\kappa(\eta)\lambda_{K_n}h_n} e^{\frac{C\epsilon_n}{\gamma}\left(\lambda_{K_n}^{4+\frac{k}{2}}+1\right)h_n}\right).$$

By Weyl's law we have that $\lambda_{K_n} \sim K_n^{2/k}$, thus

$$\mathcal{R}_n \leq \frac{C}{n}\left(n^2 e^{-cK_n^{2/k}h_n} e^{\frac{\tilde{C}\epsilon_n}{\gamma}K_n^{\frac{8}{k}+1}}\right).$$

Recalling the conditions (iv), (v) and (ii) in Theorem 6, as well as the scaling $K_n = (\log(n))^q$ we get

$$\mathcal{R}_n \leq \frac{C}{n}\left(n^2 e^{-c(\log(n))^{\frac{2q}{k}-\alpha}}\right)$$

$$= \frac{Ch_n^{3/2}}{n} \left(\frac{n^{2-c(\log(n))^{\frac{2q}{k}-1-\alpha}}}{h_n^{3/2}} \right)$$
$$\leq \frac{Ch_n^{3/2}}{n} \left(n^{2-c(\log(n))^{\frac{2q}{k}-1-\alpha}} (\log(n))^{\frac{3\alpha}{2}} \right).$$

So $\mathcal{R}_n = O\left(\frac{h_n^{3/2}}{n}\right)$ because by the definition of α in (iv) in Theorem 6 we have $\frac{2q}{k} - 1 - \alpha > 0$. We are left with proving item (ii) in Theorem 4 for both $e^{-t\Delta_n}(\cdot)$ and P_n. We prove it for $e^{-t\Delta_n}(\cdot)$, the proof for P_n being analogous. The proof is divided into three steps.

Step 1. We claim that with probability greater than $1 - a_1\epsilon_n^{-6k} \exp(-a_2 n\epsilon_n^{k+4}) - a_3 n \exp(-a_4 n (\lambda_{K_n} + 1)^{-k})$

$$\max_{x,y\in V_n} \left| H_{\epsilon_n}^n(h_n, x, y) - \frac{\rho(y)}{n} H(\kappa(\eta)h_n, x, y) \right| = o\left(\frac{\sqrt{h_n}}{n}\right). \tag{36}$$

To show (36) we pick two points $x, y \in V_n$ and we compute

$$\left| H_{\epsilon_n}^n(h_n, x, y) - \frac{\rho(y)}{n} H(\kappa(\eta)h_n, x, y) \right| \leq \left| H_{\epsilon_n}^{K_n}(h_n, x, y) - \frac{\rho(y)}{n} H(\kappa(\eta)h_n, x, y) \right|$$
$$+ \left| \sum_{l=K_n+1}^{n} e^{-h_n\lambda_n^l} v_n^l(x) v_n^l(y) \frac{d_n(y)}{n} \right|.$$

From Lemma 9 we get that the first term on the right-hand side is $o\left(\frac{\sqrt{h_n}}{n}\right)$ with probability greater than $1 - a_1\epsilon_n^{-6k} \exp(-a_2 n\epsilon_n^{k+4}) - a_3 n \exp(-a_4 n (\lambda_{K_n} + 1)^{-k})$, while the second term is estimated in the same way as the term $\mathcal{R}_n$ in the previous part of the proof.

Step 2. We choose an optimal transport map

$$T_n \in \underset{T_\# \nu = \nu_n}{\operatorname{argmin}} \sup_{x\in M} d_M(x, T(x)),\ \theta_n := \sup_{x\in M} d_M(x, T_n(x)).$$

We claim that, with probability greater than $1 - a_1\epsilon_n^{-6k} \exp(-a_2 n\epsilon_n^{k+4}) - a_3 n \exp(-a_4 n (\lambda_{K_n} + 1)^{-k})$, we have for every $f \in C^\infty(M)$,

$$\max_{x\in V_n} \left| e^{-h_n\Delta_n} f(x) - e^{-\kappa(\eta)h_n\Delta_{\rho^2}} f(x) \right| \leq L_1 \sup_M |f| \frac{\theta_n}{\sqrt{h_n}} e^{\frac{2\theta_n \operatorname{diam}(M)}{h_n}}.$$
$$+ \sup_M |f| o(\sqrt{h_n}) + L_2 \left(\sup_M |f| + \operatorname{Lip}(f) \right) \theta_n \tag{37}$$

where the constants L_1, L_2 and the function in $o(\sqrt{h_n})$ depend only on M.

To show (37), we work under the assumption that we are in the event in which the estimate of Step 1 holds true; this happens with probability greater than

$$1 - a_1\epsilon_n^{-6k} \exp(-a_2 n\epsilon_n^{k+4}) - a_3 n \exp(-a_4 n (\lambda_{K_n} + 1)^{-k}).$$

We take $f \in C^\infty(M)$ and $x \in V_n$. Then by using the triangle inequality

$$|e^{-h_n\Delta_n} f(x) - e^{-\kappa(\eta)h_n\Delta_{\rho^2}} f(x)|$$

$$
\begin{aligned}
&= \left| \sum_{y\in V_n} H^n_{\epsilon_n}(h_n,x,y)f(y) - \int_M H(\kappa(\eta)h_n,x,y)f(y)\rho^2(y)d\operatorname{Vol}_M(y)\right| \\
&\le \sum_{y\in V_n}\left| H^n_{\epsilon_n}(h_n,x,y)f(y) - \frac{\rho(y)}{n}H(\kappa(\eta)h_n,x,y)f(y)\right| \\
&\quad + \left|\sum_{y\in V_n}\frac{\rho(y)}{n}H(\kappa(\eta)h_n,x,y)f(y) - \int_M H(\kappa(\eta)h_n,x,y)f(y)\rho^2(y)d\operatorname{Vol}_M(y)\right|.
\end{aligned}
$$

For the first term on the right-hand side, we use the estimate in Step 1 to infer

$$
\begin{aligned}
\sum_{y\in V_n}\left| H^n_{\epsilon_n}(h_n,x,y)f(y) - \frac{\rho(y)}{n}H(\kappa(\eta)h_n,x,y)f(y)\right| &\le n\sup_M|f|o\left(\frac{\sqrt{h_n}}{n}\right) \\
&= \sup_M|f|o(\sqrt{h_n}).
\end{aligned}
$$

For the second term, we recall that $(T_n)_\#\nu = \nu_n$, thus

$$
\begin{aligned}
&\left|\sum_{y\in V_n}\frac{\rho(y)}{n}H(\kappa(\eta)h_n,x,y)f(y) - \int_M H(\kappa(\eta)h_n,x,y)f(y)\rho^2(y)d\operatorname{Vol}_M(y)\right| \\
&= \left|\int_M H(\kappa(\eta)h_n,x,T_n(y))f(T_n(y))\rho(T_n(y))d\nu(y) - \int_M H(\kappa(\eta)h_n,x,y)f(y)\rho(y)d\nu(y)\right|.
\end{aligned}
$$

By the smoothness of ρ and f, we observe that

$$
\left|\int_M H(\kappa(\eta)h_n,x,y)\left(f(T_n(y))\rho(T_n(y)) - f(y)\rho(y)\right)d\nu(y)\right| \le L_2\left(\sup_M|f| + \operatorname{Lip}(f)\right)\theta_n,
$$

so we are left with showing that

$$
\begin{aligned}
&\left|\int_M (H(h_n,x,T_n(y)) - H(h_n,x,y))f(T_n(y))\rho(T_n(y))d\nu(y)\right| \\
&\le L_1\sup_M|f|\frac{\theta_n}{\sqrt{h_n}}e^{\frac{\theta_n\operatorname{diam}(M)}{h_n}}.
\end{aligned}
\tag{38}
$$

To prove (38) we fix $x,y\in M$ and we consider the length minimizing constant-speed geodesic $\sigma_{n,y}:[0,1]\to M$ from y to $T_n(y)$, i.e.,

$$
\operatorname{Length}(\sigma_{n,y}|_{[0,s]}) = d_M(y,\sigma_{n,y}(s)).
$$

By the fundamental theorem of calculus, the Cauchy–Schwarz inequality and the boundedness of ρ we obtain

$$
\begin{aligned}
&\left|\int_M (H(h_n,x,T_n(y)) - H(h_n,x,y))f(T_n(y))\rho(T_n(y))d\nu(y)\right| \\
&\le C\sup_M|f|\int_0^1\int_M|\nabla H(h_n,x,\sigma_{n,y}(s))||\dot\sigma_{n,y}(s)|d\nu(y)ds
\end{aligned}
$$

$$\leq C\theta_n \sup_M |f| \int_0^1 \int_M \frac{\hat{Q}_1}{\sqrt{h_n}\mu(B_{\sqrt{h_n}}(x))} \exp\left(-\frac{d_M^2(x,\sigma_{n,y}(s))}{\hat{Q}_2 h_n}\right) d\nu(y)ds, \tag{39}$$

where in the last line we used the fact that the speed of the constant-speed geodesic $\sigma_{n,y}$ is equal to its length – which can be bounded by $C\theta_n$ by definition of θ_n – and we estimated the gradient of the heat kernel by an application of Theorem 25. We now observe that by the reverse triangle inequality

$$\begin{aligned} |d_M^2(x,\sigma_{n,y}(s)) - d_M^2(x,y)| &= (d_M(x,y) - d_M(x,\sigma_{n,y}(s)))(d_M(x,\sigma_{n,y}(s)) + d_M(x,y)) \\ &\leq 2\theta_n d_M(x,y). \end{aligned}$$

Inserting this estimate into (39) and using the Gaussian lower bound for the heat kernel from Theorem 25 yields

$$\begin{aligned} &\left| \int_M (H(h_n,x,T_n(y)) - H(h_n,x,y)) f(T_n(y))\rho(T_n(y)) d\nu \right| \\ &\leq C\frac{\theta_n}{\sqrt{h_n}} e^{\frac{2\theta_n \operatorname{diam}(M)}{h_n}} \sup_M |f| \int_M H(\tilde{Q}h_n,x,y)d\nu(y) \\ &\leq L_1 \sup_M |f| \frac{\theta_n}{\sqrt{h_n}} e^{\frac{2\theta_n \operatorname{diam}(M)}{h_n}}. \end{aligned}$$

Step 3. Conclusion. To conclude the proof of the theorem from (37) one clearly just needs to prove that

$$\limsup_{n\to+\infty} \frac{\theta_n}{h_n^{3/2}} < +\infty.$$

We first treat the case $k \geq 3$. Observe that, by Theorem 29

$$\limsup_{n\to+\infty} \frac{n^{1/k}\theta_n}{\log^{1/k}(n)} < +\infty.$$

In particular, using also assumption (vi)

$$\limsup_{n\to+\infty} \frac{\theta_n}{h_n^{3/2}} = \limsup_{n\to+\infty} \left(\frac{n^{1/k}\theta_n}{\log^{1/k}(n)} \frac{\log^{1/k}(n)}{\epsilon_n n^{1/k}} \frac{\epsilon_n}{h_n^{3/2}} \right) < +\infty,$$

provided

$$\limsup_{n\to+\infty} \frac{\epsilon_n}{h_n^{3/2}} < +\infty. \tag{40}$$

To check that (40) is satisfied, we observe that by the assumptions (iv) and (v) in Theorem 6 we get

$$\limsup_{n\to+\infty} \frac{\epsilon_n}{h_n^{3/2}} \leq \limsup_{n\to+\infty} (\log(n))^{\frac{3}{2}\alpha-\beta},$$

the right-hand side of which is finite since assumption (i) in Theorem 6 implies $\frac{3}{2}\alpha - \beta \leq 0$. For the case $k = 2$ we proceed analogously. Recall that by Theorem 29

$$\limsup_{n\to+\infty} \frac{n^{1/2}\theta_n}{\log^{3/4}(n)} < +\infty.$$

In particular, using also assumption (vi) in Theorem 6 we obtain

$$\limsup_{n\to+\infty} \frac{\theta_n}{h_n^{3/2}} = \limsup_{n\to+\infty} \left(\frac{\theta_n n^{1/2}}{\log^{3/4}(n)} \left(\frac{\log(n)}{\epsilon_n^8 n} \right)^{1/2} \frac{\epsilon_n^4 \log^{1/4}(n)}{h_n^{3/2}} \right) < +\infty,$$

provided

$$\limsup_{n\to+\infty} \frac{\epsilon_n^4 \log^{1/4}(n)}{h_n^{3/2}} < +\infty.$$

To show this, we estimate ϵ_n using assumption (v) in Theorem 6 and esimate h_n using assumption (iv) in Theorem 6

$$\limsup_{n\to+\infty} \frac{\epsilon_n^4 \log^{1/4}(n)}{h_n^{3/2}} \le \limsup_{n\to+\infty} \left(\log(n)\right)^{\frac{1}{4}+\frac{3}{2}\alpha-4\beta} < +\infty,$$

which follows from (i) in Theorem 6. ■

Proof of Lemma 9. In this proof, we denote by γ the quantity $\gamma := \inf_{i\in\mathbf{N}}(\lambda_{i+1}-\lambda_i)$, which is positive by Item (ii) in Theorem 6.

To show (15), fix two points $x, y \in V_n$. By using the expansion (33) and the triangle inequality we have

$$\left| H_{\epsilon_n}^{K_n}(h_n, x, y) - \frac{\rho(y)}{n} H(\kappa(\eta)h_n, x, y) \right| \le \mathbf{I}_n + \mathbf{II}_n,$$

where we define

$$\mathbf{I}_n = \left| \sum_{l=1}^{K_n-1} e^{-h_n\lambda_n^l} v_n^l(x) v_n^l(y) \frac{d_n(y)}{n} - e^{-h_n\kappa(\eta)\lambda^l} f_l(x) f_l(y) \frac{\rho(y)}{n} \right|,$$

$$\mathbf{II}_n = \left| \sum_{l=K_n}^{+\infty} e^{-h_n\kappa(\eta)\lambda^l} f_l(x) f_l(y) \frac{\rho(y)}{n} \right|.$$

We now proceed to show that these two terms are both of order $o\left(\frac{\sqrt{h_n}}{n}\right)$.
To control term $\mathbf{II}_n$ we follow the ideas in Dunson et al. (2021) and Bérard et al. (1994). By the Cauchy–Schwarz inequality and by the fact that ρ is bounded we get

$$\mathbf{II}_n \le \frac{C}{n} \max_{z\in M} \sum_{l=K_n}^{+\infty} e^{-h_n\kappa(\eta)\lambda_l} f_l^2(z).$$

To control the right hand side, fix $z \in M$. We define a measure ω_z on $\mathbf{R}$ by

$$\omega_z := \sum_{l=K_n}^{+\infty} f_l^2(z) \delta_{\lambda_l}(d\lambda).$$

Then an integration by parts yields

$$\sum_{l=K_n}^{+\infty} e^{-h_n\kappa(\eta)\lambda_l} f_l^2(z) = \int_{\mathbf{R}} e^{-\kappa(\eta)h_n\lambda} d\omega_z(d\lambda)$$

$$
\begin{aligned}
&= \left[e^{-\kappa(\eta)h_n\lambda}\omega_z([0,\lambda])\right]_{\lambda=0}^{+\infty} + \int_{\mathbf{R}} \kappa(\eta)h_n e^{-\kappa(\eta)h_n\lambda}\omega_z([0,\lambda])d\lambda \\
&\leq \limsup_{\lambda\to+\infty}\left(e^{-h_n\kappa(\eta)\lambda}\sum_{\lambda_{K_n}\leq\lambda_l\leq\lambda} f_l^2(z)\right) \\
&\quad + \int_{\lambda_{K_n}}^{+\infty} h_n\kappa(\eta)e^{-h_n\kappa(\eta)\lambda}\omega_z([0,\lambda])d\lambda.
\end{aligned}
$$

Now we use Theorem 25 to show that the first term on the right hand side vanishes. Recalling the notation $\mu := \xi\,\mathrm{Vol}_M$, and using the Gaussian upper bounds in Theorem 25 we get in particular

$$
\begin{aligned}
\sum_{\lambda_{K_n}\leq\lambda_l\leq\lambda} f_l^2(z) \leq e\sum_{0\leq\lambda_l\leq\lambda} e^{-\frac{\lambda_l}{\lambda}} f_l^2(z) &\leq eH\left(\frac{1}{\lambda},z,z\right) \\
&\leq \frac{C}{\mu(B_{\lambda^{-1/2}}(x)))} \leq C\lambda^{\frac{k}{2}},
\end{aligned}
\tag{41}
$$

so that indeed

$$
\limsup_{\lambda\to+\infty} e^{-h_n\frac{\kappa(\eta)}{2}\lambda}\sum_{\lambda_{K_n}\leq\lambda_l\leq\lambda} f_l^2(z) \leq \limsup_{\lambda\to+\infty} e^{-h_n\frac{\kappa(\eta)}{2}\lambda} C\lambda^{\frac{k}{2}} = 0.
$$

We thus obtain, using (41) once more with λ_{K_n} replaced by zero,

$$
\begin{aligned}
\mathbf{II}_n &\leq \frac{C}{n}\int_{\lambda_{K_n}}^{+\infty} h_n\kappa(\eta)e^{-h_n\kappa(\eta)\lambda}\lambda^{k/2}d\lambda \\
&= \frac{C}{n}\left(h_n\kappa(\eta)\right)^{-\frac{k}{2}}\int_{\kappa(\eta)h_n\lambda_{K_n}}^{+\infty} e^{-\lambda}\lambda^{k/2}d\lambda \\
&\leq \frac{C}{n}h_n^{-\frac{k}{2}}\int_{ch_nK_n^{2/k}}^{+\infty} e^{-\lambda}\lambda^{k/2}d\lambda,
\end{aligned}
$$

where we used Weyl's law in the last step. If $ch_nK_n^{\frac{2}{k}} - \frac{k}{2} \geq 1$, we can estimate the right hand side by

$$
\frac{C}{n}h_n^{-\frac{k}{2}}\left(ch_nK_n^{\frac{2}{k}}\right)^{\frac{k}{2}+1} e^{-ch_nK_n^{\frac{2}{k}}} = \frac{\tilde{C}}{n}K_ne^{-A}A,
$$

where $A = ch_nK_n^{\frac{2}{k}}$. Now we follow the reasoning in the proof of (Dunson et al., 2021, Theorem 3) to obtain $K_nAe^{-A} \leq \frac{1}{K_n}e^{-\frac{A}{2}}$ provided $A \geq 8\log(K_n)$, which is satisfied because of our assumption (iv) in Theorem 6. Thus, using again our assumptions on h_n

$$
\mathbf{II}_n \leq \frac{\tilde{C}\sqrt{h_n}}{n}\left(\frac{e^{-c(\log(n))^{\frac{2q}{k}-\alpha}}}{(\log(n))^q\sqrt{h_n}}\right)
$$

$$\leq \frac{\tilde{C}\sqrt{h_n}}{n}\left(e^{-c(\log(n))^{\frac{2q}{k}-\alpha}}(\log(n))^{\frac{\alpha}{2}-q}\right).$$

Thus we obtain that $\mathbf{II}_n = o\left(\frac{\sqrt{h_n}}{n}\right)$ because of the definition of α.

Regarding the term $\mathbf{I}_n$, we use the triangle inequality, to decompose this into four terms

$$\mathbf{I}_n \leq \mathbf{I}_n^a + \mathbf{I}_n^b + \mathbf{I}_n^c + \mathbf{I}_n^d,$$

where

$$\mathbf{I}_n^a = \left|\sum_{l=1}^{K_n-1}\left(e^{-h_n\lambda_n^l} - e^{-\kappa(\eta)h_n\lambda_l}\right)\frac{\rho(y)}{n}f_l(x)f_l(y)\right|,$$

$$\mathbf{I}_n^b = \left|\sum_{l=1}^{K_n-1} e^{-h_n\lambda_n^l}\left(C_1\frac{\rho(y)}{n} - \frac{d_n(y)}{n}\right)\frac{f_l(x)}{C_1^{1/2}}\frac{f_l(y)}{C_1^{1/2}}\right|,$$

$$\mathbf{I}_n^c = \left|\sum_{l=1}^{K_n-1} e^{-h_n\lambda_n^l}\frac{d_n(y)}{n}\left(\frac{f_l(x)}{C_1^{1/2}} - v_n^l(x)\right)\frac{f_l(y)}{C_1^{1/2}}\right|,$$

$$\mathbf{I}_n^d = \left|\sum_{l=1}^{K_n-1} e^{-h_n\lambda_n^l}\frac{d_n(y)}{n}v_n^l(x)\left(\frac{f_l(y)}{C_1^{1/2}} - v_n^l(y)\right)\right|.$$

We now proceed at estimating these four terms.

Term $\mathbf{I}_n^a$. We observe that $\lambda_n^1 = \lambda_1 = 0$, thus in the sum we can neglect the term corresponding to $l = 1$, i.e.

$$\mathbf{I}_n^a \leq \frac{C}{n}\sum_{l=2}^{K_n-1}\left|e^{-h_n\lambda_n^l} - e^{-h_n\kappa(\eta)\lambda_l}\right| \|f_l\|_{C^0(M)}^2.$$

Since $s \mapsto e^{-s}$ is 1-Lipschitz continuous on $[0, +\infty)$, for every $2 \leq l \leq K_n - 1$ we have

$$\left|e^{-h_n\lambda_n^l} - e^{-\kappa(\eta)h_n\lambda_l}\right| \leq |\lambda_n^l - \kappa(\eta)\lambda_l| h_n \leq Q_5\frac{\|f_l\|_{C^3(M)}}{\gamma}\epsilon_n h_n,$$

where the last inequality holds with probability greater than $1 - Q_1\epsilon_n^{-6k}\exp(-Q_2 n\epsilon_n^{k+4})$ $- Q_3 n\exp(-Q_4 n\left(\lambda_{\bar{l}}+1\right)^{-k})$ because of Theorem 27. In particular using also Theorem 24 to control the C^0 and C^3 norm of the eigenfunctions and using the fact that for $l \leq K_n$ we have $\lambda_l \leq \lambda_{K_n}$ we can bound

$$\mathbf{I}_n^a \leq \frac{Ch_n}{n}\left(\frac{K_n\left(\lambda_{K_n}^{1+\frac{k}{2}}+1\right)^2\left(\lambda_{K_n}^{4+\frac{k}{2}}+1\right)\epsilon_n}{\gamma}\right).$$

From this, we obtain that $\mathbf{I}_n^a = o\left(\frac{\sqrt{h_n}}{n}\right)$, because by our assumptions on ϵ_n in (v) of Theorem 6 and our assumptions on the spectral gap in (ii) of Theorem 6 we clearly have

$$\left(\frac{K_n\left(\lambda_{K_n}^{1+\frac{k}{2}}+1\right)^2\left(\lambda_{K_n}^{4+\frac{k}{2}}+1\right)\epsilon_n}{\gamma}\right) = o(1).$$

Term $\mathbf{I}_n^b$. Using Theorem 27, Theorem 28 and Theorem 24 we have that with probability greater than $1 - Q_1\epsilon_n^{-6k}\exp(-Q_2 n\epsilon_n^{k+4}) - Q_3 n\exp(-Q_4 n\left(\lambda_{\bar{l}}+1\right)^{-k}) - Q_6\epsilon_n^{-k}\exp(-Q_7 n\epsilon_n^{k+2})$, for each $1 \leq l \leq K_n - 1$ we can estimate

$$\begin{aligned}
&\left| e^{-h_n\lambda_n^l}\left(C_1\frac{\rho(y)}{n} - \frac{d_n(y)}{n}\right)\frac{f_l(x)}{C_1^{1/2}}\frac{f_l(y)}{C_1^{1/2}}\right| \\
&\leq \frac{C}{n}e^{-h_n\kappa(\eta)\lambda_l}e^{-h_n\left(\lambda_n^l-\kappa(\eta)\lambda_l\right)}\|C_1\rho - d_n\|_{L^\infty(G_n)}\|f_l\|^2_{L^\infty(M)} \\
&\leq \frac{C}{n}e^{Ch_n\frac{\left(\lambda_{K_n}^{4+\frac{k}{2}}+1\right)\epsilon_n}{\gamma}}\left(\lambda_{K_n}^{1+\frac{k}{2}}+1\right)^2\epsilon_n.
\end{aligned}$$

In particular, multiplying and dividing by $\sqrt{h_n}$ and summing over $l = 1, ..., K_n$, we obtain

$$\mathbf{I}_n^b \leq \frac{C\sqrt{h_n}}{n}\left(\frac{K_n}{\sqrt{h_n}}e^{ch_n\frac{\left(\lambda_{K_n}^{4+\frac{k}{2}}+1\right)\epsilon_n}{\gamma}}\left(\lambda_{K_n}^{1+\frac{k}{2}}+1\right)^2\epsilon_n\right).$$

By Weyl's law and our by assumptions (v), (iv) and (ii) in Theorem 6, this is again an $o\left(\frac{\sqrt{h_n}}{n}\right)$ term.

The terms $\mathbf{I}_n^c, \mathbf{I}_n^d$ are treated similarly. In particular $\mathbf{I}_n = o\left(\frac{\sqrt{h_n}}{n}\right)$ provided we are in the event in which Theorem 27 and Theorem 28 apply. This happens with probability greater than

$$\begin{aligned}
&1 - Q_1\epsilon_n^{-6k}\exp(-Q_2 n\epsilon_n^{k+4}) - Q_3 n\exp(-Q_4 n\left(\lambda_{\bar{l}}+1\right)^{-k}) - Q_6\epsilon_n^{-k}\exp(-Q_7 n\epsilon_n^{k+2}) \\
&\geq 1 - (Q_1+Q_6)\epsilon_n^{-6k}\exp(-\min(Q_2,Q_7)n\epsilon_n^{k+4}) - Q_3 n\exp(-Q_4 n\left(\lambda_{\bar{l}}+1\right)^{-k}) \\
&= 1 - a_1\epsilon_n^{-6k}\exp(-a_2 n\epsilon_n^{k+4}) - a_3 n\exp(-a_4 n\left(\lambda_{K_n}+1\right)^{-k}),
\end{aligned}$$

provided n is large enough, this concludes our argument for (15). ■

Proof of Corollary 8. We know from Theorem 6 that for n large enough, assumptions (i), (ii), (iii) of Theorem 4 hold for both the choices of the operators $e^{-t\Delta_n}$ and P_n on the graph G_n on an event A_n such that

$$\mathbb{P}(A_n) \geq 1 - C\epsilon_n^{-6k}\exp(-\frac{1}{C}n\epsilon_n^{k+4}) - Cn\exp(-\frac{n}{CK_n^2}).$$

For $\overline{n} \in \mathbf{N}$ large enough we consider the set

$$C_{\overline{n}} := \bigcap_{n\geq\overline{n}} A_n.$$

Theorem 4 says that, in the event $C_{\overline{n}}$, for both the choices of the operators $e^{-t\Delta_n}$ and P_n we have that (6) and (7) hold true. Observe that

$$\mathbb{P}(C_{\overline{n}}) \geq 1 - \sum_{n\geq\overline{n}} C\epsilon_n^{-6k}\exp(-\frac{1}{C}n\epsilon_n^{k+4}) - Cn\exp(-\frac{1}{C}nK_n^{-2}),$$

In particular, we have that

$$\begin{aligned}\mathbb{P}\left(\{u^* \text{ and } u_* \text{ satisfy (6) and (7)}\}\right) &\geq \mathbb{P}\left(\bigcup_{\overline{n}\in\mathbf{N}} C_{\overline{n}}\right) = \lim_{\overline{n}\to+\infty} \mathbb{P}(C_{\overline{n}}) \\ &\geq 1 - \lim_{\overline{n}\to+\infty} \sum_{n\geq\overline{n}} \Bigg(C\epsilon_n^{-6k} \exp(-\frac{1}{C} n\epsilon_n^{k+4}) \\ &\qquad - Cn\exp(-\frac{1}{C} nK_n^{-2}) \Bigg).\end{aligned} \tag{42}$$

We thus just need to show that

$$\lim_{\overline{n}\to+\infty} \sum_{n\geq\overline{n}} \left(C\epsilon_n^{-6k} \exp(-\frac{1}{C} n\epsilon_n^{k+4}) - Cn\exp(-\frac{1}{C} nK_n^{-2}) \right) = 0,$$

in other words, we need to prove that the series is convergent. To this end, observe that

$$\begin{aligned}C\epsilon_n^{-6k}\exp(-\frac{1}{C}n\epsilon_n^{k+4}) &= C\exp\left(-6k\log(\epsilon_n) - \frac{1}{C}n\epsilon_n^{k+4}\right) \\ &= C\exp\left(\log(n)\left(-6k\frac{\log(\epsilon_n)}{\log(n)} - \frac{1}{C}\frac{n\epsilon_n^{k+4}}{\log(n)}\right)\right) \\ &= Cn^{\left(-6k\frac{\log(\epsilon_n)}{\log(n)} - \frac{1}{C}\frac{n\epsilon_n^{k+4}}{\log(n)}\right)}.\end{aligned}$$

In a similar way, we have

$$Cn\exp(-\frac{1}{C}n\left(\lambda_{K_n}+1\right)^{-k}) = Cn^{\left(1-\frac{1}{C}\frac{nK_n^{-2}}{\log(n)}\right)}.$$

To prove the convergence of the series appearing in (42) it is sufficient to show

$$\lim_{n\to+\infty}\left(-6k\frac{\log(\epsilon_n)}{\log(n)} - \frac{1}{C}\frac{n\epsilon_n^{k+4}}{\log(n)}\right) = \lim_{n\to+\infty}\left(1 - \frac{1}{C}\frac{nK_n^{-2}}{\log(n)}\right) = -\infty.$$

The second limit is easily treated by recalling that $K_n = \log^q(n)$. To treat the first limit, observe that by assumption (14) in Corollary 8 we have

$$\lim_{n\to+\infty}\frac{n\epsilon_n^{k+4}}{\log(n)} = +\infty.$$

To conclude the proof, we show that

$$\inf_{n\in\mathbf{N}}\frac{\log(\epsilon_n)}{\log(n)} > -\infty. \tag{43}$$

Indeed, we have

$$\frac{\log(\epsilon_n)}{\log(n)} = \frac{\log\left(\frac{\epsilon_n n^{\frac{1}{k+4}}}{\log^{\frac{1}{k+4}}(n)}\right)}{\log(n)} - \frac{1}{k+4} + \frac{\log\log(n)}{\log(n)},$$

The first term is bounded from below because it is asymptotically nonnegative by (14) . The last term converges to zero as $n \to +\infty$. Thus (43) holds and the proof is complete. ■

B.3 MBO on manifolds

In this subsection, we prove the results of Appendix A.

Proof of Theorem 20. We let $\hat{x} := \exp_x(z(x)\nu(x))$. Then we have

$$\frac{1}{2} + \omega_1\sqrt{h} = \int_{\Omega_0} H(\kappa h, \hat{x}, y)\rho^2(y) d\,\mathrm{Vol}_M$$

By the Gaussian upper bounds on the heat kernel in Theorem 25, we have that $d_M(\hat{x}, \partial\Omega_0) \leq \tilde{C}\sqrt{h}$, for a fixed constant $\tilde{C}$, independent of Ω_0. In particular, we infer from the asymptotic expansion of the heat kernel in Theorem 26 that

$$\frac{1}{2} + \omega_1\sqrt{h} = \int_{\Omega_0} \frac{e^{-\frac{d_M^2(\hat{x},y)}{4\kappa h}}}{(4\pi\kappa h)^{k/2}} v_0(\hat{x}, y)\rho^2(y) d\,\mathrm{Vol}_M + O(h). \tag{44}$$

Since $d(\hat{x}, \partial\Omega_0) \leq \tilde{C}h$, and $\mathrm{diam}(\Omega_0) \leq \frac{\mathrm{inj}(M)}{2}$, we can rewrite the integral in (44) in exponential coordinates around $\hat{x}$, i.e.

$$\frac{1}{2} + \omega_1 \circ \exp_{\hat{x}} \sqrt{h} = \int_{\tilde{\Omega}_0} \frac{e^{-\frac{|y|^2}{4\kappa h}}}{(4\pi\kappa h)^{k/2}} v_0(\hat{x}, \exp_{\hat{x}}(y))\rho^2(\exp_{\hat{x}}(y)) dy + O(h),$$

where $\tilde{\Omega}_0 := \exp_{\hat{x}}^{-1}(\Omega_0)$. Recalling that $v_0(\hat{x}, \hat{x}) = \frac{1}{\rho^2(\hat{x})}$, a Taylor expansion of the function $y \mapsto v_0(\hat{x}, \exp_{\hat{x}}(y))\rho^2(\exp_{\hat{x}}(y))$ around zero reveals that

$$\frac{1}{2} + \omega_1 \circ \exp_{\hat{x}} \sqrt{h} = \int_{\tilde{\Omega}_0} \frac{e^{-\frac{|y|^2}{4\kappa h}}}{(4\pi\kappa h)^{k/2}} dy + O(\sqrt{h}).$$

In other words, there exists a bounded function ω_2 on $\mathbf{R}^k$ such that

$$\frac{1}{2} + \omega_2\sqrt{h} = \int_{\tilde{\Omega}_0} \frac{e^{-\frac{|y|^2}{4\kappa h}}}{(4\pi\kappa h)^{k/2}} dy.$$

In other words, we have that $0 \in \partial E$, where

$$E = \left\{ v \in \mathbf{R}^k |\ \frac{1}{2} + \omega_2(v)\sqrt{h} \leq \int_{\tilde{\Omega}_0} \frac{e^{-\frac{|v-y|^2}{4\kappa h}}}{(4\pi\kappa h)^{k/2}} dy \right\},$$

and thus the normal distance $z(x)$ coincides with the normal distance of $\partial\tilde{\Omega}_0$ and E at the point $\exp_{\hat{x}}^{-1}(x) \in \partial\tilde{\Omega}_0$. The conclusion of the proof is then obtained by applying the following result. ■

Proposition 23 *Let $\Omega \subset \mathbf{R}^k$ be a smooth open set. Let E be obtained by applying one step of MBO with diffusion coefficient $\kappa > 0$, bounded drift $\omega : \mathbf{R}^k \to \mathbf{R}$ and step size $h > 0$. Let $x \in \partial\Omega$. Let $\nu(x)$ be the outer unit normal to $\partial\Omega$ at x, define*

$$z(x) := \begin{cases} \sup\{l \in \mathbf{R}^- |\ x + l\nu(x) \in E\} & \text{if } x \notin E \\ \inf\{l \in \mathbf{R}^+ |\ x + l\nu(x) \notin E\} & \text{if } x \in E. \end{cases}$$

Then we have

$$|z(x)| \le \tilde{C}h,$$

where the constant $\tilde{C}$ depends only on k, κ and the C^0-norm of the second fundamental form of $\partial\Omega$.

Proposition 23 is a weaker version of (Fuchs and Laux, 2022, Theorem 4.1), which makes rigorous the original ideas in Mascarenhas (1992). In those works, the authors identify the exact first order coefficient of the expansion of $z(x)$ in h. Since we do not need this, we present a proof of our weaker statement.

Proof of Proposition 23. For ease of notation, we assume that $\kappa = 1$. We treat the case when $z(x) > 0$, the other case is similar. First of all, we observe that $z(x) \le \tilde{C}_k\sqrt{h}$, for a constant $\tilde{C}_k$ depending just on the dimension k. We now choose a coordinate system in which $x = 0$ and $\nu(x) = e_k$. We may find an open set U containing the origin and a smooth function $\zeta : \mathbf{R}^{k-1} \to \mathbf{R}$ such that $\zeta(0) = 0, D\zeta(0) = 0$ and

$$U \cap \Omega = \left\{ v \in \mathbf{R}^k |\ v_k < \zeta(v_1, ..., v_{k-1}) \right\}.$$

Using the fact that $z(x) = O(\sqrt{h})$ and the exponential decay of the heat kernel, we have that there exists a bounded function $\omega : \mathbf{R}^k \to \mathbf{R}$ such that

$$\frac{1}{2} + \omega((0, z(x)))\sqrt{h} = \int_{\mathbf{R}^{k-1}} \int_{-\infty}^{\zeta(y)+z(x)} \frac{e^{-\frac{|y|^2+|s|^2}{4h}}}{(4\pi h)^{k/2}} dsdy. \tag{45}$$

Recalling that the Gaussian integrates to 1/2 over half-spaces, we get that (45) reads

$$\omega((0, z(x)))\sqrt{h} = \int_{\mathbf{R}^{k-1}} \int_{0}^{\zeta(y)+z(x)} \frac{e^{-\frac{|y|^2+|s|^2}{4h}}}{(4\pi h)^{k/2}} dsdy.$$

Since $\zeta(0) = 0$ and $D\zeta(0) = y$, there exists a bounded function ζ_1 such that $\zeta(v) = \zeta_1(v)|v|^2$. We also observe that

$$e^{-t} \ge 1 - t,\ t \ge 0.$$

In particular

$$\begin{aligned}
\omega((0, z(x)))\sqrt{h} &\ge \frac{1}{(4\pi h)^{k/2}} \int_{\mathbf{R}^{k-1}} e^{-\frac{|y|^2}{4h}} \int_0^{\zeta(y)+z(x)} \left(1 - \frac{s^2}{4h}\right) dsdy \\
&= \frac{1}{(4\pi h)^{k/2}} \int_{\mathbf{R}^{k-1}} e^{-\frac{|y|^2}{4h}} \left(\zeta_1(y)|y|^2 + z(x) - \frac{1}{12h}\left(\zeta_1(y)|y|^2 + z(x)\right)^3\right) dy.
\end{aligned}$$

By using the change of variable $y \to \sqrt{h}y$ we obtain

$$\omega((0, z(x))\sqrt{h} \ge \frac{1}{h^{1/2}}\left(z(x) + \frac{q_1}{h}z(x)^3 + q_2 h + q_3 h^2 + q_4 z(x)^2\right),$$

where q_1, q_2, q_3, q_4 are coefficients depending on the first six moments of the function $y \mapsto e^{-|y|^2}$. By multiplying both sides by $\sqrt{h}$ we get

$$\omega((0, z(x))h - \left(q_2 h + q_3 h^2 + q_4 z(x)^2\right) \ge z(x) + \frac{q_1}{h}z(x)^3.$$

By applying (Fuchs and Laux, 2022, Lemma 6.1) (which holds true even if we additionally consider a bounded drift ω), we have that $z(x) = O(h^{3/2})$. In particular, for h small enough

$$\frac{1}{2} < 1 - \frac{q_1}{h} z(x)^2,$$

in other words

$$2\omega((0, z(x))h - 2\left(q_2 h + q_3 h^2 + q_4 z(x)^2\right) \geq z(x) \geq 0,$$

from which we conclude that $z(x) = O(h)$. ■

Proof of Corollary 21. Denote by $\tilde{C}_{r,x_0}$ the constant obtained by applying Theorem 20 to $\Omega_0 = B_r(x_0)$. Since $\tilde{C}_{r,x_0}$ depends on Ω_0 only through the C^0 norm of the second fundamental form S_{r,x_0} of $\partial B_r(x_0)$, it is sufficient to show that this can be bounded independently of $\frac{R}{2} \leq r \leq R$ and $x_0 \in M$. We clearly have that

$$(0, \operatorname{diam}(M)) \times M \ni (r, x_0) \mapsto \|S_{r,x_0}\|_{C^0}$$

is a continuous function. It is thus bounded on the compact set

$$W := \left\{(r, x) \in (0, +\infty) \times M : \frac{R}{2} \leq r \leq R, x \in M\right\}.$$

This completes the proof of Corollary 21. ■

Proof of Theorem 22. For ease of notation, let us assume that $\kappa = 1$. We start by observing that

$$\int_{M \setminus B_{h_n^{\frac{1}{4}}}(z_{h_n})} H(h_n, z_{h_n}, y)\xi(y) d\operatorname{Vol}_M(y) = o\left(\sqrt{h_n}\right).$$

This is proved by using the Gaussian bounds from Theorem 25, as we did in (Laux and Lelmi, 2021, Theorem 3, Step 2). In particular, both in (i) and in (ii) of Theorem 22 we can replace the domain of integration with

$$\{\varphi_{h_n}(s_{h_n} - h_n, \cdot) \geq 0\} \cap B_{h_n^{\frac{1}{4}}}(z_{h_n}).$$

In this way, the sequence of integrals can be computed in normal coordinates around z_{h_n}, i.e.,

$$\begin{aligned}
&\int_{\{\varphi_{h_n}(s_{h_n} - h_n, \cdot) \geq 0\} \cap B_{h_n^{\frac{1}{4}}}(z_{h_n})} H(h_n, z_{h_n}, y)\xi(y) d\operatorname{Vol}_M(y) \\
&= \int_{\{\tilde{\varphi}_{h_n}(s_{h_n} - h_n, \cdot) \geq 0\} \cap B_{h_n^{\frac{1}{4}}}(0)} H(h_n, z_{h_n}, \exp_{z_{n_n}}(y))\xi(\exp_{z_{n_n}}(y))\sqrt{\det(g)} dy,
\end{aligned}$$

where we set

$$\tilde{\varphi}_{h_n}(t, y) := \varphi_{h_n}(t, \exp_{z_{h_n}}(y)), \; y \in B_{\frac{\operatorname{inj}(M)}{2}}(0).$$

Using the asymptotic expansion for the heat kernel in Theorem 26, it is easy to see that

$$\int_{\{\tilde{\varphi}_{h_n}(s_{h_n} - h_n, \cdot) \geq 0\} \cap B_{h_n^{\frac{1}{4}}}(0)} H(h_n, z_{h_n}, \exp_{z_{n_n}}(y))\xi(\exp_{z_{n_n}}(y))\sqrt{\det(g)} dy$$

$$= \int_{\{\tilde\varphi_{h_n}(s_{h_n}-h_n,\cdot)\geq 0\}\cap B_{h_n^{\frac14}}(0)} \frac{e^{-\frac{|y|^2}{4h_n}}}{(4\pi h_n)^{k/2}} v_0(z_{h_n}, \exp_{z_{h_n}}(y))\xi(\exp_{z_{n_n}}(y))\sqrt{\det(g)}dy$$
$$+ o(\sqrt{h}_n).$$

In particular, in both (i) and (ii) in Theorem 22 the integrals may be substituted with

$$\int_{\{\tilde\varphi_{h_n}(s_{h_n}-h_n,\cdot)\geq 0\}\cap B_{h_n^{\frac14}}(0)} \frac{e^{-\frac{|y|^2}{4h_n}}}{(4\pi h_n)^{k/2}} v_0(z_{h_n}, \exp_{z_{h_n}}(y))\xi(\exp_{z_{n_n}}(y))\sqrt{\det(g)}dy.$$

These integrals may be furthermore decomposed into the sums $\mathbb{I}_n + \mathbb{II}_n$,

$$\mathbb{I}_n := \int_{\{\tilde\varphi_{h_n}(s_{h_n}-h_n,\cdot)\geq 0\}\cap B_{h_n^{\frac14}}(0)} \frac{e^{-\frac{|y|^2}{4h_n}}}{(4\pi h_n)^{k/2}} dy,$$
$$\mathbb{II}_n := \int_{\{\tilde\varphi_{h_n}(s_{h_n}-h_n,\cdot)\geq 0\}\cap B_{h_n^{\frac14}}(0)} \frac{e^{-\frac{|y|^2}{4h_n}}}{(4\pi h_n)^{k/2}} (w_n(y)-1)dy,$$

where we define

$$w_n(y) := v_0(z_{h_n}, \exp_{z_{h_n}}(y))\xi(\exp_{z_{n_n}}(y))\sqrt{\det(g)}.$$

We claim that

$$\lim_{n\to+\infty} \mathbb{II}_n = \begin{cases} 0 & \text{if } \nabla\varphi(s,z)=0, \\ \frac{1}{2\sqrt{\pi}|\nabla\varphi(s,z)|}\langle \frac{\nabla\xi}{\xi}(z), \nabla\varphi(s,z)\rangle & \text{otherwise.} \end{cases} \tag{46}$$

Using (49) we see that

$$w_n(y) = \sqrt{\frac{\xi(\exp_{z_{n_n}}(y))\det(g)}{\xi(z_{h_n})\det(d_{\exp^{-1}_{z_{h_n}}(y)}(\exp_{z_{h_n}}))}}.$$

In particular, denoting $\tilde\xi_n = \xi\circ\exp_{z_{h_n}}$ and $D_n := \det(d_{\exp^{-1}_{z_{h_n}}(y)}(\exp_{z_{h_n}}))$ we get

$$Dw_n = \frac{1}{2w_n(y)} \frac{\Big((D_y\tilde\xi_n)\det(g) + \tilde\xi_n D_y\det(g))\tilde\xi_n(0)D_n - \tilde\xi_n\det(g)\tilde\xi_n(0)D_yD_n\Big)}{\tilde\xi_n(0)^2D_n^2}.$$

We now recall that, in normal coordinates $g(z_{h_n}) = Id$, $Dg(z_{h_n}) = 0$, in particular

$$Dw_n(z_{h_n}) = \frac12\frac{D\tilde\xi_n}{\tilde\xi_n}(0),$$

and by a Taylor expansion

$$Dw_n(y) = 1 + \frac12\frac{D\tilde\xi_n}{\tilde\xi_n}(0)\cdot y + O(|y|^2);$$

in particular, we infer that

$$\mathbb{III}_n = \frac{1}{2}\frac{D\tilde{\xi}_n}{\tilde{\xi}_n}(0)\cdot\int_{\{\tilde{\varphi}_{h_n}(s_{h_n}-h_n,\cdot)\geq 0\}\cap B_{h_n^{\frac{1}{4}}}(0)} \frac{e^{-\frac{|y|^2}{4h_n}}}{(4\pi h_n)^{k/2}} y dy + O(h_n).$$

Now we claim that

$$\begin{aligned}
&\lim_{n\to+\infty}\frac{1}{2\sqrt{h_n}}\frac{D\tilde{\xi}_n}{\tilde{\xi}_n}(0)\cdot\int_{\{\tilde{\varphi}_{h_n}(s_{h_n}-h_n,\cdot)\geq 0\}\cap B_{h_n^{\frac{1}{4}}}(0)} \frac{e^{-\frac{|y|^2}{4h_n}}}{(4\pi h_n)^{k/2}} y dy \\
&= \frac{1}{2\sqrt{\pi}|\nabla\varphi(s,z)|}\frac{D\tilde{\xi}}{\tilde{\xi}}(0)\cdot D\tilde{\varphi}(s,0),
\end{aligned} \tag{47}$$

where $\tilde{\xi} = \xi\circ\exp_z$. Of course (47) gives (46).

To see that (47) holds, we start by changing variable in the integral by setting $y = \frac{y}{\sqrt{h_n}}$, which gives that the argument in the limit equals

$$\frac{1}{2}\frac{D\tilde{\xi}_n}{\tilde{\xi}_n}(0)\cdot\int_{\left\{y|\ \tilde{\varphi}_{h_n}(s_{h_n}-h_n,\sqrt{h_n}y)\geq 0\right\}\cap B_{h_n^{-\frac{1}{4}}}(0)} \frac{e^{-\frac{|y|^2}{4}}}{(4\pi)^{k/2}} y dy.$$

We now let R_n be a sequence of orthogonal matrices such that $R_n^T e_1 = \frac{D\tilde{\xi}_n(0)}{|D\tilde{\xi}_n(0)|}$ and without loss of generality we assume that the sequence converges to an orthogonal matrix R. We change variable by setting $y = R_n^T y$ and we get that the argument of the limit becomes

$$\frac{|D\tilde{\xi}_n(0)|}{2}\int_{\mathcal{C}_n\cap B_{h_n^{-\frac{1}{4}}}(0)} \frac{e^{-\frac{|y|^2}{4}}}{(4\pi)^{k/2}} y_1 dy,$$

where we define

$$\mathcal{C}_n := \left\{y\in\mathbf{R}^k|\ \tilde{\varphi}_{h_n}(s_{h_n}-h_n, R_n\sqrt{h_n}y)\geq 0\right\}.$$

We now observe that, by Taylor expanding $\tilde{\varphi}_{h_n}(t_{h_n}-\cdot,\cdot)$ around $(0,0)$

$$\begin{aligned}
\tilde{\varphi}_{h_n}(s_{h_n}-h_n, R_n\sqrt{h_n}y) =& \delta_{h_n} + \sqrt{h_n}R_n^T D\tilde{\varphi}_{h_n}(s_{h_n},0)\cdot y \\
&- h_n\partial_s\tilde{\varphi}_{h_n}(s_{h_n},0) + o(|y|^2 + h_n^2),
\end{aligned}$$

thus

$$\begin{aligned}
\mathcal{C}_n = \Bigg\{y\in\mathbf{R}^k|\ & \frac{\delta_{h_n}}{\sqrt{h_n}} + R_n^T D\tilde{\varphi}_{h_n}(s_{h_n},0)\cdot y \\
&- \sqrt{h_n}\partial_s\tilde{\varphi}_{h_n}(s_{h_n},0) + o(\sqrt{h_n}|y|^2 + h_n^{\frac{3}{2}})\geq 0\Bigg\}.
\end{aligned}$$

Recalling assumption (19) this re-reads

$$\mathcal{C}_n = \Big\{ y \in \mathbf{R}^k \mid \; R_n^T D\tilde{\varphi}_{h_n}(s_{h_n}, 0) \cdot y + o(1) \geq 0 \Big\}.$$

Observe also that

$$\begin{aligned} R_n D\tilde{\xi}_n(0) &= |D\tilde{\xi}_n(0)| e_1 \\ &= \sqrt{\langle \nabla \xi(z_{h_n}), \nabla \xi(z_{h_n}) \rangle} e_1 \underset{n \to +\infty}{\to} \sqrt{\langle \nabla \xi(z), \nabla \xi(z) \rangle} e_1, \end{aligned}$$

but also

$$R_n D\tilde{\xi}_n(0) = D\tilde{\xi} \circ R_n^T(0) = D\xi \circ \exp_{z_{h_n}} \circ R_n^T(0)) \underset{n \to +\infty}{\to} RD(\xi \circ \exp_z)(0).$$

In other words we must have $D\tilde{\xi}(0) = |D\tilde{\xi}(0)| R^T e_1$. In particular

$$\begin{aligned} \lim_{n \to +\infty} \frac{|D\tilde{\xi}_n(0)|}{2} \int_{\mathcal{C}_n \cap B_{h_n^{-\frac{1}{4}}}(0)} \frac{e^{-\frac{|y|^2}{4}}}{(4\pi)^{k/2}} y_1 dy &= \frac{|D\tilde{\xi}(0)|}{2} \int_{\{y \mid \; R^T D\tilde{\varphi}(s,0) \cdot y \geq 0\}} \frac{e^{-\frac{|y|^2}{4}}}{(4\pi)^{k/2}} y_1 dy \\ &= \frac{|D\tilde{\xi}(0)|}{2} \int_{\{y \mid \; D\tilde{\varphi}(s,0) \cdot y \geq 0\}} \frac{e^{-\frac{|y|^2}{4}}}{(4\pi)^{k/2}} Ry \cdot e_1 dy \\ &= \frac{1}{2} \frac{D\tilde{\xi}}{\tilde{\xi}}(0) \cdot \int_{\{y \mid \; D\tilde{\varphi}(s,0) \cdot y \geq 0\}} \frac{e^{-\frac{|y|^2}{4}}}{(4\pi)^{k/2}} y dy. \end{aligned}$$

If $\nabla \varphi(t, z) = 0$, then the last integral is zero, being component-wise the integral over the whole space of on odd-function. Otherwise we change variable according to $y = O^T y$, where O is an orthogonal matrix such that $OD\tilde{\varphi}(s, 0) = |D\tilde{\varphi}(s, 0)| e_1$, which gives that the last integral equals

$$\begin{aligned} \frac{1}{2} \frac{OD\tilde{\xi}}{\tilde{\xi}}(0) \cdot \int_{\{y \mid \; y_1 \geq 0\}} \frac{e^{-\frac{|y|^2}{4}}}{(4\pi)^{k/2}} y dy &= \frac{1}{2} \frac{OD\tilde{\xi}}{\tilde{\xi}}(0) \cdot e_1 \frac{1}{\sqrt{\pi}} \\ &= \frac{1}{2\sqrt{\pi} |D\tilde{\varphi}(s, 0)|} \frac{D\tilde{\xi}}{\tilde{\xi}}(0) \cdot D\tilde{\varphi}(s, 0). \end{aligned}$$

We are now in a position to prove (i) and (ii) in Theorem 22.

Item (i). By the discussion above, the left hand side of (20) may be substituted with

$$\begin{aligned} &\liminf_{n \to +\infty} \frac{1}{\sqrt{h_n}} \left(\frac{1}{2} - \mathbb{I}_n - \mathbb{III}_n \right) \\ &\geq \liminf_{n \to +\infty} \frac{1}{\sqrt{h_n}} \left(\frac{1}{2} - \mathbb{I}_n \right) - \frac{1}{2\sqrt{\pi} |\nabla \varphi(s, z)|} \left\langle \frac{\nabla \xi}{\xi}(z), \nabla \varphi(s, z) \right\rangle, \end{aligned}$$

where we used (46) in the second line. To estimate

$$\liminf_{n \to +\infty} \frac{1}{\sqrt{h_n}} \left(\frac{1}{2} - \mathbb{I}_n \right)$$

we can use (Barles and Georgelin, 1995, Proposition 4.1) applied with

$$(t_h, x_h) = (s_h, 0),\ (t, x) = (s, 0),\ \phi_h(t, \cdot) = \tilde{\varphi}_h(t, \cdot).$$

The only difference is that here we do not assume that $\phi(t_h, x_h) = 0$, but $\phi(t_h, x_h) = o(\sqrt{h})$ - one can check that the result holds true also with this modification by the same proof of (Barles and Georgelin, 1995, Proposition 4.1). In particular, we get

$$\begin{aligned}
&\liminf_{n\to+\infty} \frac{1}{\sqrt{h_n}}\left(\frac{1}{2} - \int_{\{\varphi_{h_n}(t_{h_n}-h_n,\cdot)\geq 0\}} H(h_n, z_{h_n}, y)\xi(y) d\operatorname{Vol}_M\right) \\
&\geq \frac{1}{2\sqrt{\pi}|D\tilde{\varphi}(s,0)|}\left(\partial_t\tilde{\varphi} + \Delta\tilde{\varphi} - \frac{D\tilde{\varphi}\cdot D^2\tilde{\varphi}D\tilde{\varphi}}{|D\tilde{\varphi}|^2} - \frac{D\tilde{\xi}}{\tilde{\xi}}\cdot D\tilde{\varphi}\right),
\end{aligned}$$

which is equal to the right hand side of (20) because we are using exponential coordinates around z (recall our convention $\Delta = -\sum_{i=1}^k \partial^2_{ii}$).

Item (ii). Once again, by the above discussion, we can assume that

$$\frac{1}{2} - \mathbb{I}_n \leq o(\sqrt{h_n}),$$

and the result follows by applying (Barles and Georgelin, 1995, Proposition 4.1) with

$$(t_h, x_h) = (s_h, 0),\ (t, x) = (s, 0),\ \phi_h(t, \cdot) = \tilde{\varphi}_h(t, \cdot).$$

In this case, there are two differences from the original version (Barles and Georgelin, 1995, Proposition 4.1). First of all, we again do not assume that $\phi_h(t_h, x_h) = 0$, but we assume $\phi_h(t_h, x_h) = o(\sqrt{h})$. Then, we assume that $\frac{1}{2} - \mathbb{I}_n \leq o(\sqrt{h_n})$ and not the stronger $\frac{1}{2} - \mathbb{I}_n \leq 0$. But a quick inspection of the proof of (Barles and Georgelin, 1995, Proposition 4.1) reveals that these changes are irrelevant for the argument to work. ■

Appendix C. Miscellaneous results

C.1 Results on weighted manifolds

Hereafter we collect some results about weighted Laplacians and heat kernels on closed manifolds. Let (M, g) be a k-dimensional, compact Riemannian manifold endowed with a measure $\mu := \xi \operatorname{Vol}_M$, with $\xi \in C^\infty(M)$, $\xi > 0$. We denote by Δ_ξ the associated Laplacian, which is defined on $f \in C^\infty(M)$ as

$$\Delta_\xi f := -\frac{1}{\xi}\operatorname{div}(\xi\nabla f).$$

We denote by H the corresponding heat kernel, i.e., H is a real valued function defined on $(0, +\infty) \times M \times M$ such that for any $u \in L^2(M)$ the function

$$e^{-t\Delta_\xi}u(x) := T(t)u(x) = \int_M H(t, x, y)u(y)d\mu(y),$$

defined for $(t,x) \in (0,+\infty) \times M$, is the unique solution to the Cauchy problem

$$\begin{cases} \partial_t v = -\Delta_\xi v & \text{in } (0,+\infty)\times M, \\ v(0,x) = u(x) & \text{on } M, \end{cases}$$

where the initial value at $t=0$ is attained in the sense that

$$\lim_{t\downarrow 0} e^{-t\Delta_\xi} u = u \text{ in } L^2(M).$$

We will use the following results.

Theorem 24 *Let M, ξ be as above. Let f be an $L^2(\xi)$-normalized eigenfunction of Δ_ξ corresponding to the eigenvalue λ, then for each integer $m \geq 0$*

$$\|f\|_{C^m(M)} \leq C_{M,m}\left(\lambda^{m+1+\frac{k}{2}} + 1\right).$$

Theorem 25 *Let M, ξ be as above. There exists constants $Q_1, Q_2, Q_3, Q_4, \hat{Q}_1, \hat{Q}_2 > 0$ such that for every $t > 0$ and all $x, y \in M$,*

$$\frac{Q_1}{\mu(B_{\sqrt{t}}(x))} e^{-\frac{d_M^2(x,y)}{Q_2 t}} \leq H(t,x,y) \leq \frac{Q_3}{\mu(B_{\sqrt{t}}(x))} e^{-\frac{d_M^2(x,y)}{Q_4 t}}.$$

$$|\nabla_x H(t,x,y)| \leq \frac{\hat{Q}_1}{\sqrt{t}\mu(B_{\sqrt{t}}(x))} \exp\left(-\frac{d_M^2(x,y)}{\hat{Q}_2 t}\right).$$

Theorem 26 *Let M, ξ be as above. There exist functions $v_j \in C^\infty(M\times M), j \in \mathbf{N}$, such that for every $N > l + \frac{k}{2}$ there exists a constant $\tilde{C}_N < \infty$ such that*

$$\left|\nabla^l\left(H(t,x,y) - \frac{e^{-\frac{d_M^2(x,y)}{4t}}}{(4\pi t)^{k/2}} \sum_{j=0}^N v_j(x,y) t^j\right)\right| \leq \tilde{C}_N t^{N+1-\frac{k}{2}}, \tag{48}$$

provided $d(x,y) \leq \frac{\operatorname{inj}(M)}{2}$. Moreover we have

$$v_0(x,y) = \frac{1}{\sqrt{\xi(x)\xi(y)\det(d_{\exp_x^{-1}(y)}\exp_x)}}. \tag{49}$$

Theorem 24 follows by the Sobolev embedding theorem and the L^2-regularity theory for elliptic equations on manifolds. Theorem 25 follows from the Li–Yau inequality for weighted manifolds Setti (1992). The asymptotic expansion in Theorem 26 follows by constructing the heat kernel by means of the *parametrix* method: this construction is technical and we refer to Rosenberg (1997), where this is carried out for the case $\xi = 1$. Here we just sketch the first part of the construction for a general density ξ, which gives (49). The idea is that when x, y are close enough, say $d(x,y) < \frac{\operatorname{inj}(M)}{2}$, a good approximation for the heat kernel should be given by

$$H_N(t,x,y) := G_t(x,y)\left(v_0(x,y) + \ldots + t^N v_N(x,y)\right), \tag{50}$$

for smooth functions v_j and $t > 0$. Here

$$G_t(x,y) := \frac{e^{-\frac{d_M^2(x,y)}{4t}}}{(4\pi t)^{k/2}}.$$

Since the Ansatz (50) should be an approximation of the heat kernel, we would like to have

$$0 = \partial_t H_N + \Delta_\xi H_N, \tag{51}$$

where Δ_ξ denotes the weighted Laplacian with respect to the y-variable. We now compute the right hand side of the above equation by using exponential coordinates around x: we denote them by $(r,\theta) \in [0,R) \times \mathbb{S}^{k-1}$. Observe that

$$\begin{aligned}\partial_t H_N &= \partial_t G_t(v_0 + ... + t^N v_N) + G_t(v_1 + ... + N t^{N-1} v_N)\\ &= \left(\frac{r^2}{4t^2} - \frac{k}{2t}\right) G_t(v_0 + ... + t^N v_N) + G_t(v_1 + ... + N t^{N-1} v_N).\end{aligned}$$

Furthermore

$$\begin{aligned}\Delta_\xi H_N = G_t &\left(\Delta_\xi v_0 + ... + t^N \Delta_\xi v_N\right)\\ &+ \Delta_\xi G_t(v_0 + ... + t^N v_N) - 2\langle \nabla G_t, \left(\nabla v_0 + ... + t^N \nabla v_N\right)\rangle.\end{aligned}$$

Using Gauss' Lemma and the fact that G_t is independent of θ we get

$$\begin{aligned}2\langle \nabla G_t, \left(\nabla v_0 + ... + t^N \nabla v_N\right)\rangle &= 2\partial_r G_t(\partial_r v_0 + ... + t^N \partial_r v_N)\\ &= -\frac{r}{t} G_t(\partial_r v_0 + ... + t^N \partial_r v_N).\end{aligned}$$

We also observe that by definition of Δ_ξ and by using again Gauss' Lemma and the independence of G_t from θ

$$\Delta_\xi G_t = \Delta G_t - \langle \frac{\nabla \xi}{\xi}, \nabla G_t \rangle = \Delta G_t + \frac{r}{2t}\frac{\partial_r \xi}{\xi} G_t.$$

We define

$$D(y) := \det(d_{\exp_x^{-1}(y)} \exp_x).$$

Using the expression of the Laplacian in spherical coordinates and the invariance of G_t with respect to θ we get

$$\Delta G_t = -\frac{\partial^2 G_t}{\partial r^2} - \partial_r G_t \left(\frac{\partial_r D}{D} + \frac{k-1}{r}\right) = -\left(\frac{r^2}{4t^2} - \frac{k}{2t}\right) G_t + \frac{r}{2t}\frac{\partial_r D}{D} G_t.$$

Putting things together we have

$$\begin{aligned}\partial_t H_N + \Delta_\xi H_N = G_t \Bigg(&(v_1 + ... + N t^{N-1} v_N) - (\Delta_\xi v_0 + ... t^N \Delta_\xi v_N)\\ &+ \frac{r}{2t} \partial_r \log(D\xi)(v_0 + ... + t^N v_N) + \frac{r}{t}(\partial_r v_0 + ... + \partial_r v_N)) \Bigg).\end{aligned}$$

Although we cannot get (51) exactly, we can choose v_j in such a way that

$$\partial_t H_N + \Delta_\xi H_N = G_t t^N \Delta_\xi v_N.$$

In other words, we choose the coefficients in such a way that

$$\frac{r}{2t}\partial_r \log(D\xi) v_0 + \frac{r}{t}\partial_r v_0 = 0, \tag{52}$$

$$j t^{j-1} v_j - t^{j-1}\Delta_\xi v_{j-1} + t^{j-1}\frac{r}{2}\partial_r \log(D\xi) v_j + r t^{j-1}\partial_r v_j = 0, \text{ for } 1 \le j \le N. \tag{53}$$

Once one solves (52), one can show inductively that (53) admits a smooth solution v_j. It is easily seen that (52) can be solved explicitly to give

$$v_0(x,y) = \frac{1}{\sqrt{\xi(x)\xi(y)\det(d_{\exp_x^{-1}(y)}\exp_x)}}.$$

From here, the construction of the heat kernel and the estimate (48) follow verbatim as in Rosenberg (1997).

C.2 Results on random geometric graphs

In this subsection we use the setting and the notation of Section 3, with the points $\{x_i\}_{i=1}^{+\infty}$ being given by i.i.d. random points on M, distributed according to a probability distribution $\nu = \rho \operatorname{Vol}_M \in \mathcal{P}(M)$, with $\rho \in C^\infty(M)$, $\rho > 0$. The following two results are proved in Calder and García Trillos (2022); Calder et al. (2022) for the unnormalized graph Laplacian, but the proof of the statements extends when we work with the random walk Laplacian. Hereafter, given $l \in \mathbf{N}$, we set

$$\gamma_l := \inf_{j<l, j\in\mathbf{N}} (\lambda_{j+1} - \lambda_j).$$

Theorem 27 (Theorem 2.6 in Calder et al. (2022) for the unnormalized Laplacian) *In the above-mentioned setting, if additionally, the eigenvalues of Δ_{ρ^2} are simple, then for every $\bar{l} \in \mathbf{N}$ we have that with probability greater than*

$$1 - Q_1\epsilon_n^{-6k}\exp(-Q_2 n\epsilon_n^{k+4}) - Q_3 n \exp(-Q_4 n \left(\lambda_{\bar{l}} + 1\right)^{-k})$$

we have for every $l \le \bar{l}$

$$|\lambda_n^l - \kappa(\eta)\lambda_l| + \max_{z\in V_n}\left| v_n^l(z) - \frac{f_l(z)}{C_1^{1/2}}\right| \le Q_5 \frac{\|f_l\|_{C^3(M)}}{\gamma_l}\epsilon_n.$$

Theorem 28 (Corollary 3.7 in Calder et al. (2022)) *In the above-mentioned setting, if n is large enough, with probability greater than $1 - Q_6\epsilon_n^{-k}\exp(-Q_7 n\epsilon_n^{k+2})$, we have that*

$$\max_{z\in V_n} |d_{n,\epsilon_n}(z) - C_1\rho(z)| \le Q_8\epsilon_n.$$

We also recall the following result, which may be easily derived from (García Trillos et al., 2020, Theorem 2).

Theorem 29 *Let (M, g) be a k-dimensional closed Riemannian manifold. Let $\rho \in C^\infty(M)$, $\rho > 0$ such that $\nu := \rho \, \mathrm{Vol}_M \in \mathcal{P}(M)$. Let $\{X_i\}_{i \in \mathbf{N}}$ be i.i.d. random points in M distributed according to ν and let $\nu_n := \frac{1}{n} \sum_{i=1}^n \delta_{X_i}$ be the associated empirical measures. Then there is a constant $C > 0$ such that almost surely there exist transport maps T_n such that $(T_n)_\# \nu = \nu_n$ and*

$$\begin{cases} \limsup_{n \to +\infty} \frac{n^{1/2} \sup_{x \in M} d_M(x, T_n(x))}{log^{3/4}(n)} \leq C & \textit{if } k = 2, \\ \limsup_{n \to +\infty} \frac{n^{1/k} \sup_{x \in M} d_M(x, T_n(x))}{log^{1/k}(n)} \leq C & \textit{if } k \geq 3. \end{cases}$$