Convergence of graph Dirichlet energies and graph Laplacians on intersecting manifolds of varying dimensions

Leon Bungert[*] Dejan Slepčev[†]

September 30, 2025

Abstract

We study Γ-convergence of graph Dirichlet energies and spectral convergence of graph Laplacians on unions of intersecting manifolds of potentially different dimensions. Our investigation is motivated by problems of machine learning, as real-world data often consist of parts or classes with different intrinsic dimensions. An important challenge is to understand which machine learning methods adapt to such varied dimensionalities. We investigate the standard unnormalized and the normalized graph Dirichlet energies. We show that the unnormalized energy and its associated graph Laplacian asymptotically only sees the variations within the manifold of the highest dimension. On the other hand, we prove that the normalized Dirichlet energy converges to a (tensorized) Dirichlet energy on the union of manifolds that adapts to all dimensions simultaneously. We also establish the related spectral convergence and present a few numerical experiments to illustrate our findings.

Keywords. graph Laplacian, Dirichlet form, nonlocal Dirichlet energy, Gamma-convergence, spectral convergence, discrete-to-continuum limit, multi-manifold learning
MSC2020 classes. 49J55, 49J45, 60D05, 49R50, 68R10, 62G20

Contents

[*]Institute of Mathematics, Center of Artifical Intelligence and Data Science (CAIDAS), University of Würzburg, Emil-Fischer-Str. 40, 97074 Würzburg, Germany. Email: leon.bungert@uni-wuerzburg.de
[†]Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh, United States, Email: slepcev@math.cmu.edu

1 Introduction

1.1 Motivation and Discussion

We investigate Γ-convergence of graph Dirichlet energies and spectral convergence of graph Laplacians on unions of intersecting manifolds of potentially different dimensions. In the process, we also refine the results on Γ-convergence of nonlocal Dirichlet energies on metric measure spaces, when specialized to unions of intersecting manifolds. The motivation for our work arises in machine learning where graph Laplacians are used to capture and utilize the intrinsic geometry of the data [6, 7, 10, 23, 24, 41, 44, 57]. In many (real life) data sets the intrinsic dimensionality of components or clusters may differ [1, 9, 18, 43] and, furthermore, the clusters can have some intersection. For example, in sets of hand-written digits the intrinsic dimensionality of sets of digit 4 is greater than that of digit 1. Furthermore, there is typically some overlap between sets of digits 1 and 7. This lead to the development of specific algorithms tailored to discover the multi-manifold structure of the data [3, 5, 20, 21, 34, 35, 52, 53, 54].

While there have been many theoretical studies of graph Laplacians and their convergence, including [4, 11, 14, 15, 16, 22, 27, 28, 30, 36, 41, 42, 45, 46, 55], there are only a few works [29, 50] that theoretically analyze Laplacian-based algorithms for data of different intrinsic dimensionality. In particular, none of the rigorous works investigated the asymptotics of the standard normalized graph Laplacians on unions of manifolds. Here we show that the asymptotic behavior, as the number of available data points grows, of different, popular, forms of graph Laplacians greatly differs when applied to data sampled from manifolds of different dimensions. We investigate the standard unnormalized and normalized graph Laplacians, and show that the unnormalized graph Laplacian and the associated unnormalized Dirichlet energy, asymptotically only see the variations within the manifold of the highest dimension, Theorems 2.5 and 2.6. On the other hand, we prove that the normalized Dirichlet energy converges to a Dirichlet energy on the union of manifolds that adapts to all dimensions simultaneously, Theorems 2.2 and 2.3.

Our results have important consequences to the understanding of graph Laplacian based machine learning algorithms. Namely the following are desirable property of algorithms: (i) they capture information about parts of all dimensions present and (ii) separate smooth manifolds that intersect at a positive angle, as these are thought to represent different data clusters. We show that the standard normalized graph Laplacians has both of these properties, except in the case that intersecting manifolds have the same intrinsic dimension and that the intersection is of codimension one, when (ii) fails. We note that having such intersection would be a very unusual situation for data lying in a high-dimensional ambient space, as is typically the case for images and other data classes. On the other hand, the standard, unnormalized, graph Laplacian only sees the variations in data of the highest dimension. Thus, our results provide a strong argument for using the normalized graph Laplacian over the unnormalized one.

Outline. We review the literature on convergence of graph Laplacians in Section 1.2.1 and on convergence of nonlocal Laplacians in Section 1.2.2. We introduce the setup and state the main results in Section 2. In particular, we describe the geometric setup in Section 2.1, graph construction in Section 2.2, introduce the graph and nonlocal Dirichlet energies in Section 2.3, and state the results on Γ-convergence (Theorem 2.2) and compactness (Theorem 2.1) for graph-based Dirichlet energies in Section 2.4. The result on spectral convergence of the normalized graph Laplacian (Theorem 2.3) is stated and proved in Section 2.5, while the one for the unnormalized graph Laplacian (Theorem 2.6) is established in Section 2.6. In Section 2.7 we illustrate our theoretical findings with numerically computed graph Laplacian spectra. Results on Γ-convergence (Theorem 3.2) and compactness (Theorem 3.1) for nonlocal Dirichlet energies are presented and proved in Section 3. Finally, the proofs of our main results, Theorem 2.1 and Theorem 2.2, are presented in Section 4.

In the appendix we establish or recall some the auxiliary facts. Specifically in Section A.1 we characterize the angle between intersecting manifolds, in Section A.2 we recall the TL^p metric that allows one to compute the discrete and continuum setting, in Section A.3 we recall existing results on Γ-convergence of graph Laplacians for data on manifolds, in Section A.4 we show that the boundedness of the energies as $n \to \infty$ in the codimension-one case implies that the desired trace condition is satisfied at the intersection, in Section A.5 we establish density of smooth functions in Sobolev space H^1 on union of intersecting manifolds in the setting where the manifolds are of same dimension and the intersection is a codimension-one submanifold, and in Section A.6 we show the density of Lipschitz functions in the same space, but for the case that the codimension of the intersection in at least one of the manifolds is two or more.

1.2 Related literature

1.2.1 Graph Laplacians

Graph Laplacians play an important role in machine learning. Most of data encountered in practice are high dimensional: signals include many measurements and images have many pixels. Fortunately, most data possess some intrinsic low-dimensional structure that makes machine learning tasks tractable with a limited amount of data. To access and encode the low-dimensional structure in the data, a large family of algorithms is based on first creating a graph by connecting nearby data points. The graphs provide an implicit description of the geometry of the data which, in turn, enables computationally efficient approaches to many tasks. In this work, we study commonly used weighted geometric graphs, where edge weights are a function of the distance between data points, which are the vertices of the graph, as described in Section 2.2.

One of the key objects associated to such graphs is the graph Laplacian which plays a key role in a variety of machine learning methods such as spectral clustering [6, 7, 41, 44], diffusion maps [23], and graph neural networks [10, 24, 57]. The main forms of Laplacians considered are the unnormalized graph Laplacian (2.18), the normalized graph Laplacian (2.13), and the random walk Laplacian which is spectrally equivalent to the normalized one (see [41]). The quadratic forms associated to these operators are the unnormalized graph Dirichlet energy (2.17) and the normalized graph Dirichlet energy (2.6).

Providing theoretical guaranties for such methods spurred on intense research that has seen remarkable progress over the past two decades. The works considered the limit of these and related Laplacians under the assumption that data lie in an open set in Euclidean space or on a smooth manifold embedded in the Euclidean space. Early works established criteria and error estimates for convergence of graph Laplacians towards the weighted Laplacians, when applied to a fixed smooth function [36, 45], a notion known as pointwise convergence. Subsequently, based on the techniques developed in [32], Γ-convergence of graph Dirichlet energies, which we recall in Theorem A.4, was established on Euclidean domains [33] and manifolds [40]. Spectral convergence establishes criteria for the eigenvalues and eigenvectors of graph Laplacians to converge to eigenvalues and eigenfunctions of weighted Laplacians or weighted Laplace–Beltrami operators in manifold setting, respectively. It is implied by Γ-convergence via the use of the Courant–Fischer formula for eigenpairs [27, 33], but can also be established using operator perturbation estimates, see [8, 14, 42, 46]. A number of recent works established error estimates on the convergence of eigenvalues and eigenvectors of graph Laplacians [4, 15, 22, 28, 30, 55].

The works above make the manifold hypothesis, namely assume that data lie on a low-dimensional manifold in the high-dimensional ambient space, or study the simpler case that the data lie in an open set. As we argue in the introduction, real data are often comprised of parts that have different intrinsic dimensions. To come closer to capturing the structure of such data, in this paper we consider the case that data lie on a union of, potentially intersecting, manifolds whose intrinsic dimension can differ. There are only a few papers in the literature that investigate graph Laplacians on samples of data from intersecting manifolds.

Ullrich [50] considers glued manifolds of differing dimensions. However, his goal and objects of interest are different. Namely, we show that on unions of intersecting manifolds (similar results would hold for glued ones) if the dimensions differ, the limiting problem reduces to the one of a disjoint union of the component manifolds. In particular, the resulting diffusion would remain on the manifold it started on. On the other hand, Ullrich studies how the Laplacian should be reweighted near the intersection (meaning that the diffusion is accelerated) so that the resulting process would visit both manifolds. He is able to obtain precise conditions on the scaling of the weights with respect to the spaces involved for the mixing to happen. We remark that from the perspective of machining learning tasks such as clustering and classification, it is desirable if the manifolds representing different clusters would be seen as disjoint in the limit.

García Trillos, He, and Li [29] study the asymptotic behavior of graph Laplacians on samples of intersecting manifolds with the goal of modifying the graph construction so that the limiting problem would be one on the disjoint union of manifolds. In particular, they consider graph constructions that have few connections between manifolds ("sparse outer connectivity" in their terminology). They show that for such graphs the unnormalized graph Laplacian on unions of manifolds of the same dimension indeed spectrally converges to the (weighted) Laplace–Beltrami operator on the disjoint union of manifolds. Furthermore, they provide error estimates for eigenvalues and eigenvectors. When the manifold dimensions differ, they showed that the limit problem only sees the manifolds of the highest dimension, similarly to the limit in such situation in Theorem 2.6. Finally, they provide an example of the graph construction that satisfies the conditions needed. There, nodes in "annular proximity graphs with angle constraints" are connected by an edge if there exists a relatively straight path in a standard random geometric graph between them.

1.2.2 Nonlocal functionals

The connection between nonlocal functionals and Dirichlet forms has been explored extensively in the literature [2, 26, 37, 38, 48, 49]. Korevaar and Schoen [37] introduced the functional, now known as the Korevaar–Schoen energy,

$$E_\varepsilon^{KS}(u) := \int_{\mathcal{M}} \fint_{B(x,\varepsilon)} \frac{(u(x)-u(y))^2}{\varepsilon^2} \, \mathrm{d}\mu(y) \, \mathrm{d}\mu(x) \tag{1.1}$$

in order to consider Sobolev spaces on spaces lacking smoothness. We note that the functional can be rewritten as

$$E_\varepsilon^{KS}(u) = \frac{1}{2\varepsilon^2}\int_{\mathcal{M}}\int_{\mathcal{M}}\left(\frac{1}{\eta_\varepsilon(|x-\cdot|)\star\mu} + \frac{1}{\eta_\varepsilon(|y-\cdot|)\star\mu}\right)\eta_\varepsilon(|x-y|)(u(x)-u(y))^2\,\mathrm{d}\mu(y)\,\mathrm{d}\mu(x) \tag{1.2}$$

for $\eta(s) = \chi_{[0,1]}$. The structure of this energy is similar to that of the normalized nonlocal Dirichlet energy (2.8).

In the sequence of works that followed, Sturm [48, 49] and Kumagai–Sturm [38] used the Korevaar–Schoen energy and related nonlocal Dirichlet energies to define Dirichlet forms on spaces lacking manifold structure, such as fractals (this required replacing ε^2 by a more general function $h(\varepsilon^2)$). In particular, Dirichlet forms are defined as Γ-limits of nonlocal energies along sequences. The limiting form may depend on the sequence considered. If the measures considered satisfy a doubling property, the authors are able to obtain more detailed properties of the resulting forms. Alonso and Baudoin [2] have studied the Dirichlet energies on Cheeger spaces and shown compactness of sequences and have improved Γ-convergence to Mosco convergence. Cheeger spaces are metric measure spaces where the measure satisfies a doubling property and where Poincaré inequality with a Lipschitz semi-norm replacing the modulus of the gradient holds. We remark that in all of these works the Γ-convergence follows from abstract results, and the limit may depend on the subsequence taken.

The contribution of our work in the context of Dirichlet forms on general spaces is that we provide a precise characterization of the limiting form via classical (gradient-based) Dirichlet forms on individual manifolds. This provides detailed information and useful examples of Dirichlet forms on metric measure spaces which do not satisfy measure-doubling property in general. In particular, when two manifolds $\mathcal{M}^{(1)}, \mathcal{M}^{(2)} \subset \mathbb{R}^d$ embedded in $\mathbb{R}^d$ have different dimensions and their intersection $\mathcal{M}^{(12)} := \mathcal{M}^{(1)} \cap \mathcal{M}^{(2)}$ is non-empty, their union is a metric measure space (equipped with the volume measure inherited from the volume measures of the individual manifolds) which does not satisfy the volume doubling property. To see this, consider points x in $\mathcal{M}^{(2)}$ at distance r from $\mathcal{M}^{(12)}$ and compare $\mu(B(x,r))$ and $\mu(B(x,2r))$ while letting $r \to 0$. We also note that our limiting form does not depend on the sequence considered. Finally, our compactness result Theorem 3.1 is new.

2 Setup and main results

2.1 Geometry of the space

For notational simplicity we just consider a space which is a union of two intersecting manifolds. Let $\mathcal{M} := \mathcal{M}^{(1)} \cup \mathcal{M}^{(2)}$ be the union of two compact, smooth Riemannian manifolds without boundary of dimension $d^{(i)}$, $i = 1, 2$, embedded in Euclidean space, $\mathcal{M}^{(i)} \subset \mathbb{R}^N$. Without a loss of generality we will assume in throughout the paper that $1 \leq d^{(1)} \leq d^{(2)}$. We assume that the manifolds intersect in a *nondegenerate* way, which means that their intersection $\mathcal{M}^{(12)} := \mathcal{M}^{(1)} \cap \mathcal{M}^{(2)}$ is a manifold of dimension $d^{(12)} < d^{(1)}$ embedded in $\mathbb{R}^N$ and for every $x \in \mathcal{M}^{(12)}$ the dimension of $\operatorname{span}(T_x\mathcal{M}^{(1)} \cup T_x\mathcal{M}^{(2)})$ is $d^{(1)} + d^{(2)} - d^{(12)}$. We note that the smallest principal angle

$$\begin{aligned}\theta := \min_{x\in\mathcal{M}^{(12)}} \Big\{\angle(u,v)\ : u \in T_x\mathcal{M}^{(1)} \cap (T_x\mathcal{M}^{(1)} \cap T_x\mathcal{M}^{(2)})^\perp,\ u \neq 0 \\ v \in T_x\mathcal{M}^{(2)} \cap (T_x\mathcal{M}^{(1)} \cap T_x\mathcal{M}^{(2)})^\perp,\ v \neq 0\Big\},\end{aligned} \tag{2.1}$$

is positive. Here, $\angle(u,v) := \arccos\frac{|\langle u,v\rangle|}{|u||v|}$ is the angle between two vectors $u, v \in \mathbb{R}^N$. We point to Theorem A.1 in the appendix for an important consequence of the positivity of the principal angle.

We consider random samples of a probability measure supported on $\mathcal{M}$. More precisely, consider probability measures $\mu^{(i)} \in \mathcal{P}(\mathcal{M}^{(i)})$ supported in the manifolds $\mathcal{M}^{(i)}$, for $i = 1, 2$. We assume that $\mu^{(i)}$ has a density $\rho^{(i)}$ with respect to the volume measure $\mathrm{Vol}^{(i)}$ on $\mathcal{M}^{(i)}$ and as a consequence it holds $\mu^{(i)}(\mathcal{M}^{(12)}) = 0$. We pose the following set of standard assumptions for the densities:

Assumption 1. The probability densities $\rho^{(i)}$ with $i \in \{1, 2\}$ satisfy the following:

(i) $\rho^{(i)}$ is Lipschitz continuous on $\mathcal{M}^{(i)}$.

(ii) $\rho^{(i)}$ is bounded from above and below by positive numbers: $\frac{1}{C_\rho} \leq \rho^{(i)} \leq C_\rho$ for some $C_\rho > 0$.

To describe the proportions of samples from each of the two manifolds, we introduce numbers $\alpha^{(i)} > 0$ for $i = 1, 2$, satisfying $\alpha^{(1)} + \alpha^{(2)} = 1$. We use them to define a mixture model for the union of manifolds by defining the probability measure $\mu \in \mathcal{P}(\mathcal{M})$ via $\mu(A) := \sum_{i=1}^{2} \alpha^{(i)} \mu^{(i)}(A \cap \mathcal{M}^{(i)})$ for Borel subsets $A \subset \mathcal{M}$. The measure μ has density ρ with respect to the volume measure Vol on $\mathcal{M}$, defined via $\mathrm{Vol} \llcorner_{\mathcal{M}^{(i)}} := \mathrm{Vol}^{(i)}$, where

$$\rho|_{\mathcal{M}^{(i)}} := \alpha^{(i)} \rho^{(i)} \quad \text{for } i \in \{1, 2\}. \tag{2.2}$$

Let V_n be a set of n *i.i.d.* samples of the probability measure μ. Let $n^{(i)}$ be the number of samples that lie in $\mathcal{M}^{(i)}$ for $i = 1, 2$. Note that $n^{(i)} \sim \mathrm{Binomial}(n, \alpha^{(i)})$ and hence the points in $V_n \cap \mathcal{M}^{(i)}$ constitute a so-called Binomial point process. Then, almost surely, $n^{(1)} + n^{(2)} = n$ and by the law of large numbers

$$\lim_{n \to \infty} \frac{n^{(i)}}{n} = \alpha^{(i)} \qquad i \in \{1, 2\}. \tag{2.3}$$

2.2 Graph construction

Given a set of n points V_n in $\mathbb{R}^N$ (samples of $\mathcal{M}$), a weight profile $\eta : [0, \infty) \to [0, \infty)$ and a bandwidth parameter $\varepsilon > 0$, we construct a weighted graph whose vertices are points in V_n and where the vertices $x, y \in V_n$ are connected by an edge with the weight $\eta_\varepsilon(|x - y|)$. Here $|\cdot|$ denotes the Euclidean distance in $\mathbb{R}^N$. We define $\eta_\varepsilon(\cdot) = \eta(\cdot/\varepsilon)$. We remark that the weight function $\eta_\varepsilon(|x - y|)$ will be the only quantity which explicitly depends on the ambient Euclidean metric in $\mathbb{R}^N$.

We note that our definition of η_ε does not contain a dimension-dependent scaling of the bandwidth ε. Many papers include such scaling, as it is needed when working with the unnormalized graph Dirichlet energy or Laplacian. However, such scaling is not relevant for the normalized energies we consider. When dealing with the unnormalized Laplacian we explicitly include the dimensional scaling in the statement of Theorem 2.6. We impose the following requirements on the weight function η that are standard in the field of graph-based learning.

Assumption 2. The function $\eta : [0, \infty) \to [0, \infty)$ satisfies:

(i) $t \mapsto \eta(t)$ is non-increasing.

(ii) $\eta(0) > 0$ and η is continuous at 0.

(iii) $\operatorname{supp} \eta = [0, 1]$.

We first remark that condition (iii) can be, trivially, replaced by the requirement that the support is compact. Moreover, condition (iii) could be replaced by a decay condition at infinity. However, this would necessitate estimating the tail terms and hence make the proofs more complicated. For $i = 1, 2$ we define the following moments

$$\sigma_\eta^{(i)} := \int_{\mathbb{R}^{d^{(i)}}} \eta(|x|)|x_1|^2 \,\mathrm{d}x, \tag{2.4a}$$

$$\beta_\eta^{(i)} := \int_{\mathbb{R}^{d^{(i)}}} \eta(|x|) \,\mathrm{d}x, \tag{2.4b}$$

which are finite if η satisfies Assumption 2. Finally, we let

$$\deg_{n,\varepsilon}(x) := \frac{1}{n} \sum_{y \in V_n} \eta_\varepsilon(|x-y|) \tag{2.5}$$

denote the weighted degree of x, or in other words, the average of all edge weights adjacent to x.

2.3 Discrete and continuum normalized Dirichlet energies

For $\varepsilon > 0$, on the graph with vertices V_n and edge weights $\eta_\varepsilon(|x-y|)$, the discrete normalized Dirichlet energy of a function $u : V_n \to \mathbb{R}$ is defined by

$$E_{n,\varepsilon}(u) := \frac{1}{n^2\varepsilon^2} \sum_{x,y \in V_n} \eta_\varepsilon(|x-y|) \left| \frac{u(x)}{\sqrt{\deg_{n,\varepsilon}(x)}} - \frac{u(y)}{\sqrt{\deg_{n,\varepsilon}(y)}} \right|^2. \tag{2.6}$$

The empirical measure of points in V_n is the probability measure

$$\mu_n := \frac{1}{n} \sum_{x \in V_n} \delta_x$$

Since $n^{(i)}$ are both positive with (very) high probability, μ_n can be written as a convex combination of two probability measures

$$\mu_n = \sum_{i=1}^{2} \frac{n^{(i)}}{n} \mu_n^{(i)} \qquad \text{where } \ \mu_n^{(i)} := \frac{1}{n^{(i)}} \sum_{x \in V_n \cap \mathcal{M}^{(i)}} \delta_x,$$

are the empirical measures of the data points that lie in $\mathcal{M}^{(i)}$ for $i \in \{1, 2\}$. We can rewrite and reinterpret $E_{n,\varepsilon}$ as functional on $L^2(\mu_n)$ as follows

$$E_{n,\varepsilon}(u) := \frac{1}{\varepsilon^2} \int_{\mathcal{M}} \int_{\mathcal{M}} \eta_\varepsilon(|x-y|) \left| \frac{u(x)}{\sqrt{\eta_\varepsilon(|x-\cdot|) \star \mu_n}} - \frac{u(y)}{\sqrt{\eta_\varepsilon(|y-\cdot|) \star \mu_n}} \right|^2 \mathrm{d}\mu_n(x)\mathrm{d}\mu_n(y), \tag{2.7}$$

where we used the notation

$$\eta_\varepsilon(|x-\cdot|) \star \mu_n := \int_{\mathcal{M}} \eta_\varepsilon(|x-z|) \,\mathrm{d}\mu_n(z) = \deg_{n,\varepsilon}(x).$$

This interpretation gives rise to a nonlocal functional defined on $L^2(\mathcal{M})$ which arises as large sample limit of $E_{n,\varepsilon}$ and takes the form

$$\mathcal{E}_\varepsilon(u) := \frac{1}{\varepsilon^2} \int_{\mathcal{M}} \int_{\mathcal{M}} \eta_\varepsilon(|x-y|) \left| \frac{u(x)}{\sqrt{\eta_\varepsilon(|x-\cdot|) \star \mu}} - \frac{u(y)}{\sqrt{\eta_\varepsilon(|y-\cdot|) \star \mu}} \right|^2 \mathrm{d}\mu(x)\mathrm{d}\mu(y), \tag{2.8}$$

where $\eta_\varepsilon(|x-\cdot|)\star\mu := \int_{\mathcal{M}} \eta_\varepsilon(|x-z|)\,\mathrm{d}\mu(z)$.

The main result of this paper is the proof that the functionals $E_{n,\varepsilon}$ Γ-converge as $n\to\infty$ and $\varepsilon\to 0$ to a normalized Dirichlet energy $\mathcal{E}: L^2(\mathcal{M})\to[0,\infty]$, defined as

$$\mathcal{E}(u) = \begin{cases} \displaystyle\sum_{i=1}^{2} \frac{\sigma_\eta^{(i)}}{\beta_\eta^{(i)}} \int_{\mathcal{M}^{(i)}} \left|\nabla_{\mathcal{M}^{(i)}}\left(\frac{u}{\sqrt{\alpha^{(i)}\rho^{(i)}}}\right)\alpha^{(i)}\rho^{(i)}\right|^2 \mathrm{dVol}^{(i)} & \text{if } u\in H^1_{\sqrt{\mu}}(\mathcal{M}),\\ +\infty & \text{else.}\end{cases} \tag{2.9}$$

Here, $\alpha^{(i)}$ are as in (2.2), $\sigma_\eta^{(i)}$ as well as $\beta_\eta^{(i)}$ are defined in (2.4), and the Sobolev space $H^1_{\sqrt{\mu}}(\mathcal{M})$ is defined as

$$\begin{aligned} H^1_{\sqrt{\mu}}(\mathcal{M}) := \Bigg\{ & u\in L^2(\mathcal{M}) \,:\, \frac{u|_{\mathcal{M}^{(i)}}}{\sqrt{\rho^{(i)}}} \in H^1(\mathcal{M}^{(i)}) \text{ for } i=1,2 \text{ and} \\ & \text{if } d^{(1)}=d^{(2)}=d^{(12)}+1 \text{ then } \mathrm{Tr}^{(1)}\left(\frac{u}{\sqrt{\alpha^{(1)}\rho^{(1)}}}\right) = \mathrm{Tr}^{(2)}\left(\frac{u}{\sqrt{\alpha^{(2)}\rho^{(2)}}}\right)\Bigg\}, \end{aligned} \tag{2.10}$$

where $\mathrm{Tr}^{(i)}: H^1(\mathcal{M}^{(i)})\to H^{1/2}(\mathcal{M}^{(12)})$ is the trace operator from each manifold $\mathcal{M}^{(i)}$ for $i=1,2$ to the intersection $\mathcal{M}^{(12)}$. Note that the trace condition applies only in the special case that manifolds are of the same dimension and their intersection is a codimension-1 submanifold in each. Then we are asking for the regularity of $u/\sqrt{\rho}$ on the union of the manifolds, while u itself can have a jump discontinuity between the manifolds (if $\alpha^{(1)}\rho^{(1)}$ and $\alpha^{(2)}\rho^{(2)}$ differ of $\mathcal{M}^{(12)}$).

Remark 2.1. We note that if $\eta(s)=c\exp(-s^2/2)$, that is, if η is any constant multiple of the Gaussian with unit variance, then the factor $\frac{\sigma_\eta^{(i)}}{\beta_\eta^{(i)}}=1$ regardless of i (or in other words regardless of the dimension). In particular the weighted Laplacian corresponding to $\mathcal{E}$ does not have dimension-dependent factors. We note that this example does not fit the Assumption 2 (iii), but as we remarked the theoretical results are expected to hold for η that decays sufficiently fast to zero.

For the Γ-convergence to hold we require, as is standard, that $\varepsilon=\varepsilon_n$ converges to zero sufficiently slowly as $n\to\infty$. In particular we assume

Assumption 3. The sequence of graph bandwidths $(\varepsilon_n)_{n\in\mathbb{N}}\subset(0,\infty)$ satisfies

$$\ell_n \ll \varepsilon_n \ll 1, \tag{2.11}$$

where we let $d:=\max\left\{d^{(1)},d^{(2)}\right\}$ denote the maximal dimension of the manifolds and define

$$\ell_n := \begin{cases} \sqrt{\dfrac{\log\log n}{n}} & \text{if } d=1,\\ \left(\dfrac{\log n}{n}\right)^{\frac{1}{d}} & \text{if } d\geq 2.\end{cases}$$

Remark 2.2. Note that with the notation $a_n \ll b_n$ (or equivalently $b_n \gg a_n$) for two sequences of non-negative real numbers $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$ we mean that

$$\lim_{n\to\infty}\frac{a_n}{b_n}=0.$$

We also remark that the condition (2.11) is essentially sharp in terms of scaling, as it is known that even on a single manifold with $d\geq 2$ the functionals do not converge to a desirable limit if $\varepsilon_n \ll \left({}^{\log n}/_{n}\right)^{1/d}$, see [33]. For continuum limits in the challenging percolation regime $\varepsilon_n\sim\left({}^{\log n}/_{n}\right)^{1/d}$ on Poisson point clouds we refer to [12] for the graph infinity Laplacian and to [4] for the spectrum of the graph Laplacian.

Remark 2.3. With probability of order $n^{-\alpha/2}$ for arbitrary $\alpha > 2$ the lower bound for ε_n in (2.11) upper bounds the Wasserstein-∞ distance of the absolutely continuous measure μ and the empirical sample measure μ_n, see [28, 31] for $d \geq 2$ and [47] for the case $d = 1$. As a consequence, there exists a constant $C > 0$ such that with probability one there are maps $T_n : \mathcal{M} \to \mathcal{M}$ which satisfy $(T_n)_\sharp \mu = \mu_n$ and

$$\limsup_{n\to\infty} \frac{\|\mathrm{id} - T_n\|_{L^\infty(\mathcal{M})}}{\ell_n} \leq C. \tag{2.12}$$

2.4 Main results

The following are our main results:

Theorem 2.1 (Compactness)**.** *Under Assumptions 1 to 3 let $u_n \subset L^2(\mu_n)$ be a sequence such that*

$$\sup_{n\in\mathbb{N}} E_{n,\varepsilon_n}(u_n) < \infty, \qquad \sup_{n\in\mathbb{N}} \|u_n\|_{L^2(\mu_n)} < \infty.$$

Then almost surely $(u_n)_{n\in\mathbb{N}}$ possesses a subsequence which converges in the $TL^2(\mathcal{M})$-sense.

We note that the definition of $TL^2(\mathcal{M})$ metric is recalled in Section A.2.

Theorem 2.2 (Γ-convergence)**.** *Under Assumptions 1 to 3 almost surely the functionals E_{n,ε_n}, defined in* (2.6)*, Γ-converge to $\mathcal{E}$, defined in* (2.9)*, in the $TL^2(\mathcal{M})$-sense, meaning that*

1. *Almost surely for all sequences $(u_n) \subset L^2(\mu_n)$ with $u_n \xrightarrow{TL^2} u$ for $u \in L^2(\mathcal{M})$ it holds*

$$\mathcal{E}(u) \leq \liminf_{n\to\infty} E_{n,\varepsilon_n}(u_n);$$

2. *Almost surely for all $u \in L^2(\mathcal{M})$ there exists a so-called recovery sequence $(u_n) \subset L^2(\mu_n)$ with $u_n \xrightarrow{TL^2} u$ and*

$$\limsup_{n\to\infty} E_{n,\varepsilon_n}(u_n) \leq \mathcal{E}(u).$$

Section 4 is devoted to the proofs of these two theorems. In fact, the central difficulties of proving Γ-convergence of the graph functional $E_{n,\varepsilon}$ to $\mathcal{E}$ already arise when one aims to prove Γ-convergence of its nonlocal counterpart $\mathcal{E}_\varepsilon$ defined in (2.8). Since describing a proof of Γ-convergence for these continuum nonlocal functionals allows us to clearly present the main ideas and may be of independent interest, we dedicate Section 3 to nonlocal functionals, where we present arguments for proving their Γ-convergence as $\varepsilon \to 0$.

2.5 Convergence of eigenvalues and eigenvectors of normalized graph Laplacians

The results on Γ-convergence can be directly applied to establish the convergence of eigenvalues and eigenvectors of the normalized graph Laplacian on the samples of the union of manifolds to the eigenvalues and eigenfunctions of the weighted Laplace–Beltrami operator on the union of manifolds. This proof relies on a variational description of the spectral problem via the Courant–Fisher formula and follows, almost verbatim, the proof of convergence of eigenvalues and eigenfunctions in Sections 3.1 and 3.2 of [33]. We therefore just state the result and refer to [33] for technical details.

We first note that the normalized graph Laplacian associated to the Dirichlet energy (2.6) has the following form: For $u : V_n \to \mathbb{R}$ and $x \in V_n$

$$L^N_{n,\varepsilon} u(x) := \frac{2}{n\varepsilon^2} \left(u(x) - \sum_{y \in V_n} \frac{\eta_\varepsilon(|x-y|)}{\sqrt{\deg_{n,\varepsilon}(x)}\sqrt{\deg_{n,\varepsilon}(y)}} u(y) \right) \tag{2.13}$$

In describing the limiting eigenvalue problem we distinguish two cases:

Case 1, $d^{(2)} - d^{(12)} \geq 2$**.** In this case the limiting problem is a pair of eigenvalue problems on individual manifolds; the problems are entirely decoupled. That is the eigenvalue problem for the operator corresponding to functional defined in (2.9) is as follows: For $u \in H^1_{\sqrt{\mu}}(\mathcal{M})$ (which in this case is equivalent to the restriction of u to $\mathcal{M}^{(i)}$ lying in $H^1(\mathcal{M}^{(i)})$), we have for $i = 1, 2$ on $\mathcal{M}^{(i)}$:

$$-\frac{\sigma_\eta^{(i)}}{\beta_\eta^{(i)}} \frac{1}{{\rho^{(i)}}^{\frac{3}{2}}} \operatorname{div}_{\mathcal{M}^{(i)}} \left({\rho^{(i)}}^2 \nabla_{\mathcal{M}^{(i)}} \left(\frac{u^{(i)}}{\sqrt{\rho^{(i)}}} \right) \right) = \lambda u^{(i)}. \tag{2.14}$$

We note that if $(\lambda, u^{(i)})$ is a normalized (in $L^2_{\rho^{(i)}}$) eigenpair for the problem (2.14) considered on the *individual* manifold $\mathcal{M}^{(i)}$, and if we define u to be $u^{(i)}/\sqrt{\alpha^{(i)}}$ on $\mathcal{M}^{(i)}$ and zero on the other manifold, then (λ, u) is a normalized (in L^2_ρ) eigenpair for the full problem (2.14). From the standard elliptic theory (see for example [25, 33]) we know that on each individual manifold the problem (2.14) has a sequence of eigenpairs $(\lambda_k^{(i)}, u_k^{(i)})$ such that $0 = \lambda_1^{(i)} < \cdots \leq \lambda_k^{(i)} \leq \lambda_{k+1}^{(i)} \leq \cdots$ and $\{u_k^{(i)}\}_{k \in \mathbb{N}}$ form an orthonormal basis of $L^2_{\rho^{(i)}}(\mathcal{M}^{(i)})$. Since $L^2_\rho(\mathcal{M}) \cong L^2_{\rho^{(1)}}(\mathcal{M}^{(1)}) \times L^2_{\rho^{(2)}}(\mathcal{M}^{(2)})$, extending $u_k^{(i)}$ by zero to the other manifold, as above, and combining the resulting eigenvectors for (2.14) into one sequence provides an orthornomal basis for $L^2_\rho(\mathcal{M})$. We conclude that there exists a sequence of eigenvalues $0 = \lambda_1 = \lambda_2 < \lambda_3 \leq \lambda_4 \leq \cdots$ which converges to infinity, and a corresponding orthonormal in L^2_ρ sequence of eigenfunctions.

Case 2, $d^{(1)} = d^{(2)} =: d$ **and** $d - d^{(12)} = 1$**.** In this case the limiting eigenvalue problem still takes the form (2.14), just that the fact that $u \in H^1_{\sqrt{\mu}}(\mathcal{M})$ couples the two problems on $\mathcal{M}^{(1)}$ and $\mathcal{M}^{(2)}$. This problem too has a nondecreasing sequence of nonnegative eigenvalues and a corresponding orthonormal basis of eigenfunctions. To establish this it is convenient to restate the Dirichlet form and the eigenvalue problem in terms of $v = \frac{u}{\sqrt{\rho}}$, where ρ is as in (2.2). Then $v \in H^1(\mathcal{M})$ and the Dirichlet form in (2.9) can be restated as

$$\mathcal{E}_v(v) := \begin{cases} \dfrac{\sigma_\eta}{\beta_\eta} \displaystyle\int_{\mathcal{M}} |\nabla_{\mathcal{M}} v|^2 \rho^2 \, \mathrm{dVol} & \text{if } v \in H^1(\mathcal{M}), \\ +\infty & \text{else,} \end{cases} \tag{2.15}$$

where we dropped the indices on σ_η and β_η as they do not depend on i, and $\nabla_{\mathcal{M}}$ is considered to mean $\nabla_{\mathcal{M}^{(i)}}$ for points in $\mathcal{M}^{(i)}$. The eigenvalue problem is

$$-\frac{\sigma_\eta}{\beta_\eta} \frac{1}{\rho^2} \operatorname{div}_{\mathcal{M}} \left(\rho^2 \nabla_{\mathcal{M}} v \right) = \lambda v. \tag{2.16}$$

The proof of the spectral theorem for the Laplacian in Euclidean domains (see for example [25]) readily extends to the present situation. Since this is long, but elementary, we omit the full proof. We just note that $\mathcal{F}_v(v) = \mathcal{E}_v(v) + \|v\|^2_{L^2(\mathcal{M})}$ is a quadratic form that is equivalent to the squared H^1-norm. Thus, the existence of solutions of elliptic problems on $\mathcal{M}$ follows using the Riesz representation lemma. The compactness of the solution operator as a mapping from $H^1(\mathcal{M})$ to $L^2(\mathcal{M})$ follows immediately by Rellich–Kondrachov theorem since if $w \in H^1(\mathcal{M})$, then the restriction of w to $\mathcal{M}^{(i)}$ belongs to $H^1(\mathcal{M}^{(i)})$ which compactly embeds in $L^2(\mathcal{M}^{(i)})$. Thus one can use the Fredholm alternative for compact operators to show that the operator has a complete discrete spectrum.

Theorem 2.3. *Suppose Assumptions 1 to 3 hold. Let $\lambda_k^{(n)}$ be the k-th eigenvalue of L^N_{n,ε_n} and let $u_k^{(n)}$ be the associated eigenfunction, normalized in $L^2(\mu_n)$. Let λ_k be be the k-th eigenvalue for Equation* (2.14)*.*

Then almost surely

(i) (Convergence of eigenvalues) For every $k \in \mathbb{N}$,

$$\lim_{n\to\infty} \lambda_k^{(n)} = \lambda_k.$$

(ii) (Convergence of eigenfunctions) The sequence of eigenfunctions $\{u_k^{(n)}\}_{n\in\mathbb{N}}$ (as well as any of its subsequences) has a convergent subsequence and the limit of any such subsequence is a normalized eigenfunction for Equation (2.14) *corresponding to eigenvalue λ_k.*

The proof of the theorem relies on variational description of the spectrum using the Courant–Fisher formula and the Γ-convergence of the associated energies. This argument is carried out in detail in the proofs of Theorems 1.2 and 1.5 in [33].

2.6 The unnormalized (standard) graph Laplacian: Γ-convergence and spectral convergence

As we indicate in the introduction, we advocate for the use of the normalized graph Dirichlet energy (and hence the associated normalized graph Laplacian) as a tool to preserve information from all manifolds in a multi-manifold setting. In order to be able to contrast the behavior of the normalized graph Laplacian to the unnormalized one, we here study the asymptotic behavior (as $n \to \infty$) of the unnormalized graph Dirichlet energy and the associated graph Laplacian. The main point is that when $d^{(1)} < d^{(2)}$ then only the information about $\mathcal{M}^{(2)}$ is retained in the limit.

Following the setting of Section 2.3, we define the discrete Dirichlet energy of a graph function $u : V_n \to \mathbb{R}$ as

$$F_{n,\varepsilon}(u) := \frac{1}{n^2\varepsilon^2} \sum_{x,y\in V_n} \eta_\varepsilon(|x-y|)\,|u(x)-u(y)|^2 . \tag{2.17}$$

The associated (unnormalized) graph Laplacian has the form: For $x \in V_n$

$$L^U_{n,\varepsilon} u(x) := \frac{2}{n\varepsilon^2} \sum_{y\in V_n} \eta_\varepsilon(|x-y|)\,(u(x)-u(y)) . \tag{2.18}$$

We note that the definitions above differ, by a power of ε, from the definition of unnormalized Dirichlet energy and unnormalized Laplacian in the works that study their convergence, such as [11, 13, 14, 33, 39, 40], cf. Theorem A.4 in the appendix. In those works one would also divide by ε^d where d is the dimension of the space considered. In the setting of the manifolds with different dimension there is no power of ε that we can divide by to properly normalize the functional. Thus we here define the unnormalized Dirichlet energy and the associated Laplacian without the dimension dependent rescaling factor. The proper rescalings are stated and discussed in the results below, namely Theorems 2.4 to 2.6.

The Γ-limit of a suitable rescaling of the energies $F_{n,\varepsilon}$ is given by the following Dirichlet energy $\mathcal{G}$:

$L^2(\mathcal{M}) \to [0, \infty]$:

$$\mathcal{G}(u) = \begin{cases} \sum_{i=1}^{2} \sigma_\eta^{(i)} \int_{\mathcal{M}^{(i)}} |\nabla_{\mathcal{M}^{(i)}} u|^2 \left(\alpha^{(i)} \rho^{(i)}\right)^2 \mathrm{dVol}^{(i)} & \text{if } d^{(1)} = d^{(2)} \text{ and } u \in H^1(\mathcal{M}), \\ \sigma_\eta^{(2)} \int_{\mathcal{M}^{(2)}} |\nabla_{\mathcal{M}^{(2)}} u|^2 \left(\alpha^{(2)} \rho^{(2)}\right)^2 \mathrm{dVol}^{(2)} & \text{if } d^{(1)} < d^{(2)},\ u \in H^1(\mathcal{M}),\ u|_{\mathcal{M}^{(1)}} \equiv const. \\ +\infty & \text{else.} \end{cases} \tag{2.19}$$

Here, since the densities $\rho^{(i)}$ do not appear within the gradient term, the Sobolev space $H^1(\mathcal{M})$ is defined as in (2.10) for $\mu \equiv const.$, i.e.,

$$H^1(\mathcal{M}) := \Big\{ u \in L^2(\mathcal{M}) \ : \ u|_{\mathcal{M}^{(i)}} \in H^1(\mathcal{M}^{(i)}) \text{ for } i = 1, 2 \text{ and} \\ \text{if } d^{(1)} = d^{(2)} = d^{(12)} + 1 \text{ then } \mathrm{Tr}^{(1)} u = \mathrm{Tr}^{(2)} u \Big\}. \tag{2.20}$$

We now turn to Γ-convergence and compactness of unnormalized graph Dirichlet energy.

Theorem 2.4 (Compactness)**.** *Under Assumptions 1 to 3 let $u_n \subset L^2(\mu_n)$ be a sequence such that*

$$\sup_{n \in \mathbb{N}} \varepsilon_n^{-d^{(2)}} F_{n,\varepsilon_n}(u_n) < \infty, \qquad \sup_{n \in \mathbb{N}} \|u_n\|_{L^2(\mu_n)} < \infty.$$

If $d^{(1)} = d^{(2)} = d^{(12)} + 1$, then almost surely $(u_n)_{n \in \mathbb{N}}$ possesses a subsequence which converges in the $TL^2(\mathcal{M})$-sense. If $d^{(1)} < d^{(2)}$, then almost surely $(u_n|_{\mathcal{M}^{(2)}})_{n \in \mathbb{N}}$ possesses a subsequence which converges in the $TL^2(\mathcal{M}^{(2)})$-sense.

Theorem 2.5 (Γ-convergence)**.** *Under Assumptions 1 to 3 almost surely the functionals $\varepsilon_n^{-d^{(2)}} F_{n,\varepsilon_n}$ Γ-converge to $\mathcal{G}$ in the $TL^2(\mathcal{M})$-sense, meaning that*

(i) Almost surely for all sequences $(u_n) \subset L^2(\mu_n)$ with $u_n \xrightarrow{TL^2} u$ for $u \in L^2(\mathcal{M})$ it holds

$$\mathcal{G}(u) \le \liminf_{n \to \infty} \varepsilon_n^{-d^{(2)}} F_{n,\varepsilon_n}(u_n);$$

(ii) Almost surely for all $u \in L^2(\mathcal{M})$ there exists a so-called recovery sequence $(u_n) \subset L^2(\mu_n)$ with $u_n \xrightarrow{TL^2} u$ and

$$\limsup_{n \to \infty} \varepsilon_n^{-d^{(2)}} F_{n,\varepsilon_n}(u_n) \le \mathcal{G}(u).$$

The steps of the proof of the Γ-convergence of the unnormalized graph Dirichlet energy are either completely analogous to the normalized graph Laplacian or are much simpler and follow the standard techniques presented here and in [33]. We omit them for brevity.

We note that if the manifold dimensions are unequal, the Γ-limit $\mathcal{G}$ forces constant functions on the lower-dimensional manifold and just depends on the Dirichlet energy on the higher-dimensional manifold. This behavior is consistent with the spectral convergence for multi-manifold clustering with adapted weight constructions developed in [29].

Remark 2.4. Note that another potential scaling would be to consider $\varepsilon_n^{-d^{(1)}} F_{n,\varepsilon_n}$. However, in this case, we would only get compactness in the $TL^2(\mathcal{M}^{(1)})$ topology and we would lose it on the higher-dimensional manifold $\mathcal{M}^{(2)}$. The corresponding Γ-limit would be the standard Dirichlet energy on $\mathcal{M}^{(1)}$.

Remark 2.5. Note that in the case $d^{(1)} = d^{(2)} = d^{(12)} + 1$ the Γ-limit $\mathcal{G}$ is very similar to $\mathcal{E}$ in (2.9) in that it is a sum of Dirichlet energies on the two manifolds together with trace constraint on their intersection. The differences between the unnormalized and the normalized models become more apparent in case of different dimensions $d^{(1)} < d^{(2)}$, where the Γ-limit $\mathcal{G}$ of the unnormalized energies forces functions to be constant on $\mathcal{M}^{(1)}$ but is otherwise independent of the lower dimensional manifold.

To describe the eigenvalue problem associated to the unnormalized Dirichlet energy consider the following equation for $u \in H^1(\mathcal{M})$, which takes the following form on $\mathcal{M}^{(i)}$ for $i = 1, 2$

$$L^U u := -\frac{\sigma_\eta^{(i)}}{\alpha^{(i)}\rho^{(i)}} \operatorname{div}_{\mathcal{M}^{(i)}} \left((\alpha^{(i)}\rho^{(i)})^2 \nabla_{\mathcal{M}^{(i)}} u \right) = \lambda u \tag{2.21}$$

If $d^{(1)} = d^{(2)}$ the eigenfunctions are nonzero $u \in H^1(\mathcal{M})$ which solve Equation (2.21) on all of $\mathcal{M}$. If $d^{(1)} < d^{(2)}$, then eigenfunctions u are nonzero functions in $H^1(\mathcal{M}) = H^1(\mathcal{M}^{(1)}) \times H^1(\mathcal{M}^{(2)})$, with the operator in the equation on $\mathcal{M}^{(1)}$ replaced by

$$L^U u = \begin{cases} 0 & \text{if } u = \text{const.} \\ \infty & \text{otherwise.} \end{cases}$$

We remark that when $d^{(1)} < d^{(2)}$, $\lambda = 0$ is an eigenvalue of multiplicity 2 with eigenspace consisting of functions which are constant on both $\mathcal{M}^{(1)}$ and $\mathcal{M}^{(2)}$. Furthermore if λ_k is the k-th eigenvalue of this problem and $\lambda_{k,2}$ the k-th eigenvalue of Equation (2.21) considered on $\mathcal{M}^{(2)}$ alone, then for $k > 1$, $\lambda_k = \lambda_{k-1,2}$. Furthermore, the corresponding eigenfunctions satisfy $u_k = u_{k-1,2}$ on $\mathcal{M}^{(2)}$ and $u_k = 0$ on $\mathcal{M}^{(1)}$ for $k > 2$.

Theorem 2.6. *Suppose Assumptions 1 to 3 hold. Let $\lambda_k^{(n)}$ be the k-th eigenvalue of $\varepsilon_n^{-d^{(2)}} L^U_{n,\varepsilon_n}$ and let $u_k^{(n)}$ be the associated eigenfunction, normalized in $L^2(\mu_n)$. Let λ_k be be the k-th eigenvalue for Equation* (2.21)*. Then almost surely*

(i) (Convergence of eigenvalues) For every $k \in \mathbb{N}$,

$$\lim_{n\to\infty} \lambda_k^{(n)} = \lambda_k.$$

(ii) (Convergence of eigenfunctions) The sequence of eigenfunctions $\{u_k^{(n)}\}_{n\in\mathbb{N}}$ (as well as any of its subsequences) has a convergent subsequence and the limit of any such subsequence is a normalized eigenfunction for Equation (2.14) *corresponding to eigenvalue λ_k.*

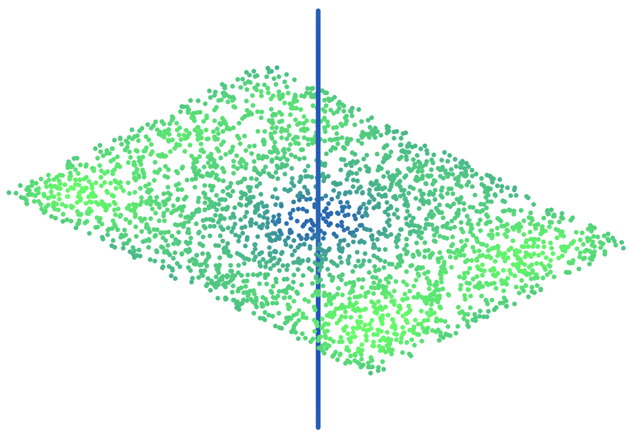
We remark that in the cases of different dimensions a result of similar nature has been obtained in [29, Theorem 10], where the authors also obtain error estimates for eigenvalues and eigenvectors.

Similarly to the proof of Theorem 2.3, the proof of Theorem 2.6 is a straightforward modification of arguments arguments in the proof of [33, Theorems 1.2].

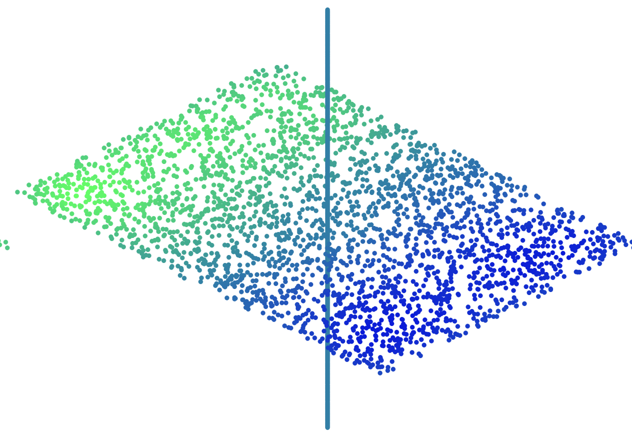
Here we make a remark regarding the eigenvectors corresponding to the first two eigenvalues when $d^{(1)} < d^{(2)}$. Then the limiting problem has eigenspace of dimension two corresponding to eigenvalue zero, spanned by normalized eigenfunctions $u_1 = c\chi_{\mathcal{M}}$ and $u_2 = c_1\chi_{\mathcal{M}^{(1)}} + c_2\chi_{\mathcal{M}^{(2)}}$. Note that $\lambda_1^{(n)} = 0$ with high probability for all n large and thus, due to the connectivity of the graph, $u_1^{(n)}$ is a constant function for all n large. Hence, any subsequential limit of $\{u_1^{(n)}\}_{n=1,2,\ldots}$ belongs to $\pm u_1$. Thus the convergence of eigenvectors of Theorem 2.6 and orthogonality imply that any subsequential limit of $\{u_2^{(n)}\}_{n=1,2,\ldots}$ belongs to $\pm u_2$. In particular we see that $u_2^{(n)}$ separate the manifolds when n is large, as expected. We note that this is not yet the case for n and ε considered in experiments for Figure 2.

2.7 Numerical illustrations

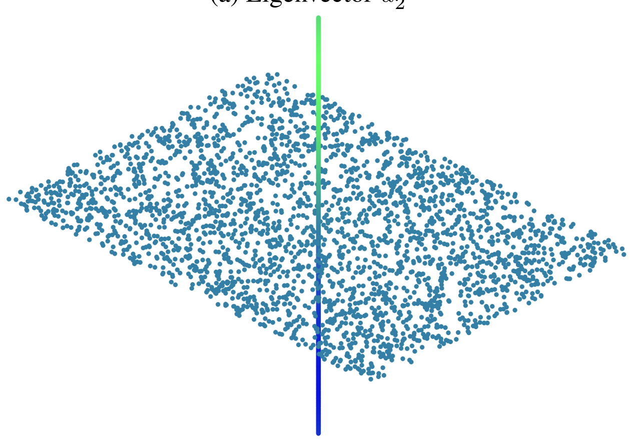
In this section we illustrate some of the findings on a union of a 2D rectangle and a line segment, $\mathcal{M} = \mathcal{M}^{(1)} \cup \mathcal{M}^{(2)}$ where $\mathcal{M}_1 = \{0\} \times \{0\} \times [-0.65, 0.65]$ and $\mathcal{M}_2 = [-0.7, 0.7] \times [-0.5, 0.5] \times \{0\}$. While this does not exactly fit the assumptions of our theorems as $\mathcal{M}^{(1)}$ and $\mathcal{M}^{(2)}$ have a boundary, we note that the theory is expected to carry over to manifolds with boundary where for eigenvalue problem one obtains Neumann conditions at the boundary, as [33] establishes for Euclidean domains.



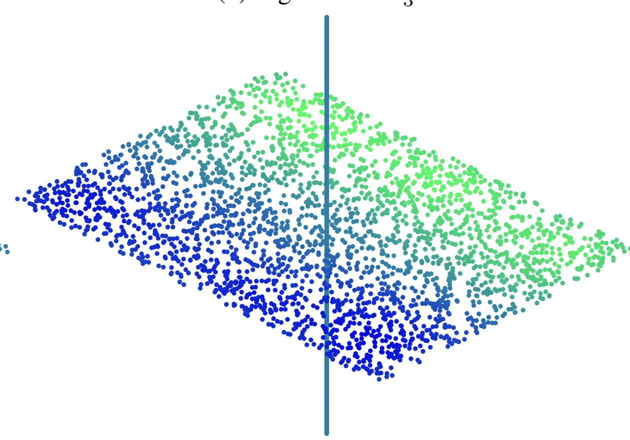
(a) Eigenvector $u_2^{(n)}$



(b) Eigenvector $u_3^{(n)}$



(c) Eigenvector $u_4^{(n)}$



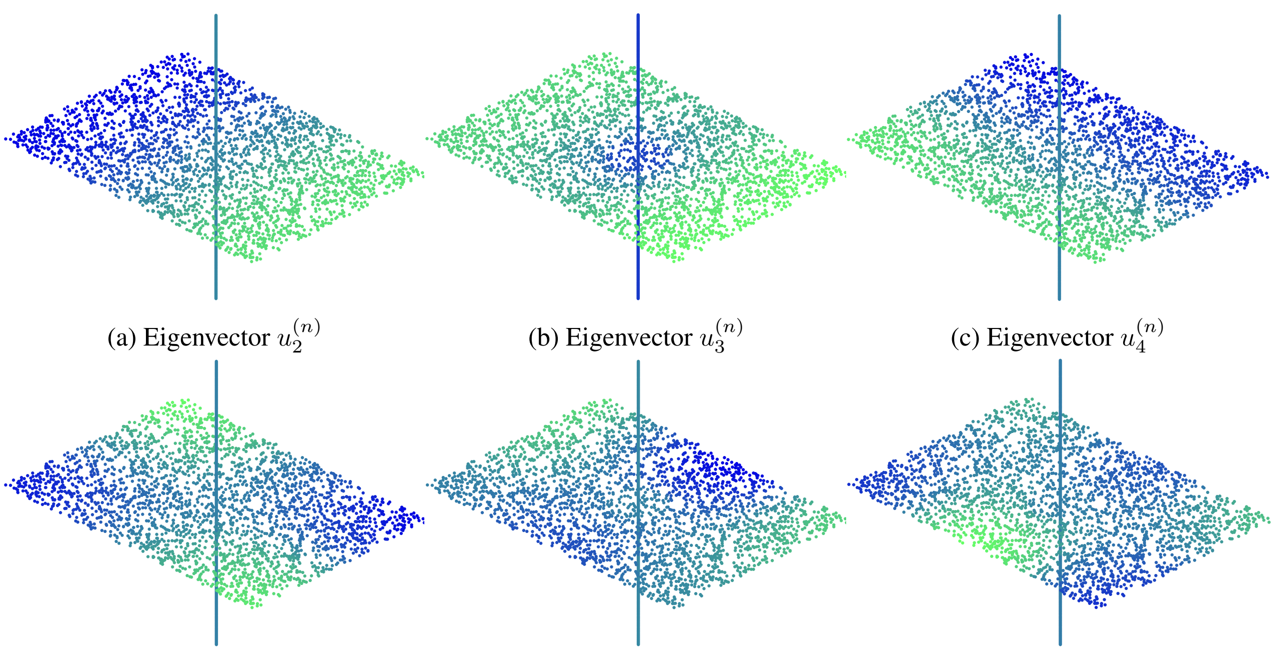
(d) Eigenvector $u_5^{(n)}$

Figure 1: Second to fifth eigenvector of the normalized Laplacian (2.13) for $\eta = \chi_{[0,1]}$, $\varepsilon = 0.13$, $n^{(1)} = 2,600$ and $n^{(2)} = 2,800$.

In our experiments we take $\eta = \chi_{[0,1]}$ and $\varepsilon = 0.13$ and hence use random geometric graphs. Direct computation gives $\sigma^{(1)} = 2/3$, $\beta^{(1)} = 2$, $\sigma^{(2)} = \pi/4$, and $\beta^{(2)} = \pi$. We consider uniform measure on both the rectangle and the line segment. Thus for the normalized Laplacian the first few eigenvalues of the problem on $\mathcal{M}^{(1)}$ are $0, \frac{\pi^2}{3\cdot 1.69}, \frac{4\cdot\pi^2}{3\cdot 1.69}, \ldots$, and for the problem on $\mathcal{M}^{(2)}$ are $0, \frac{\pi^2}{4\cdot 1.96}, \frac{\pi^2}{4}, \frac{\pi^2}{4\cdot 1.96} +$

$\frac{\pi^2}{4}, \ldots$. Comparing the values gives that the eigenvalues for the joint problem Equation (2.15) on $\mathcal{M}$ are $0, 0, \frac{\pi^2}{4 \cdot 1.96}, \frac{\pi^2}{3 \cdot 1.69}, \frac{\pi^2}{4}, \frac{\pi^2}{4 \cdot 1.96} + \frac{\pi^2}{4}, \ldots$. We observe that this is in agreement with Figure 1, where we see that second eigenvector approximates function constant on each manifold and is separating the manifolds, the third and fifth eigenvectors correspond to Laplacian eigenfunction on the rectangle and fourth eigenvector corresponds to the first nonconstant Laplacian eigenfunction on the line segment, where all problems are considered with Neumann boundary conditions. In particular we see that the normalized Laplacian is able to see variations in data of both dimensions. Let us also remark that the fact that for $u_2^{(n)}$ the values on $\mathcal{M}^{(1)}$ are almost constant and the transition between the values $\mathcal{M}^{(1)}$ and the bulk of $\mathcal{M}^{(2)}$ happens in $\mathcal{M}^{(2)}$, close to the intersection is similar to the nature of transition layers constructed in the limsup arguments in Section 3.3 and Section 4.3.

The eigenvectors for the unnormalized graph Laplacian are displayed on Figure 2. We observe that only variations within the rectangle are captured by the first seven eigenvectors. This is in agreement with the theoretical prediction. From the perspective of using unnormalized Laplacians in machine learning tasks, this is undesirable, as the information contained in $\mathcal{M}^{(1)}$ is largely lost (beyond just recognizing that data belong to $\mathcal{M}^{(1)}$). We note that the almost piecewise constant eigenvector corresponds to the third eigenvalue, while our results show that such vectors could converge to the second eigenfunction of the limiting eigenvalue problem as the number of points goes to infinity and $\varepsilon_n \to 0$. The reason for this discrepancy in the nature of the second eigenvector/eigenfunction is that we are still far from the limiting regime when $\varepsilon \to 0$; in particular the transition layers in $u_2^{(n)}$ and $u_3^{(n)}$ on Figure 2 have comparable lengths. From the perspective of machine learning tasks it would be desirable if the second eigenvector was able to distinguish the manifolds. In our experiments, we observed that this is more of an issue for the unnormalized Laplacian than for the normalized Laplacian. Heuristic explanation is that the degrees of the nodes are high near the intersection, so for the normalized Laplacian the differences are suppressed near the intersection making the energy of the transition layer lower.



(a) Eigenvector $u_2^{(n)}$ (b) Eigenvector $u_3^{(n)}$ (c) Eigenvector $u_4^{(n)}$

(d) Eigenvector $u_5^{(n)}$ (e) Eigenvector $u_6^{(n)}$ (f) Eigenvector $u_7^{(n)}$

Figure 2: Second to seventh eigenvector of the unnormalized Laplacian (2.18) for $\eta = \chi_{[0,1]}$, $\varepsilon = 0.13$, $n^{(1)} = 2,600$, and $n^{(2)} = 2,800$.

3 Continuum nonlocal-to-local limit

In this section, we consider nonlocal normalized Dirichlet energies, $\mathcal{E}_\varepsilon$, on unions of manifolds and their limit $\mathcal{E}$ as the bandwidth $\varepsilon \to 0$. This formally corresponds to the case $n = \infty$ in the graph-based setting of the previous section. These results are of independent interest and add to the rich field connecting nonlocal and local energies that we discussed in the introduction. However, our main aim here is to provide the key ideas of the proof of our main results Theorems 2.1 and 2.2, in an accessible way.

Let us first recall the nonlocal and local normalized Dirichlet energies, defined in (2.8) and (2.9), respectively: for $u \in L^2(\mathcal{M})$

$$\mathcal{E}_\varepsilon(u) = \frac{1}{\varepsilon^2}\int_{\mathcal{M}}\int_{\mathcal{M}} \eta_\varepsilon(|x-y|)\left|\frac{u(x)}{\sqrt{\eta_\varepsilon(|x-\cdot|)\star\mu}} - \frac{u(y)}{\sqrt{\eta_\varepsilon(|y-\cdot|)\star\mu}}\right|^2 \mathrm{d}\mu(x)\mathrm{d}\mu(y),$$

$$\mathcal{E}(u) = \begin{cases} \displaystyle\sum_{i=1}^{2}\frac{\sigma_\eta^{(i)}}{\beta_\eta^{(i)}}\int_{\mathcal{M}^{(i)}}\left|\nabla_{\mathcal{M}^{(i)}}\left(\frac{u}{\sqrt{\alpha^{(i)}\rho^{(i)}}}\right)\alpha^{(i)}\rho^{(i)}\right|^2 \mathrm{dVol}^{(i)} & \text{if } u \in H^1_{\sqrt{\mu}}(\mathcal{M}),\\ +\infty & \text{else.}\end{cases}$$

The proofs of Γ-convergence on union of manifolds uses the existing results of Γ-convergence on a single manifold. For the nonlocal-to-local convergence of the nonlocal Dirichlet energy in the continuum setting the relevant Γ-convergence result is Theorem A.3 from [40]. Similarly, the proofs of our main results Theorems 2.1 and 2.2, which we present in Section 4, use Theorem A.4. in the appendix which provides a Γ-convergence and a compactness property for the standard graph Dirichlet energy on a single manifold. More precisely we shall use Theorem A.5 which generalizes Theorem A.4 to Binomial point processes.

We first state the compactness result, which is similar to Theorem 2.1.

Theorem 3.1. *Under Assumptions 1 and 2 for any sequence $(\varepsilon_n)_{n\in\mathbb{N}} \subset (0,\infty)$ converging to zero and any sequence of functions $(u_n)_{n\in N} \subset L^2(\mathcal{M})$ satisfying*

$$\sup_{n\in\mathbb{N}} \mathcal{E}_{\varepsilon_n}(u_n) < \infty, \qquad \sup_{n\in\mathbb{N}} \|u_n\|_{L^2(\mathcal{M})} < \infty,$$

there exists a subsequence of which converges in $L^2(\mathcal{M})$.

The main result of this section is the Γ-convergence of nonlocal normalized Dirichlet energies.

Theorem 3.2. *Suppose that Assumptions 1 and 2 hold and $(\varepsilon_n)_{n\in\mathbb{N}} \subset (0,\infty)$ converges to zero. Then $\mathcal{E}_{\varepsilon_n}$ Γ-converges in $L^2(\mathcal{M})$ to $\mathcal{E}$.*

For proving these theorems it will be convenient to denote by $\mathcal{G}_\varepsilon(\cdot;\mathcal{M}^{(i)})$ and $\mathcal{G}(\cdot;\mathcal{M}^{(i)})$ the nonlocal and the standard Dirichlet energy from Theorem A.3, defined in the manifold $\mathcal{M}^{(i)}$ and equipped with the probability measures $\mu^{(i)} = \rho^{(i)}\mathrm{Vol}^{(i)}$ for $i \in \{1,2\}$.

3.1 Proof of compactness for nonlocal normalized Dirichlet energies

Here we provide a proof of Theorem 3.1, where the main difficulty arises when proving compactness on the larger dimensional manifold $\mathcal{M}^{(2)}$ close to its intersection with $\mathcal{M}^{(1)}$, where the convolutions $x \mapsto \varepsilon_n^{-d^{(2)}}\eta_{\varepsilon_n}(|x-\cdot|)\star\mu$ scale like $\varepsilon_n^{d^{(1)}-d^{(2)}}$ and hence are potentially unbounded.

Proof of Theorem 3.1. Defining the sequence of functions $v_n \in L^2(\mathcal{M})$ via

$$v_n(x) := \frac{u_n(x)}{\sqrt{\varepsilon_n^{-d^{(i)}} \eta_{\varepsilon_n}(|x-\cdot|) \star \mu}}$$

for $x \in \mathcal{M}^{(i)}$ and $i \in \{1,2\}$, and using that $\mu \llcorner \mathcal{M}^{(i)} = \alpha^{(i)}\mu^{(i)}$ we have that

$$\mathcal{E}_{\varepsilon_n}(u_n) \geq \left(\alpha^{(1)}\right)^2 \mathcal{G}_{\varepsilon_n}(v_n;\mathcal{M}^{(1)}) + \left(\alpha^{(2)}\right)^2 \mathcal{G}_{\varepsilon_n}(v_n;\mathcal{M}^{(2)}).$$

In the case that the manifolds have the same dimension $d^{(1)} = d^{(2)} =: d$ we can bound the L^2-norm of v_n by noting that thanks to Assumption 2 there exists $\tau, C_\eta > 0$ such that $\eta(t) \geq C_\eta$ for all $t \leq \tau$ and hence we have

$$\eta_{\varepsilon_n}(|x-\cdot|) \star \mu \geq C_\rho \int_{\mathcal{M}} \eta_{\varepsilon_n}(|x-y|)\,\mathrm{dVol}(y) \geq C_\rho C_\eta \mathrm{Vol}(B_{\tau\varepsilon}(x) \cap \mathcal{M}) \sim \varepsilon^d,$$

where $a(\varepsilon) \sim b(\varepsilon)$ means $0 < \liminf_{\varepsilon\to 0} a/b \leq \limsup_{\varepsilon \to 0} a/b < \infty$. This implies $\|v_n\|_{L^2(\mathcal{M})} \leq C\|u_n\|_{L^2(\mathcal{M})}$ for a constant C that depends on ρ, η and $\mathcal{M}$.

If the dimensions of the manifolds are unequal, we have

$$\varepsilon_n^{-d^{(1)}} \eta_{\varepsilon_n}(|x-\cdot|) \star \mu \sim O(\varepsilon_n^{d^{(2)}-d^{(1)}}) + 1 \sim 1 \tag{3.1}$$

for $x \in \mathcal{M}^{(1)}$ close to the intersection and

$$\varepsilon_n^{-d^{(2)}} \eta_{\varepsilon_n}(|x-\cdot|) \star \mu \sim \varepsilon_n^{d^{(1)}-d^{(2)}} \to \infty \tag{3.2}$$

for $x \in \mathcal{M}^{(2)}$ close to the intersection. Hence, $\|v_n\|_{L^2(\mathcal{M})}$ is bounded in all cases, and thus Theorem A.3 implies that (up to a subsequence that we do not relabel) $v_n \to v$ in $L^2(\mathcal{M})$. Taking (3.1) and the definition of v_n into account, the dominated convergence theorem implies $u_n \to u$ in $L^2(\mathcal{M}^{(1)})$, where

$$u(x) := \sqrt{\alpha^{(i)}\beta_\eta^{(i)}\rho^{(i)}(x)}\, v(x), \qquad \text{for } x \in \mathcal{M}^{(i)}. \tag{3.3}$$

To argue that this convergence also holds true in $L^2(\mathcal{M}^{(2)})$, we first introduce some notation: Let $A_{\varepsilon_n} := \mathcal{M}^{(1)} \cup \{x \in \mathcal{M}^{(2)} : d(x, \mathcal{M}^{(1)}) > \varepsilon_n\}$ and note that on A_{ε_n}, for $x \in \mathcal{M}^{(i)}$

$$\varepsilon_n^{-d^{(i)}} \eta_{\varepsilon_n}(|x-\cdot|) \star \mu \sim 1 \tag{3.4}$$

Define, furthermore, $u_n^{good} := \chi_{A_{\varepsilon_n}} u_n$ and $u_n^{bad} := u_n - u_n^{good}$.

Let us make a few observations: Since $v_n \to v$ in $L^2(\mathcal{M})$, we also have $\chi_{A_{\varepsilon_n}} v_n \to v$ in $L^2(\mathcal{M})$. Furthermore, we note that by standard properties of the convolution (see, e.g., in [56, Theorem 2.2]) we have

$$\varepsilon_n^{-d^{(i)}} \eta_{\varepsilon_n}(|x-\cdot|) \star \mu \llcorner \mathcal{M}^{(i)} \to \alpha^{(i)}\beta_\eta^{(i)}\rho^{(i)}(x) \quad \text{uniformly in } \mathcal{M}^{(i)} \text{ as } n \to \infty. \tag{3.5}$$

Thus, due also to (3.1), the following function on $\mathcal{M}$

$$x \mapsto \chi_{A_{\varepsilon_n}}(x)\varepsilon_n^{-d^{(i)}} \eta_{\varepsilon_n}(|x-\cdot|) \star \mu - \chi_{A_{\varepsilon_n}}(x)\alpha^{(i)}\beta_\eta^{(i)}\rho^{(i)}(x) \tag{3.6}$$

(where i corresponds to the manifold x belongs to) converges uniformly to zero. From these observations, and taking the definition of v_n and (3.4) into account, it follows that u_n^{good} converges in $L^2(\mathcal{M})$ to $u \in$

$L^2(\mathcal{M})$, defined in (3.3). So to prove $L^2(\mathcal{M})$ convergence of u_n (along a subsequence) it remains to show that $u_n^{bad} \to 0$ in $L^2(\mathcal{M}^{(2)})$, at least along a subsequence. Note that we already know $u_n^{bad} \to 0$ in $L^2(\mathcal{M}^{(1)})$ since $u_n \to u$ and $u_n^{good} \to u$ in the same topology.

The idea that guides the remainder of the proof is the following: If the L^2 norm of u_n^{bad} is bounded from below by a positive number, since the energy $\mathcal{E}_{\varepsilon_n}(u_n)$ is small, the cross-term of the energy would also force the L^2 norm of u_n restricted to a vanishing (as $n \to \infty$) region of $\mathcal{M}^{(1)}$ close to the intersection to be bounded from below by a positive number. However, this is not possible since u_n^{good} converges in $L^2(\mathcal{M}^{(1)})$. The precise argument is a bit more complicated since we are not able to execute the above in one step, but instead iterate over an expanding sequence of sets.

Let us define the "good" set $G^1_{\varepsilon_n} := \{x \in \mathcal{M}^{(1)} : d(x, \mathcal{M}^{(2)}) < \varepsilon_n\}$ on which we already know that $u_n \to u$ as well as the "strip" $S^1_{\varepsilon_n} := \{x \in \mathcal{M}^{(2)} : d(x, \mathcal{M}^{(1)}) < \varepsilon_n/2\}$ of points where we want to deduce this convergence. For $k > 1$ we update the good and the strip sets iteratively via $G^k_{\varepsilon_n} := G^{k-1}_{\varepsilon_n} \cup S^{k-1}_{\varepsilon_n}$ and $S^k_{\varepsilon_n} = \{x \in \mathcal{M}^{(2)} : d(x, G^{k-1}_{\varepsilon_n}) < \varepsilon_n/2\}$.

Let $\bar{k} > 1$ be the smallest k such that for all $n \in \mathbb{N}$, $\{x \in \mathcal{M}^{(2)} : d(x, \mathcal{M}^{(1)}) \leq \varepsilon_n\} \subset G^k_{\varepsilon_n}$. We know by Theorem A.1 that $\bar{k}$ is finite, but that depending on the angle of intersection of manifolds may be bigger than two.

We show by induction that $\|u_n\|_{L^2(G^k_{\varepsilon_n})} \to 0$ as $n \to \infty$. Since $\mathcal{M} \setminus A_{\varepsilon_n} = \{x \in \mathcal{M}^{(2)} : d(x, \mathcal{M}^{(1)})\} \leq \varepsilon_n\} \subset G^{\bar{k}}_{\varepsilon_n}$ this implies that $\|u_n^{bad}\|_{L^2(\mathcal{M})} \to 0$, which is the desired conclusion.

Case $k = 1$**.** We note that the fact that $\|u_n\|_{L^2(G^1_{\varepsilon_n})} \to 0$ follows from the fact that u_n restricted to $\mathcal{M}^{(1)}$ converge in L^2.

Case $k > 1$**.** Observe that for all $k > 1$, due to the definition of the sets G^k_ε and S^k_ε, and Assumption 2(iii), we have that for all $x \in S^k_{\varepsilon_n}$, $\eta_{\varepsilon_n}(|x - \cdot|) \star \chi_{G^k_{\varepsilon_n}} > c_k > 0$ and hence $\eta_{\varepsilon_n}(|x - \cdot|) \star (\mu \llcorner G^k_{\varepsilon_n}) \geq C_k \eta_{\varepsilon_n}(|x - \cdot|) \star \mu$ for some $C_k > 0$. Furthermore, for all $x \in \mathcal{M}$ we trivially have that $\eta_{\varepsilon_n}(|x - \cdot|) \star (\mu \llcorner S^k_{\varepsilon_n}) \leq \eta_{\varepsilon_n}(|x - \cdot|) \star \mu$. Therefore

$$
\begin{aligned}
& C_k \|u_n\|^2_{L^2(\mu \llcorner S^k_{\varepsilon_n})} - \|u_n\|^2_{L^2(\mu \llcorner G^k_{\varepsilon_n})} \\
&\leq \int_{S^k_{\varepsilon_n}} \frac{\eta_{\varepsilon_n}(|x - \cdot|) \star (\mu \llcorner G^k_{\varepsilon_n})}{\eta_{\varepsilon_n}(|x - \cdot|) \star \mu} u_n^2(x) \mathrm{d}\mu(x) - \int_{G^k_{\varepsilon_n}} \frac{\eta_{\varepsilon_n}(|y - \cdot|) \star (\mu \llcorner S^k_{\varepsilon_n})}{\eta_{\varepsilon_n}(|y - \cdot|) \star \mu} u_n^2(y) \mathrm{d}\mu(y) \\
&= \int_{S^k_{\varepsilon_n}} \int_{G^k_{\varepsilon_n}} \eta_{\varepsilon_n}(|x - y|) \left(\frac{u_n^2(x)}{\eta_{\varepsilon_n}(|x - \cdot|) \star \mu} - \frac{u_n^2(y)}{\eta_{\varepsilon_n}(|y - \cdot|) \star \mu} \right) \mathrm{d}\mu(y) \mathrm{d}\mu(x) \\
&\leq \sqrt{\int_{S^k_{\varepsilon_n}} \int_{G^k_{\varepsilon_n}} \eta_{\varepsilon_n}(|x - y|) \left| \frac{u_n(x)}{\sqrt{\eta_{\varepsilon_n}(|x - \cdot|) \star \mu}} - \frac{u_n(y)}{\sqrt{\eta_{\varepsilon_n}(|y - \cdot|) \star \mu}} \right|^2 \mathrm{d}\mu(y) \mathrm{d}\mu(x)} \\
&\quad \times \sqrt{\int_{S^k_{\varepsilon_n}} \int_{G^k_{\varepsilon_n}} \eta_{\varepsilon_n}(|x - y|) \left(\frac{u_n(x)}{\sqrt{\eta_{\varepsilon_n}(|x - \cdot|) \star \mu}} + \frac{u_n(y)}{\sqrt{\eta_{\varepsilon_n}(|y - \cdot|) \star \mu}} \right)^2 \mathrm{d}\mu(y) \mathrm{d}\mu(x)} \\
&\leq 2\varepsilon_n \sqrt{\mathcal{E}_{\varepsilon_n}(u_n)} \|u_n\|_{L^2(\mu)}
\end{aligned}
$$

Therefore

$$
\|u_n\|^2_{L^2(\mu \llcorner S^k_{\varepsilon_n})} \leq \frac{1}{C_k} \left(\|u_n\|^2_{L^2(\mu \llcorner G^k_{\varepsilon_n})} + 2\varepsilon_n \sqrt{\mathcal{E}_{\varepsilon_n}(u_n)} \|u_n\|_{L^2(\mu)} \right)
$$

Since, using also the induction assumption, the right-hand side converges to zero as $n \to \infty$, we conclude that $\|u_n\|^2_{L^2(\mu \llcorner S^k_{\varepsilon_n})} \to 0$ and therefore $\|u_n\|^2_{L^2(\mu \llcorner G^{k+1}_{\varepsilon_n})} \to 0$ as $n \to \infty$. □

For the full proof of Theorem 2.1 which follows similar ideas we refer to Section 4.1.

3.2 Proof of the nonlocal Γ-liminf inequality

Here we prove the liminf inequality part of Γ-converence claimed in Theorem 3.2. That is we prove that for any sequence $(u_n) \subset L^2(\mathcal{M})$ with $u_n \xrightarrow{L^2} u$ for $u \in L^2(\mathcal{M})$ it holds

$$\mathcal{E}(u) \leq \liminf_{n\to\infty} \mathcal{E}_{\varepsilon_n}(u_n).$$

The proof follows similar ideas as used in proving Theorem 3.1 in Section 3.1. Here, however, the main challenge consist in establishing the trace condition in the definition of $H^1_{\sqrt{\mu}}(\mathcal{M})$ in the case of codimension 1.

Proof of the liminf inequality for Theorem 3.2. Let $v_n(x) := \frac{u_n(x)}{\sqrt{\varepsilon_n^{-d^{(i)}}\eta_{\varepsilon_n}(|x-\cdot|)\star\mu}}$ as before and $v(x) := \frac{u(x)}{\sqrt{\alpha^{(i)}\beta_\eta^{(i)}\rho^{(i)}(x)}}$ for $x \in \mathcal{M}^{(i)}$ and $i \in \{1,2\}$. We first prove that $v_n \to v$ in $L^2(\mathcal{M})$ by computing

$$\begin{aligned}
&\lim_{n\to\infty}\sum_{i=1}^2 \int_{\mathcal{M}^{(i)}} |v_n - v|^2 \,\mathrm{dVol}^{(i)} \\
&= \lim_{n\to\infty}\sum_{i=1}^2 \int_{\mathcal{M}^{(i)}} \left| u_n(x) - \sqrt{\frac{\varepsilon_n^{-d^{(i)}}\eta_{\varepsilon_n}(|x-\cdot|)\star\mu}{\alpha^{(i)}\beta_\eta^{(i)}\rho^{(i)}(x)}}\,u(x)\right|^2 \frac{\mathrm{dVol}^{(i)}(x)}{\varepsilon_n^{-d^{(i)}}\eta_{\varepsilon_n}(|x-\cdot|)\star\mu} \\
&\leq 2\lim_{n\to\infty}\sum_{i=1}^2 \int_{\mathcal{M}^{(i)}} |u_n(x)-u(x)|^2 \frac{\mathrm{dVol}^{(i)}(x)}{\varepsilon_n^{-d^{(i)}}\eta_{\varepsilon_n}(|x-\cdot|)\star\mu} \\
&\quad + 2\lim_{n\to\infty}\sum_{i=1}^2 \int_{\mathcal{M}^{(i)}} \left| \frac{\sqrt{\alpha^{(i)}\beta_\eta^{(i)}\rho^{(i)}(x)} - \sqrt{\varepsilon_n^{-d^{(i)}}\eta_{\varepsilon_n}(|x-\cdot|)\star\mu}}{\sqrt{\varepsilon_n^{-d^{(i)}}\eta_{\varepsilon_n}(|x-\cdot|)\star\mu}\sqrt{\alpha^{(i)}\beta_\eta^{(i)}\rho^{(i)}(x)}}\right|^2 |u(x)|^2 \,\mathrm{dVol}^{(i)}(x).
\end{aligned}$$

Using that $u_n \to u$ in L^2 and employing the dominated convergence theorem together with (3.1), (3.2) and (3.5) for the second term, we get that $v_n \to v$ in $L^2(\mathcal{M}^{(i)})$.

By definition of v_n and v have

$$\mathcal{E}_{\varepsilon_n}(u_n) \geq \left(\alpha^{(1)}\right)^2 \mathcal{G}_{\varepsilon_n}(v_n;\mathcal{M}^{(1)}) + \left(\alpha^{(2)}\right)^2 \mathcal{G}_{\varepsilon_n}(v_n;\mathcal{M}^{(2)}).$$

and, using the liminf inequality from Theorem A.3, we obtain the inequality

$$\begin{aligned}
\sum_{i=1}^2 \frac{\alpha^{(i)}\sigma_\eta^{(i)}}{\beta_\eta^{(i)}} \int_{\mathcal{M}^{(i)}} \left| \nabla_{\mathcal{M}^{(i)}}\left(\frac{u}{\sqrt{\rho^{(i)}}}\right)\rho^{(i)}\right|^2 \mathrm{dVol}^{(i)} &= \left(\alpha^{(1)}\right)^2 \mathcal{G}_{\varepsilon_n}(v;\mathcal{M}^{(1)}) + \left(\alpha^{(2)}\right)^2 \mathcal{G}_{\varepsilon_n}(v;\mathcal{M}^{(2)}) \\
&\leq \liminf_{n\to\infty} \mathcal{E}_{\varepsilon_n}(u_n).
\end{aligned}$$

To conclude the proof of the liminf inequality, it only remains to show the trace condition in the definition of the Sobolev space $H^1_{\sqrt{\mu}}(\mathcal{M})$ from (2.10) in the case $d^{(1)} = d^{(2)} =: d = d^{(12)} + 1$. In this case the intersection $\mathcal{M}^{(12)}$ locally divides the remaining manifolds into parts $\mathcal{M}^{(i)}_\pm$, one on either side of $\mathcal{M}^{(12)}$, see Figure 3. In particular, there exists a bi-Lipschitz change of variables, Ψ, mapping the intersection of a ball around a point $x \in \mathcal{M}^{(12)}$ with $\mathcal{M}^{(1)}_+ \cup \mathcal{M}^{(2)}_+$ (the red set on the left) to an open subset of $\mathbb{R}^d$ (the red set on the right). Then the liminf inequality from Theorem A.3 implies that $u \circ \Psi^{-1}$ is a H^1-function and hence has a well-defined trace on $\Psi(\mathcal{M}^{(12)})$. Thus the two traces of u $\mathcal{M}^{(12)}$ coincide on the neighborhood of x. Since x was arbitrary the traces agree on all of $\mathcal{M}^{(12)}$.

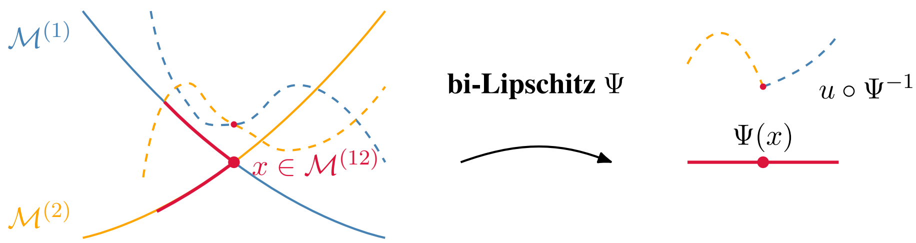


Figure 3: Validity of the trace condition as proved in Theorem A.6. The "one-sided" neighborhoods of x in $\mathcal{M}^{(1)}$ and $\mathcal{M}^{(2)}$ (red) are flattened together by Ψ. The dotted lines represent the values of the function u (left plot) and $u \circ \Psi^{-1}$ (right plot).

This idea is made rigorous in Theorem A.6 in the appendix. Applying it to v_n and v as defined above and assuming, without loss of generality, that $\liminf_{n\to\infty} \mathcal{E}_{\varepsilon_n}(u_n) < \infty$ we obtain that $\mathrm{Tr}^{(1)}(v) = \mathrm{Tr}^{(2)}(v)$ on $\mathcal{M}^{(12)}$ and therefore, using also that $\beta_\eta^{(1)} = \beta_\eta^{(2)}$, we have

$$\mathrm{Tr}^{(1)}\left(\frac{u}{\sqrt{\alpha^{(1)}\rho^{(1)}}}\right) = \mathrm{Tr}^{(2)}\left(\frac{u}{\sqrt{\alpha^{(2)}\rho^{(2)}}}\right).$$

This implies $u \in H^1_{\sqrt{\mu}}(\mathcal{M})$ and concludes the proof. □

Remark 3.1. We note that the argument does not require that $u_n \to u$ in $L^2(\mathcal{M})$, but only that

$$\int_{\mathcal{M}^{(i)}} |u_n(x) - u(x)|^2 \frac{\mathrm{dVol}^{(i)}(x)}{\varepsilon_n^{-d^{(i)}} \eta_{\varepsilon_n}(|x - \cdot|) \star \mu} \to 0, \quad n \to \infty,$$

for $i = 1, 2$, which is strictly weaker if $d^{(1)} < d^{(2)}$ since $\varepsilon_n^{-d^{(i)}} \eta_{\varepsilon_n}(|x - \cdot|) \star \mu$ is bounded from below by a positive number and, by (3.2), on $\mathcal{M}^{(2)}$ near $\mathcal{M}^{(1)}$ diverges like $\varepsilon_n^{d^{(1)}-d^{(2)}} \gg 1$.

We refer to Section 4.2 for the proof of the liminf inequality for Theorem 2.2.

3.3 Proof of nonlocal Γ-limsup inequality

Finally, we discuss the proof of the limsup inequality for the Γ-convergence claimed in Theorem 3.2, namely that for all $u \in L^2(\mathcal{M})$ there exists a so-called recovery sequence $(u_n) \subset L^2(\mathcal{M})$ such that $u_n \overset{L^2}{\to} u$ and

$$\limsup_{n\to\infty} \mathcal{E}_{\varepsilon_n}(u_n) \le \mathcal{E}(u).$$

The proof requires considering a few distinct cases with respect to the codimension of the intersection manifold $\mathcal{M}^{(12)}$. In order to avoid repetition, we do not present the argument for the critical codimension $d^{(2)} - d^{(12)} = 2$ in full generality but stick to a less technical description of the key ideas in this section. Note that later we will give the full proof of the discrete limsup inequality in Theorem 4.10.

Detailed proof sketch of the limsup inequality for Theorem 3.2. Here we go over the key steps and main difficulties in going from the standard argument on one manifold (see for example [40]) to the union of manifolds.

Smooth approximations. For proving the limsup inequality it is advantageous to reduce the argument to creating recovery sequences for the energy $\mathcal{E}(u)$ for u sufficiently regular, e.g., Lipschitz continuous. If $d^{(1)} = d^{(2)} = d^{(12)}+1$ one needs to establish the density of such functions among H^1 functions on the union of manifolds that respect the trace condition (2.10). This requires a detailed argument, which we present in Theorem A.7 in the appendix. If $d^{(2)} \geq d^{(12)} + 2$ the trace condition is of no interest and smoothing can be easily done, since the energy $\mathcal{E}$ decouples the two manifolds, cf. (2.9), and we can use the density of smooth functions in H^1 on each manifold separately. Hence, one can work with regular u on both of the manifolds and construct individual recovery sequences.

Recovery sequence for codimensions larger than two. When $d^{(1)} - d^{(12)} > 2$, i.e., the codimension of the intersection in both of the manifolds is larger than two, one can use constant recovery sequences. By density we can take u such that restriction of u to each of the manifolds is smooth. Let $v^{(i)} = u/\sqrt{\alpha^{(i)}\beta^{(i)}\rho^{(i)}}$ on $\mathcal{M}^{(i)}$. By the proof of Theorem A.3 (where a constant recovery sequence is taken)

$$\limsup_{n\to\infty} \mathcal{G}_{\varepsilon_n}(v^{(i)};\mathcal{M}^{(i)}) \leq \mathcal{G}\left(v^{(i)};\mathcal{M}^{(i)}\right) \quad \text{for } i \in \{0,1\}. \tag{3.7}$$

Then for $A_\varepsilon := \mathcal{M}^{(1)} \cup \{x \in \mathcal{M}^{(2)} \,:\, d(x,\mathcal{M}^{(1)}) > \varepsilon\}$ and $S_\varepsilon := \mathcal{M} \setminus A_\varepsilon$ we have

$$\begin{aligned}
\mathcal{E}_{\varepsilon_n}(u) &= \frac{1}{\varepsilon_n^2}\int_{\mathcal{M}}\int_{\mathcal{M}} \eta_{\varepsilon_n}(|x-y|)\left|\frac{v(x)\sqrt{\alpha^{(i)}\beta^{(i)}\rho^{(i)}(x)}}{\sqrt{\eta_{\varepsilon_n}(|x-\cdot|)\star\mu}} - \frac{v(y)\sqrt{\alpha^{(i)}\beta^{(i)}\rho^{(i)}(y)}}{\sqrt{\eta_{\varepsilon_n}(|y-\cdot|)\star\mu}}\right|^2 \mathrm{d}\mu(x)\,\mathrm{d}\mu(y)\\
&\leq \frac{1}{\varepsilon_n^2}\int_{A_{\varepsilon_n}}\int_{A_{\varepsilon_n}} \eta_{\varepsilon_n}(|x-y|)\left|\frac{v(x)\sqrt{\alpha^{(i)}\beta^{(i)}\rho^{(i)}(x)}}{\sqrt{\eta_{\varepsilon_n}(|x-\cdot|)\star\mu}} - \frac{v(y)\sqrt{\alpha^{(i)}\beta^{(i)}\rho^{(i)}(y)}}{\sqrt{\eta_{\varepsilon_n}(|y-\cdot|)\star\mu}}\right|^2 \mathrm{d}\mu(x)\,\mathrm{d}\mu(y)\\
&\quad + \frac{2}{\varepsilon_n^2}\int_{\mathcal{M}}\int_{S_{\varepsilon_n}} \eta_{\varepsilon_n}(|x-y|)\left|\frac{v(x)\sqrt{\alpha^{(i)}\beta^{(i)}\rho^{(i)}(x)}}{\sqrt{\eta_{\varepsilon_n}(|x-\cdot|)\star\mu}} - \frac{v(y)\sqrt{\alpha^{(i)}\beta^{(i)}\rho^{(i)}(y)}}{\sqrt{\eta_{\varepsilon_n}(|y-\cdot|)\star\mu}}\right|^2 \mathrm{d}\mu(x)\,\mathrm{d}\mu(y)\\
&=: \mathcal{E}^{far}_{\varepsilon_n}(u) + \mathcal{E}^{near}_{\varepsilon_n}(u).
\end{aligned}$$

Thanks to (3.5) and (3.7) we have that

$$\limsup_{n\to\infty} \mathcal{E}^{far}_{\varepsilon_n}(u) \leq \limsup_{n\to\infty} \sum_{i=1}^{2}\left(\alpha^{(i)}\right)^2 \mathcal{G}_{\varepsilon_n}(v^{(i)};\mathcal{M}^{(i)}) \leq \sum_{i=1}^{2}\left(\alpha^{(i)}\right)^2 \mathcal{G}\left(v^{(i)};\mathcal{M}^{(i)}\right) = \mathcal{E}(u).$$

So it remains to show that $\lim_{n\to\infty}\mathcal{E}^{near}_{\varepsilon_n}(u) = 0$. From (3.5), it follows that $\varepsilon_n^{-d^{(i)}}\eta_{\varepsilon_n}(|x-\cdot|)\star\mu \geq \frac{1}{2}\alpha^{(i)}\beta^{(i)}\rho^{(i)}(x)$ on $\mathcal{M}^{(i)}$, for n large enough. We estimate

$$\begin{aligned}
\mathcal{E}^{near}_{\varepsilon_n}(u) &\lesssim \frac{1}{\varepsilon_n^2}\int_{S_{2\varepsilon_n}}\int_{S_{2\varepsilon_n}} \eta_{\varepsilon_n}(|x-y|)\varepsilon_n^{-d^{(2)}}\left(v(x)^2+v(y)^2\right)\mathrm{d}\mu^{(2)}(x)\,\mathrm{d}\mu^{(2)}(y)\\
&\quad + \frac{1}{\varepsilon_n^2}\int_{\mathcal{M}^{(2)}}\int_{\mathcal{M}^{(1)}} \eta_{\varepsilon_n}(|x-y|)\left(\varepsilon_n^{-d^{(1)}}v(x)^2+\varepsilon_n^{-d^{(2)}}v(y)^2\right)\mathrm{d}\mu^{(1)}(x)\,\mathrm{d}\mu^{(2)}(y)\\
&\lesssim \frac{C_\rho\|v\|_\infty^2}{\varepsilon_n^2}\left(\mu^{(2)}(S_{2\varepsilon_n}) + \varepsilon_n^{d^{(2)}-d^{(1)}}\mu^{(1)}\left(\{x\in\mathcal{M}^{(1)} \,:\, d(x,\mathcal{M}^{(2)})<\varepsilon_n\}\right) + \varepsilon_n^{d^{(1)}-d^{(2)}}\mu^{(2)}(S_{\varepsilon_n})\right)\\
&\lesssim \frac{\|v\|_\infty^2}{\varepsilon_n^2}\left(\varepsilon_n^{d^{(2)}-d^{(12)}} + \varepsilon_n^{d^{(2)}-d^{(1)}}\varepsilon_n^{d^{(1)}-d^{(12)}} + \varepsilon_n^{d^{(1)}-d^{(2)}}\varepsilon_n^{d^{(2)}-d^{(12)}}\right)\\
&\lesssim \|v\|_\infty^2\left(\varepsilon_n^{d^{(2)}-d^{(12)}-2} + \varepsilon_n^{d^{(1)}-d^{(12)}-2}\right) \to 0 \qquad \text{as } n\to\infty.
\end{aligned}$$

Remark 3.2. We remark that the argument above can be extended to the case when $d^{(1)} - d^{(12)} \leq 2$ and $d^{(2)} - d^{(12)} > 2$. Namely it is not hard to show that functions v which are compactly supported in $\mathcal{M}^2 \setminus \mathcal{M}^{(12)}$ are dense in $H^1(\mathcal{M}^2)$ whenever $d^{(2)} - d^{(12)} \geq 2$. Thus, by a density argument, it suffices to consider v which are equal to zero in $S_{2\varepsilon_n}$ for n large enough. Thus in the estimate above we can obtain $\mathcal{E}^{near}_{\varepsilon_n}(u) \lesssim \|v\|^2_{L^\infty} \varepsilon_n^{d^{(2)}-d^{(12)}-2}$ which converges to zero as $n \to \infty$ when $d^{(2)} - d^{(12)} > 2$, regardless of $d^{(1)} - d^{(12)} > 2$.

Recovery sequence for intermediate codimensions. Here we deal with the case that $d^{(1)} - d^{(12)} \leq 2$ and $d^{(2)} - d^{(12)} \geq 2$. The critical case arises if $d^{(2)} - d^{(12)} = 2$, i.e., the larger of the two codimensions is two, where one can think of a one-dimensional manifold piercing a two-dimensional one. Even in this simple case with $d^{(1)} = 1$, $d^{(2)} = 2$, and $d^{(12)} = 0$, if one tries to create a recovery sequence a function $u \in H^1_{\sqrt{\mu}}(\mathcal{M})$ which is constant on each of the manifolds $\mathcal{M}^{(i)}$ with value $u^{(i)} \in \mathbb{R}$ but not constant on $\mathcal{M}$ as a whole, one observes that

$$\begin{aligned}\mathcal{E}_{\varepsilon_n}(u) &= \frac{1}{\varepsilon^2}\int_{\mathcal{M}^{(1)}}\int_{\mathcal{M}^{(2)}} \eta_{\varepsilon_n}(|x-y|)\left|\frac{u^{(1)}}{\sqrt{\eta_{\varepsilon_n}(|x-\cdot|)\star\mu}} - \frac{u^{(2)}}{\sqrt{\eta_{\varepsilon_n}(|x-\cdot|)\star\mu}}\right|^2 \mathrm{d}\mu(x)\,\mathrm{d}\mu(y)\\ &\sim \frac{1}{\varepsilon^2}\varepsilon^{1+2-0}\left|\frac{u^{(1)}-u^{(2)}}{\sqrt{\varepsilon}}\right|^2 \sim \left|u^{(1)}-u^{(2)}\right| \neq 0 = \mathcal{E}(u)\end{aligned}$$

Hence, the constant sequence is not a recovery sequence, unlike in the case where the codimensions are larger than two. To fix this issue, we need to interpolate the value of u on $\mathcal{M}^{(1)}$ to the higher dimensional manifold $\mathcal{M}^{(2)}$ in such a way that the energy converges.

The idea is that since the harmonic capacity of $\mathcal{M}^{(12)}$ in $\mathcal{M}^{(2)}$ is zero one can use an interpolation based on the fundamental solution of Laplacian in 2D (the codimension) to interpolate the values. Here we sketch this approach in the simplified setting outlined above. In particular, we assume $d^{(1)} = 1$, $d^{(2)} = 2$, $d^{(12)} = 0$, and that $\mathcal{M}^{(12)}$ is a single point, w.l.o.g. $\mathcal{M}^{(12)} = \{0\}$. To make computations simple we also assume that $\mathcal{M}^{(2)} := \mathbb{T}^2$ is a two-dimensional flat torus, pierced orthogonally by a one-dimensional manifold $\mathcal{M}^{(1)}$ which is flat in a neighborhood of their intersection. Let $u \in H^1_{\sqrt{\mu}}(\mathcal{M})$ be such that $u^{(i)} := u|_{\mathcal{M}^{(i)}}$ are Lipschitz continuous. We define a recovery sequence $u_n \in L^2(\mathcal{M})$ via

$$u_n(x) := u^{(1)}(x)\sqrt{\frac{\varepsilon_n^{-1}\eta_{\varepsilon_n}(|x-\cdot|)\star\mu}{\alpha^{(1)}\beta_\eta^{(1)}\rho^{(1)}(x)}}, \quad x \in \mathcal{M}^{(1)},$$

on the one-dimensional manifold $\mathcal{M}^{(1)}$ and

$$u_n(x) := \begin{cases} u^{(1)}(0)\sqrt{\dfrac{\varepsilon_n^{-1}\eta_{\varepsilon_n}(|x-\cdot|)\star\mu}{\alpha^{(1)}\beta_\eta^{(1)}\rho^{(1)}(0)}}, & \text{if } |x| < \varepsilon_n,\\ u^{(2)}(x)\sqrt{\dfrac{\varepsilon_n^{-2}\eta_{\varepsilon_n}(|x-\cdot|)\star\mu}{\alpha^{(2)}\beta_\eta^{(2)}\rho^{(2)}(x)}} + \dfrac{\log\sqrt{\varepsilon_n} - \log|x|}{\log\sqrt{\varepsilon_n} - \log\varepsilon_n}\times & \\ \times\left(u^{(1)}(0)\sqrt{\dfrac{\varepsilon_n^{-1}\eta_{\varepsilon_n}(|x-\cdot|)\star\mu}{\alpha^{(1)}\beta_\eta^{(1)}\rho^{(1)}(0)}} - u^{(2)}(x)\sqrt{\dfrac{\varepsilon_n^{-2}\eta_{\varepsilon_n}(|x-\cdot|)\star\mu}{\alpha^{(2)}\beta_\eta^{(2)}\rho^{(2)}(x)}}\right), & \text{if } \varepsilon_n \leq |x| < \sqrt{\varepsilon_n},\\ u^{(2)}(x)\sqrt{\dfrac{\varepsilon_n^{-2}\eta_{\varepsilon_n}(|x-\cdot|)\star\mu}{\alpha^{(2)}\beta_\eta^{(2)}\rho^{(2)}(x)}}, & \text{if } \sqrt{\varepsilon_n} \leq |x|, \end{cases}$$

for $x \in \mathcal{M}^{(2)}$ on the two-dimensional one. Here we assume that $n \in \mathbb{N}$ is sufficiently large so that $\varepsilon_n < 1$. Note that the square root terms are all very close to one or even equal to one (at least all but the first one)

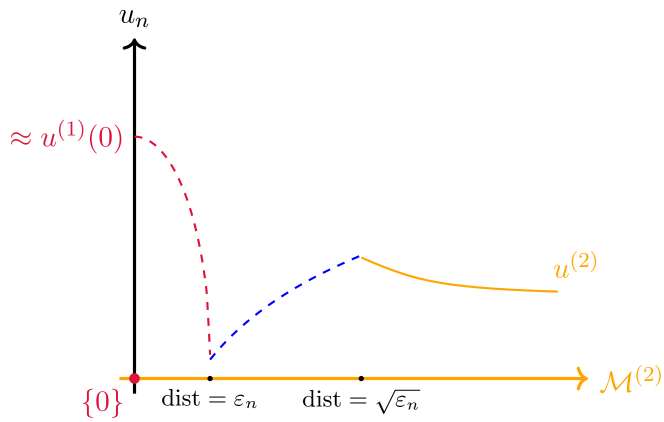


Figure 4: The recovery sequence approximates the value $u^{(1)}(0)$ on the intersection $\mathcal{M}^{(12)} = \{0\}$ and the function $u^{(2)}$ far from the intersection for $\varepsilon_n \ll 1$, interpolating logarithmically in between.

if the densities $\rho^{(i)}$ are constant on each of the two manifold. We plot this simplified situation in Figure 4 and it should be noted that the recovery sequence equals the value of $u^{(1)}$ in an ε_n-neighborhood of the intersection, logarithmically interpolates between the values of $u^{(1)}$ and $u^{(2)}$, in an annulus of diameter $\sqrt{\varepsilon_n}$, and equals $u^{(2)}$ outside of this annulus. By construction, the terms $\sqrt{\eta_{\varepsilon_n}(|x - \cdot|) \star \mu}$ will drop out once we plug u_n into the nonlocal energy $\mathcal{E}_{\varepsilon_n}$. To show that u_n is indeed a recovery sequence, we first need to argue that u_n converges to u in L^2. This is a simple consequence of the Lipschitz continuity of $u^{(i)}$ and of the behavior of the convolutions $\eta_{\varepsilon_n}(|x - \cdot|) \star \mu$ as analyzed in the compactness proof in Section 3.1. The interesting part is to prove that it satisfies the inequality $\limsup_{n\to\infty} \mathcal{E}_{\varepsilon_n}(u_n) \leq \mathcal{E}(u)$. Thanks to the regularity of u (and ρ) on each of the two manifolds, the only non-trivial region to study is the interpolation region $\varepsilon_n \leq |x| \leq \sqrt{\varepsilon_n}$. In the end, it suffices to argue that the interactions within this region and those between this interpolation region and the rest of $\mathcal{M}^{(2)}$ in the nonlocal energy go to zero. The main idea is that, since the harmonic capacity of a point in $\mathbb{R}^2$ is zero, one can use a logarithmic interpolation in the larger-dimensional manifold to match the values of the functions between the manifolds with little cost in terms of the (nonlocal) Dirichlet energy. The corresponding term can be controlled (after a change of variables using the Cheeger–Colding segment inequality in Theorem 4.7 below, and by employing the Lipschitz regularity of u and ρ) by the following integral

$$\begin{aligned}\int_{\varepsilon_n \leq |x| \leq \sqrt{\varepsilon_n}} \left| \nabla_x \frac{\log \sqrt{\varepsilon_n} - \log |x|}{\log \sqrt{\varepsilon_n} - \log \varepsilon_n} \right|^2 \mathrm{d}x &= \int_{\varepsilon_n \leq |x| \leq \sqrt{\varepsilon_n}} \left| \frac{\frac{x}{|x|^2}}{\log \sqrt{\varepsilon_n} - \log \varepsilon_n} \right|^2 \mathrm{d}x \\ &= \frac{4}{|\log \varepsilon_n|^2} \int_{\varepsilon_n \leq |x| \leq \sqrt{\varepsilon_n}} \frac{1}{|x|^2} \mathrm{d}x = \frac{8\pi}{|\log \varepsilon_n|^2} \int_{\varepsilon}^{\sqrt{\varepsilon_n}} \frac{1}{t} \mathrm{d}t = \frac{4\pi}{|\log \varepsilon_n|}\end{aligned} \tag{3.8}$$

which goes to zero as $n \to \infty$. Note that the same could not be achieved with other interpolations like, e.g., the linear one $\frac{\sqrt{\varepsilon_n} - |x|}{\sqrt{\varepsilon_n} - \varepsilon_n}$. Instead, we used the logarithmic growth of the fundamental solution of the Laplace equation in codimension two for this construction to work.

Recovery sequence for codimension one: $d^{(1)} = d^{(2)} =: d$ and $d - d^{(12)} = 1$. In this case, the energy $\mathcal{E}$ couples both manifolds through the trace condition encoded in $H^1_{\sqrt{\mu}}(\mathcal{M})$. Hence, aiming to apply a density argument we cannot simply regularize a function $u \in H^1_{\sqrt{\mu}}(\mathcal{M})$ on each of the manifolds separately since this violate the trace condition. Instead of approximating u directly, we approximate $v := u/\sqrt{\rho}$ which satisfies $\mathrm{Tr}^{(1)}(v) = \mathrm{Tr}^{(2)}(v)$. We approximate v by regular functions globally on $\mathcal{M}$ such that the trace condition is preserved and the energies of the approximations converge. By Theorem A.7 we can find a smooth approximation v_ε of v such that $v_\varepsilon|_{\mathcal{M}^{(i)}} \to v|_{\mathcal{M}^{(i)}}$ in $H^1(\mathcal{M}^{(i)})$ for $i \in \{1,2\}$ as $\varepsilon \to 0$ and $\mathrm{Tr}^{(1)}(v_\varepsilon) = \mathrm{Tr}^{(2)}(v_\varepsilon)$. We set $u_\varepsilon := v_\varepsilon\sqrt{\rho}$. Thus $\frac{u_\varepsilon}{\sqrt{\alpha^{(i)}\rho^{(i)}}} = v_\varepsilon$ on $\mathcal{M}^{(i)}$ and

$$\mathrm{Tr}^{(1)}\left(\frac{u_\varepsilon}{\sqrt{\alpha^{(1)}\rho^{(1)}}}\right) = \mathrm{Tr}^{(1)}(v_\varepsilon) = \mathrm{Tr}^{(2)}(v_\varepsilon) = \mathrm{Tr}^{(2)}\left(\frac{u_\varepsilon}{\sqrt{\alpha^{(2)}\rho^{(2)}}}\right),$$

which means that u_ε satisfies the trace condition. The Lipschitz continuity and the lower and upper bounds of $\rho^{(i)}$ then imply that $u_\varepsilon|_{\mathcal{M}^{(i)}} \in H^1(\mathcal{M}^{(i)})$ for $i \in \{1,2\}$ and hence we have showed that

$$u_\varepsilon \in H^1_{\sqrt{\mu}}(\mathcal{M}).$$

Furthermore $v_\varepsilon \to v = u/\sqrt{\rho}$ in $L^2(\mathcal{M})$ implies that $u_\varepsilon \to u$ in $L^2(\mathcal{M})$. As before we note that

$$\mathcal{E}(u) = \left(\alpha^{(1)}\right)^2 \mathcal{G}\left(\frac{u}{\sqrt{\alpha^{(1)}\beta_\eta^{(1)}\rho^{(1)}}};\mathcal{M}^{(1)}\right) + \left(\alpha^{(1)}\right)^2 \mathcal{G}^{(2)}\left(\frac{u}{\sqrt{\alpha^{(2)}\beta_\eta^{(2)}\rho^{(2)}}};\mathcal{M}^{(1)}\right). \tag{3.9}$$

Taking into account (3.9), the homogeneity of $\mathcal{G}(\cdot;\mathcal{M}^{(i)})$ as well as $\mathcal{G}(v_\varepsilon;\mathcal{M}^{(i)}) \to \mathcal{G}(v;\mathcal{M}^{(i)})$ as $\varepsilon \to 0$ by the H^1-convergence of v_ε, we deduce that $\mathcal{E}(u_\varepsilon) \to \mathcal{E}(u)$.

By the density above (combined with a diagonal argument) it suffices to construct a recovery sequence for u such that $v = u/\sqrt{\rho}$ is Lipschitz continuous. To be able to control cross terms near the intersection, instead of the constant recovery sequence, we consider the sequence

$$u_n(x) := u(x)\sqrt{\frac{\varepsilon_n^{-d}\eta_{\varepsilon_n}(|x-\cdot|) \star \mu}{\alpha^{(i)}\beta_\eta\rho^{(i)}(x)}}, \qquad x \in \mathcal{M}^{(i)},$$

where we dropped the index i on β_η since $d^{(1)} = d^{(2)} = d$. The rationale behind this choice is that the square-root term converges to one uniformly (yielding $u_n \to u$ in $L^2(\mathcal{M})$, cf. the proof of Theorem 3.1) and the numerator cancels once we plug u_n into the energy $\mathcal{E}_{\varepsilon_n}$. Using the fact that $v := \frac{u}{\sqrt{\alpha^{(i)}\beta_\eta\rho^{(i)}}}$ on $\mathcal{M}^{(i)}$, we have

$$\begin{aligned}\mathcal{E}_{\varepsilon_n}(u_n) = &\sum_{i=1}^2 \left(\alpha^{(i)}\right)^2 \mathcal{G}_{\varepsilon_n}\left(v;\mathcal{M}^{(i)}\right)\\ &+ \frac{2}{\varepsilon_n^{2+d}}\int_{\mathcal{M}^{(2)}}\int_{\mathcal{M}^{(1)}} \eta_{\varepsilon_n}(|x-y|)\,|v(x)-v(y)|^2\,\rho^{(1)}(x)\,\mathrm{dVol}^{(1)}(x)\rho^{(2)}(y)\,\mathrm{dVol}^{(2)}(y).\end{aligned}$$

Note that v is globally Lipschitz on $\mathcal{M}$. In particular, by Theorem A.3 (in the proof of which a constant recovery is used for Lipschitz functions) we have

$$\limsup_{n\to\infty}\sum_{i=1}^2 \left(\alpha^{(i)}\right)^2 \mathcal{G}_{\varepsilon_n}\left(v;\mathcal{M}^{(i)}\right) = \sum_{i=1}^2 \left(\alpha^{(i)}\right)^2 \mathcal{G}\left(v;\mathcal{M}^{(i)}\right) = \mathcal{E}(u).$$

It remains to show that the cross term in the above energy decomposition of $\mathcal{E}_{\varepsilon_n}(u_n)$ tends to zero for which we shall use Lipschitzness of v:

$$\begin{aligned}&\frac{2}{\varepsilon_n^{2+d}}\int_{\mathcal{M}^{(2)}}\int_{\mathcal{M}^{(1)}} \eta_{\varepsilon_n}(|x-y|)\,|v(x)-v(y)|^2\,\rho^{(1)}(x)\,\mathrm{dVol}^{(1)}(x)\rho^{(2)}(y)\,\mathrm{dVol}^{(2)}(y)\\ \lesssim\ &\frac{1}{\varepsilon_n^d}\int_{\mathcal{M}^{(2)}}\int_{\mathcal{M}^{(1)}} \eta_{\varepsilon_n}(|x-y|)\,\mathrm{dVol}^{(1)}(x)\,\mathrm{dVol}^{(2)}(y) \lesssim \frac{1}{\varepsilon_n^d}\varepsilon_n^{d-d^{(12)}}\varepsilon_n^d = \varepsilon_n \to 0, \qquad n\to\infty.\end{aligned}$$

$\square$

The proof of the limsup inequality for Theorem 2.2 can be found in Section 4.3 and follows the ideas and case distinctions outlined above. In particular, we refer to Theorem 4.8 therein for the above mentioned consequence of the Cheeger–Colding segment inequality that allows one to upper bound the nonlocal by the local energy.

4 Proofs of Γ-convergence and compactness for graph Dirichlet energies

In this section we prove our main results Theorem 2.1 and Theorem 2.2. We first establish the following uniform convergence for the degrees, which are nothing but kernel density estimates. Let us consider the contribution to the degree from vertices in each manifold:

$$\deg^{(i)}_{n,\varepsilon}(x) := \frac{1}{n} \sum_{y \in V_n \cap \mathcal{M}^{(i)}} \eta_\varepsilon(|x-y|).$$

Note that $\deg_{n,\varepsilon}(x) = \deg^{(1)}_{n,\varepsilon}(x) + \deg^{(2)}_{n,\varepsilon}(x)$.

Lemma 4.1 (Uniform convergence of kernel density estimators)**.** *Under Assumptions 1 to 3 it holds with probability one that*

$$\lim_{n\to\infty} \max_{i\in\{1,2\}} \sup_{x\in\mathcal{M}^{(i)}} \left| \varepsilon_n^{-d^{(i)}} \deg^{(i)}_{n,\varepsilon_n}(x) - \alpha^{(i)}\beta^{(i)}_\eta \rho^{(i)}(x) \right| = 0.$$

Proof. The result follows from [56, Theorem 2.3 (2)], combined with the Borel–Cantelli lemma. $\square$

The next result allows us to take into account the interactions between manifolds. We use it in the proof of compactness. We show that for $x \in \mathcal{M}$ the degree $\deg_{n,\varepsilon_n}(x)$ with high probability scales like $\mathrm{Vol}(B(x,\varepsilon_n)\cap\mathcal{M})$, i.e., the volume of the intersection of a Euclidean ball with radius ε_n with $\mathcal{M}$.

Lemma 4.2. *Under Assumptions 1 and 2 there exist constants $C_1, C_2, C_3, C_4 > 0$ and $\tau \in (0,1]$ only depending on ρ, η, and $\mathcal{M}$ such that for $n \in \mathbb{N}$ sufficiently large it holds*

$$\begin{aligned}
&\mathbb{P}\Big((\forall x \in V_n)\, C_1 \mathrm{Vol}(B(x,\tau\varepsilon_n)\cap\mathcal{M}) \le \deg_{n,\varepsilon_n}(x) \le C_2 \mathrm{Vol}(B(x,\varepsilon_n)\cap\mathcal{M}) \Big) \\
&\quad \ge 1 - C_3 \exp\left(-C_4 n \varepsilon_n^{d^{(2)}} + \log n \right).
\end{aligned}$$

Proof. For fixed $x \in \mathcal{M}$, applying Bernstein's inequality (see [13, Lemma 1] for a convenient version) to the random variable $\deg_{n,\varepsilon_n}(x)$ we obtain that for all $\lambda \in [0,1]$ it holds

$$\begin{aligned}
&\mathbb{P}\left(\left|\deg_{n,\varepsilon_n}(x) - \mathbb{E}\left[\deg_{n,\varepsilon_n}(x)\right]\right| > C_\rho \mathrm{Vol}(B(x,\varepsilon_n)\cap\mathcal{M})\lambda \right) \\
&\qquad \le 2\exp\left(-\frac{1}{4} C_\rho n \mathrm{Vol}(B(x,\varepsilon_n)\cap\mathcal{M})\lambda^2 \right) \le 2\exp\left(-C n \varepsilon_n^{d^{(2)}} \lambda^2 \right),
\end{aligned} \tag{4.1}$$

where $C > 0$ depends on C_ρ and $\mathcal{M}$. We also used that for $n \in \mathbb{N}$ sufficiently large the volume $\mathrm{Vol}(B(x,\varepsilon_n))$ is smallest for points $x \in \mathcal{M}^{(2)}$, in the larger dimensional manifold.

Furthermore, we can use that the vertices $y \in V_n$ are i.i.d. samples from μ to compute

$$\mathbb{E}\left[\deg_{n,\varepsilon_n}(x)\right] = \mathbb{E}\left[\frac{1}{n} \sum_{y\in V_n} \eta_{\varepsilon_n}(|x-y|) \right] = \int_{\mathcal{M}} \eta_{\varepsilon_n}(|x-y|)\mathrm{d}\mu(y).$$

Hence, thanks to Assumption 2 (arguing as in Section 3.1 there exist constants $C_1, C_2, \tau > 0$ (depending only on η) such that

$$C_1 \mathrm{Vol}(B(x,\tau\varepsilon_n)\cap\mathcal{M}) \le \mathbb{E}\left[\deg_{n,\varepsilon_n}(x)\right] \le C_2 \mathrm{Vol}(B(x,\varepsilon_n)\cap\mathcal{M}).$$

Using this estimate, we can find a suitable $\lambda \in (0,1]$ in (4.1) and constants $C_3, C_4 > 0$ such that

$$\mathbb{P}\Big(C_1\mathrm{Vol}(B(x,\tau\varepsilon_n)\cap\mathcal{M}) \le \deg_{n,\varepsilon_n}(x) \le C_2\mathrm{Vol}(B(x,\varepsilon_n)\cap\mathcal{M})\Big) \ge 1 - C_3\exp\left(-C_4 n\varepsilon_n^{d^{(2)}}\right).$$

By conditioning on $x \in V_n$, using a union bound and redefining C_1, C_2, C_3, C_4 we obtain the desired statement. □

Remark 4.1. For the probability in Theorem 4.2 to be close to one we need Assumption 3 to be satisfied which guarantees that $n\varepsilon_n^{d^{(2)}} \gg \log n$.

Remark 4.2. Under assumptions of Theorem 4.2, we note that for $x \in \mathcal{M}^{(i)}$

$$\begin{aligned}
\mathrm{Vol}(B(x,\varepsilon_n)\cap\mathcal{M}) &\le C\varepsilon_n^{d^{(i)}}, \quad \text{if } \mathrm{dist}(x,\mathcal{M}^{(3-i)}) > \varepsilon_n,\\
\mathrm{Vol}(B(x,\varepsilon_n)\cap\mathcal{M}) &\ge C\varepsilon_n^{d^{(1)}} \quad \text{if } \mathrm{dist}(x,\mathcal{M}^{(3-i)}) < \varepsilon_n/2,
\end{aligned}$$

for some $C > 0$, depending on $\mathcal{M}$. Thus it can be expected that the lower dimensional manifold dominates the scaling of the degrees $\deg_{n,\varepsilon_n}(x)$ of points x close to the intersection of the two manifolds.

From above lemmas and remarks we have the following corollary.

Corollary 4.3 (Uniform convergence of kernel density estimators II). *Under Assumptions 1 to 3, if $d^{(1)} < d^{(2)}$ it holds with probability one that*

$$\lim_{n\to\infty}\sup_{x\in\mathcal{M}^{(1)}}\left|\varepsilon_n^{-d^{(1)}}\deg_{n,\varepsilon_n}(x) - \alpha^{(1)}\beta_\eta^{(1)}\rho^{(1)}(x)\right| = 0.$$

This follows from Theorem A.7 since $\deg_{n,\varepsilon_n} = \deg^{(1)}_{n,\varepsilon_n} + \deg^{(2)}_{n,\varepsilon_n}$ and, by Theorem 4.2 and Remark 4.2, $\deg^{(2)}_{n,\varepsilon_n}(x) \le C\varepsilon_n^{d^{(2)}} \ll \varepsilon_n^{d^{(1)}}$.

Remark 4.3. Under assumptions Assumptions 1 to 3 , we almost surely have that

$$\limsup_{n\to\infty}\max_{i\in\{1,2\}}\sup_{x\in\mathcal{M}^{(i)}}\frac{1}{\varepsilon_n^{-d^{(i)}}\deg_{n,\varepsilon_n}(T_n(x))} < \infty.$$

This follows from Theorem 4.2 and Remark 4.2. In particular, if $x \in \mathcal{M}^{(1)}$ and $T_n(x) \in \mathcal{M}^{(2)}$ then since $|T_n(x) - x| < \varepsilon_n/2$ we have that $d(T_n(x),\mathcal{M}^{(1)}) < \varepsilon_n/2$ so Remark 4.2 applies.

Using the previous results we can prove the following lemma, which is needed in proving Γ-convergence and compactness.

Lemma 4.4. *Under assumptions Assumptions 1 to 3 consider $v \in L^2(\mathcal{M})$ and let $T_n : \mathcal{M} \to \mathcal{M}$ be a map which almost surely satisfies $(T_n)_\sharp\mu = \mu_n$ and* (2.12). *Then it holds with probability one that*

$$\lim_{n\to\infty}\max_{i\in\{1,2\}}\int_{\mathcal{M}^{(i)}}\left|\sqrt{\frac{\alpha^{(i)}\beta_\eta^{(i)}\rho^{(i)}(x)}{\varepsilon_n^{-d^{(i)}}\deg_{n,\varepsilon_n}(T_n(x))}} - 1\right|^2 |v(x)|^2 \,\mathrm{dVol}^{(i)}(x) = 0.$$

Proof. Fix $\delta > 0$ and let $x \in \mathcal{M}^{(i)}$ with $\operatorname{dist}(x, \mathcal{M}^{(12)}) > \delta$. We have

$$\frac{\alpha^{(i)}\beta_\eta^{(i)}\rho^{(i)}(x)}{\varepsilon_n^{-d^{(i)}}\deg_{n,\varepsilon_n}(T_n(x))} = \frac{\alpha^{(i)}\beta_\eta^{(i)}\rho^{(i)}(T_n(x))}{\varepsilon_n^{-d^{(i)}}\deg_{n,\varepsilon_n}(T_n(x))} + \frac{\alpha^{(i)}\beta_\eta^{(i)}\left(\rho^{(i)}(x) - \rho^{(i)}(T_n(x))\right)}{\varepsilon_n^{-d^{(i)}}\deg_{n,\varepsilon_n}(T_n(x))}.$$

We note that the presence of the other manifold only increases the observed degree. Therefore, combining Theorem 4.2 and Assumption 3 with the Borel–Cantelli lemma and Remark 4.2 we almost surely have

$$\liminf_{n\to\infty} \inf_{x\in\mathcal{M}} \varepsilon_n^{-d^{(i)}} \deg_{n,\varepsilon_n}(T_n(x)) > 0,$$

as established in Remark 4.3, where i corresponds to the manifold that x belongs to.

Using also Assumption 1, Equation (2.12), and Theorems 4.1 and 4.3 we get for any $\delta > 0$ and for all $x \in \mathcal{M}$ with $d(x, \mathcal{M}^{(12)}) > \delta$

$$\lim_{n\to\infty} \left| \sqrt{\frac{\alpha^{(i)}\beta_\eta^{(i)}\rho^{(i)}(x)}{\varepsilon_n^{-d^{(i)}}\deg_{n,\varepsilon_n}(T_n(x))}} - 1 \right|^2 \leq C \lim_{n\to\infty} \|T_n - \mathrm{id}\|_{L^\infty(\mathcal{M})} = 0,$$

for some $C > 0$. Hence, the dominated convergence theorem implies that with probability one

$$\lim_{n\to\infty} \int_{\{x\in\mathcal{M}^{(i)}\,:\,\operatorname{dist}(x,\mathcal{M}^{(i)})>\delta\}} \left| \sqrt{\frac{\alpha^{(i)}\beta_\eta^{(i)}\rho^{(i)}(x)}{\varepsilon_n^{-d^{(i)}}\deg_{n,\varepsilon_n}(T_n(x))}} - 1 \right|^2 |v(x)|^2 \,\mathrm{dVol}^{(i)}(x) = 0.$$

On the other hand, using the lower bound on the degrees from above, we almost surely have that

$$\begin{aligned} \limsup_{n\to\infty} \int_{\{x\in\mathcal{M}^{(i)}\,:\,\operatorname{dist}(x,\mathcal{M}^{(i)})<\delta\}} &\left| \sqrt{\frac{\alpha^{(i)}\beta_\eta^{(i)}\rho^{(i)}(x)}{\varepsilon_n^{-d^{(i)}}\deg_{n,\varepsilon_n}(T_n(x))}} - 1 \right|^2 |v(x)|^2 \,\mathrm{dVol}^{(i)}(x) \\ &\leq C \int_{\{x\in\mathcal{M}^{(i)}\,:\,\operatorname{dist}(x,\mathcal{M}^{(i)})<\delta\}} |v(x)|^2 \,\mathrm{dVol}^{(i)}(x). \end{aligned}$$

Combining the two estimates we almost surely have

$$\lim_{n\to\infty} \int_{\mathcal{M}^{(i)}} \left| \sqrt{\frac{\alpha^{(i)}\beta_\eta^{(i)}\rho^{(i)}(x)}{\varepsilon_n^{-d^{(i)}}\deg_{n,\varepsilon_n}(T_n(x))}} - 1 \right|^2 |v(x)|^2 \,\mathrm{dVol}^{(i)}(x) = 0.$$

□

4.1 Compactness

In this section we provide a proof of Theorem 2.1. We start by a simple lemma that has broader applications.

Lemma 4.5. *Let $\mathcal{M}$ and μ be as above.*[1] *Let μ_n be a sequence of probability measures on $\mathcal{M}$ converging in Wasserstein distance to μ. Let $v_n \in L^2(\mu_n)$ and $v \in L^2(\mu)$ be such that v_n converge in TL^2 distance to v. Let A_n be a sequence of sets of vanishing mass, $\mu_n(A_n) \to 0$ as $n \to \infty$. Then $v_n\chi_{A_n}$ converges to zero in the TL^2 sense.*

[1]This can be extended to a general setting of metric measure spaces.

Proof. Let $T_{n,\sharp}\mu = \mu_n$ be a stagnating sequence of transport maps. Let $H_n = T_n^{-1}(A_n)$. Note that $\mu(H_n) = \mu_n(A_n) \to 0$ as $n \to \infty$. Thus $\|v\chi_{H_n}\|_{L^2(\mu)} \to 0$ as $n \to \infty$. By properties of TL^2 convergence, [32, Lemma 3.11], we have $\int |v(x) - v_n(T_n(x)|^2 \mathrm{d}\mu(x) \to 0$ as $n \to \infty$ and thus

$$\int_{H_n} |v_n(T_n(x)|^2 \mathrm{d}\mu(x) \leq 2\int_{H_n} |v(x) - v_n(T_n(x)|^2 + |v(x)|^2 \mathrm{d}\mu(x) \to 0 \qquad \text{as } n \to \infty.$$

Hence $\int_{A_n} |v_n(y)|^2 \mathrm{d}\mu_n(y) \to 0$ as $n \to \infty$. $\square$

Proof of Theorem 2.1. Since by Assumption 2 $\eta \geq c_1\chi_{[0,c_2]}$ for some $c_1, c_2 > 0$, it suffices to show the statement for $\eta = \chi_{[0,1]}$. Let us define the functions $v_n(x) := \frac{u_n(x)}{\sqrt{\varepsilon_n^{-d^{(i)}} \deg_{n,\varepsilon_n}(x)}}$ for $x \in \mathcal{M}^{(i)} \cap V_n$. Using that the $L^2(\mu_n)$-norms of u_n are uniformly bounded and Remark 4.3, we have that almost surely the $L^2(\mu_n)$-norms of v_n are uniformly bounded as $n \to \infty$. Using $\mathcal{M} = \mathcal{M}^{(1)} \cup \mathcal{M}^{(2)}$ in (2.7) and omitting cross-terms we have

$$\begin{aligned} E_{n,\varepsilon_n}(u_n) &\geq \frac{1}{\varepsilon_n^2}\sum_{i=1}^2 \left(\frac{n^{(i)}}{n}\right)^2 \int_{\mathcal{M}^{(i)}}\int_{\mathcal{M}^{(i)}} \eta_{\varepsilon_n}(|x-y|)\left|\frac{u_n(x)}{\sqrt{\deg_{n,\varepsilon_n}(x)}} - \frac{u_n(y)}{\sqrt{\deg_{n,\varepsilon_n}(y)}}\right|^2 \mathrm{d}\mu_n^{(i)}(x)\mathrm{d}\mu_n^{(i)}(y). \\ &= \sum_{i=1}^2 \left(\frac{n^{(i)}}{n}\right)^2 \frac{1}{\varepsilon_n^{d^{(i)}+2}} \int_{\mathcal{M}^{(i)}}\int_{\mathcal{M}^{(i)}} \eta_{\varepsilon_n}(|x-y|)\left|v_n^{(i)}(x) - v_n^{(i)}(y)\right|^2 \mathrm{d}\mu_n^{(i)}(x)\mathrm{d}\mu_n^{(i)}(y). \end{aligned}$$

Theorem A.5 together with the fact that almost surely we have $\lim_{n\to\infty} \frac{n^{(i)}}{n} = \alpha^{(i)} > 0$ implies that the sequence v_n has a subsequence (which we do not relabel) which converges almost surely as $n \to \infty$ in the $TL^2(\mathcal{M})$-sense to some $v \in L^2(\mathcal{M})$.

We now argue that this implies the TL^2 convergence of u_n towards

$$u(x) := \sqrt{\alpha^{(i)}\beta_\eta^{(i)}\rho^{(i)}(x)}\, v(x), \qquad x \in \mathcal{M}^{(i)}.$$

Let us choose transport maps T_n which satisfy (2.12). Using the fact that $v_n \to v$ in $TL^2(\mathcal{M})$ as well as Theorem 4.4 we obtain almost surely

$$\begin{aligned} &\lim_{n\to\infty}\sum_{i=1}^2 \int_{\mathcal{M}^{(i)}} |u_n(T_n(x)) - u(x)|^2 \frac{\mathrm{dVol}^{(i)}(x)}{\varepsilon_n^{-d^{(i)}}\deg_{n,\varepsilon_n}(T_n(x))} \\ &\qquad = \lim_{n\to\infty}\sum_{i=1}^2 \int_{\mathcal{M}^{(i)}} \left|v_n(T_n(x)) - \sqrt{\frac{\alpha^{(i)}\beta_\eta^{(i)}\rho^{(i)}(x)}{\varepsilon_n^{-d^{(i)}}\deg_{n,\varepsilon_n}(T_n(x))}}v(x)\right|^2 \mathrm{dVol}^{(i)}(x) \\ &\qquad \leq 2\lim_{n\to\infty}\sum_{i=1}^2 \int_{\mathcal{M}^{(i)}} |v_n(T_n(x)) - v(x)|^2\, \mathrm{dVol}^{(i)}(x) \\ &\qquad\quad + 2\lim_{n\to\infty}\sum_{i=1}^2 \int_{\mathcal{M}^{(i)}} \left|1 - \sqrt{\frac{\alpha^{(i)}\beta_\eta^{(i)}\rho^{(i)}(x)}{\varepsilon_n^{-d^{(i)}}\deg_{n,\varepsilon_n}(T_n(x))}}\right|^2 |v(x)|^2\, \mathrm{dVol}^{(i)}(x) = 0 \end{aligned} \tag{4.2}$$

Let, as before, $A_\varepsilon = \mathcal{M}^{(1)} \cup \{x \in \mathcal{M}^{(2)} : d(x, \mathcal{M}^{(1)}) > \varepsilon\}$. By Theorems 4.1 and 4.3, with probability one,

$$\lim_{n\to\infty} \sup_{x\in A_{\varepsilon_n}\cap\mathcal{M}^{(i)}} \left|\varepsilon_n^{-d^{(i)}}\deg_{n,\varepsilon_n}(x) - \alpha^{(i)}\beta_\eta^{(i)}\rho^{(i)}(x)\right| = 0,$$

and hence $\varepsilon_n^{-d^{(i)}} \deg_{n,\varepsilon_n}(x)$ is bounded from below uniformly in x on A_{ε_n}. Thus (4.2) implies that

$$\lim_{n\to\infty} \int_{T_n^{-1}(A_{\varepsilon_n})} |u_n(T_n(x)) - u(x)|^2 \mathrm{d}\mu(x) \to 0 \qquad \text{as } n \to \infty.$$

Consequently, it remains to prove that $\|u_n\|_{L^2(\mu_n \llcorner A_{\varepsilon_n}^c)} \to 0$ as $n \to 0$ where $A_{\varepsilon_n}^c := \mathcal{M} \setminus A_{\varepsilon_n} = \{x \in \mathcal{M}^{(2)} \,:\, d(x, \mathcal{M}^{(1)}) < \varepsilon_n\}$. We argue as in the proof of Theorem 3.1. Let S_ε^k and G_ε^k be as in that proof.

We claim that there exists $C_k > 0$ such that for all n large enough for all $x \in S_{\varepsilon_n}^k$

$$\eta_{\varepsilon_n}(|x-\cdot|) \star (\mu_n \llcorner G_{\varepsilon_n}^k) > C_k \eta_{\varepsilon_n}(|x-\cdot|) \star \mu_n = C_k \deg_{n,\varepsilon_n}(x). \tag{4.3}$$

For $k=1$ this follows from Theorem 4.2, since for $x \in S_{\varepsilon_n}^1,\ B(x,\varepsilon_n) \cap \mathcal{M}^{(1)} = B(x,\varepsilon_n) \cap G_{\varepsilon_n}^1$. For $k > 1$ note that

$$\eta_{\varepsilon_n}(|x-\cdot|) \star (\mu_n \llcorner G_{\varepsilon_n}^k \cap \mathcal{M}^{(1)}) = \eta_{\varepsilon_n}(|x-\cdot|) \star (\mu_n \llcorner \mathcal{M}^{(1)})$$

while the fact that

$$\eta_{\varepsilon_n}(|x-\cdot|) \star (\mu_n \llcorner G_{\varepsilon_n}^k \cap \mathcal{M}^{(2)}) \geq C_k \eta_{\varepsilon_n}(|x-\cdot|) \star (\mu_n \llcorner \mathcal{M}^{(2)})$$

for some $C_k > 0$ follows from a concentration estimate analogous to ones in Theorem 4.2, since $\mathrm{Vol} \llcorner \mathcal{M}^{(2)}(B(x,\varepsilon_n) \cap G_{\varepsilon_n}^k) > C \mathrm{Vol} \llcorner \mathcal{M}^{(2)}(B(x,\varepsilon_n))$ for some $C > 0$. Combining the two facts implies (4.3). Therefore,

$$\begin{aligned}
&C_k \|u_n\|^2_{L^2(\mu_n \llcorner S_{\varepsilon_n}^k)} - \|u_n\|^2_{L^2(\mu_n \llcorner G_{\varepsilon_n}^k)} \\
&\leq \int_{S_{\varepsilon_n}^k} \frac{\eta_{\varepsilon_n}(|x-\cdot|) \star (\mu_n \llcorner G_{\varepsilon_n}^k)}{\eta_{\varepsilon_n}(|x-\cdot|) \star \mu_n} u_n^2(x) \mathrm{d}\mu_n(x) - \int_{G_{\varepsilon_n}^k} \frac{\eta_{\varepsilon_n}(|y-\cdot|) \star (\mu_n \llcorner S_{\varepsilon_n}^k)}{\eta_{\varepsilon_n}(|y-\cdot|) \star \mu_n} u_n^2(y) \mathrm{d}\mu(y) \\
&= \int_{S_{\varepsilon_n}^k} \int_{G_{\varepsilon_n}^k} \eta_{\varepsilon_n}(|x-y|) \left(\frac{u_n^2(x)}{\eta_{\varepsilon_n}(|x-\cdot|) \star \mu_n} - \frac{u_n^2(y)}{\eta_{\varepsilon_n}(|y-\cdot|) \star \mu_n} \right) \mathrm{d}\mu_n(x) \mathrm{d}\mu_n(y) \\
&\leq \sqrt{\int_{S_{\varepsilon_n}^k} \int_{G_{\varepsilon_n}^k} \eta_{\varepsilon_n}(|x-y|) \left(\frac{u_n(x)}{\sqrt{\eta_{\varepsilon_n}(|x-\cdot|) \star \mu_n}} - \frac{u_n(y)}{\sqrt{\eta_{\varepsilon_n}(|y-\cdot|) \star \mu_n}} \right)^2 \mathrm{d}\mu_n(x) \mathrm{d}\mu_n(y)} \\
&\quad \times \sqrt{\int_{S_{\varepsilon_n}^k} \int_{G_{\varepsilon_n}^k} \eta_{\varepsilon_n}(|x-y|) \left(\frac{u_n(x)}{\sqrt{\eta_{\varepsilon_n}(|x-\cdot|) \star \mu_n}} + \frac{u_n(y)}{\sqrt{\eta_{\varepsilon_n}(|y-\cdot|) \star \mu}} \right)^2 \mathrm{d}\mu_n(x) \mathrm{d}\mu_n(y)} \\
&\leq 2\varepsilon_n \sqrt{E_{n,\varepsilon_n}(u_n)} \|u_n\|_{L^2(\mu_n)},
\end{aligned}$$

where we also used that $\eta_{\varepsilon_n}(|x-\cdot|) \star \mu_n = \deg_{n,\varepsilon_n}(x)$ for $x \in \mathcal{M}$. Hence

$$\|u_n\|^2_{L^2(\mu_n \llcorner S_{\varepsilon_n}^k)} \leq \frac{1}{C_k} \left(\|u_n\|^2_{L^2(\mu_n \llcorner G_{\varepsilon_n}^k)} + 2\varepsilon_n \sqrt{E_{n,\varepsilon_n}(u_n)} \|u_n\|_{L^2(\mu_n)} \right)$$

We claim that the right-hand side converges to zero as $n \to \infty$. Note that for $k = 1$, $\|u_n\|^2_{L^2(\mu_n \llcorner G_{\varepsilon_n}^1)}$ converges to zero by Theorem 4.5 since $u_n \chi_{A_{\varepsilon_n}}$ converges in TL^2, $G_{\varepsilon_n}^1 \subset A_{\varepsilon_n}$, and $\mu_n(G_{\varepsilon_n}^1) \to 0$ as $n \to \infty$, so the right-hand side converges to zero. For $k > 1$ the conclusion follows by induction. $\square$

4.2 Γ-liminf inequality

In this section we prove the $\liminf$ inequality of Theorem 2.2, using Theorem A.6.

Proposition 4.6 (Liminf inequality)**.** *Under Assumptions 1 to 3 if* $u_n \xrightarrow{TL^2} u$*, it almost surely holds that*

$$\mathcal{E}(u) \leq \liminf_{n\to\infty} E_{n,\varepsilon_n}(u_n).$$

Proof. Without loss of generality we can assume that

$$\liminf_{n\to\infty} E_{n,\varepsilon}(u_n) < \infty \tag{4.4}$$

since otherwise there is nothing to prove. Dropping cross-terms we get

$$E_{n,\varepsilon_n}(u_n) \geq \sum_{i=1}^{2} \left(\frac{n^{(i)}}{n}\right)^2 \frac{1}{\varepsilon_n^{2+d^{(i)}}} \int_{\mathcal{M}^{(i)}} \int_{\mathcal{M}^{(i)}} \eta_{\varepsilon_n}(|x-y|)\, |v_n(x) - v_n(y)|^2 \,\mathrm{d}\mu_n^{(i)}(x)\mathrm{d}\mu_n^{(i)}(y),$$

where we define

$$v_n(x) := \frac{u_n(x)}{\sqrt{\varepsilon_n^{-d^{(i)}} \deg_{n,\varepsilon_n}(x)}}, \quad x \in \mathcal{M}^{(i)} \cap V_n.$$

Let us first argue that $v_n \to v$ in the $TL^2(\mathcal{M})$-sense where $v := \frac{u}{\sqrt{\alpha^{(i)}\beta_\eta^{(i)}\rho^{(i)}}}$ on $\mathcal{M}^{(i)}$. For this we let $T_n : \mathcal{M} \to \mathcal{M}$ be transport maps satisfying $(T_n)_\sharp \mu = \mu_n$ as well as (2.12). Using $u_n \to u$ in the TL^2-sense, together with Assumption 1 and Theorem 4.4, we obtain

$$\begin{aligned}
&\lim_{n\to\infty} \sum_{i=1}^{2} \int_{\mathcal{M}^{(i)}} |v_n(T_n(x)) - v(x)|^2 \,\mathrm{d}\mu^{(i)}(x) \\
&\quad = \lim_{n\to\infty} \sum_{i=1}^{2} \int_{\mathcal{M}^{(i)}} \left| u_n(T_n(x)) - \sqrt{\frac{\varepsilon_n^{-d^{(i)}} \deg_{n,\varepsilon_n}(T_n(x))}{\alpha^{(i)}\beta_\eta^{(i)}\rho^{(i)}(x)}} u(x) \right|^2 \frac{\mathrm{d}\mu^{(i)}(x)}{\varepsilon_n^{-d^{(i)}} \deg_{n,\varepsilon_n}(T_n(x))} \\
&\quad \leq 2 \lim_{n\to\infty} \sum_{i=1}^{2} \int_{\mathcal{M}^{(i)}} |u_n(T_n(x)) - u(x)|^2 \frac{\mathrm{dVol}^{(i)}(x)}{\varepsilon_n^{-d^{(i)}} \deg_{n,\varepsilon_n}(T_n(x))} \\
&\qquad + 2 \lim_{n\to\infty} \sum_{i=1}^{2} \int_{\mathcal{M}^{(i)}} \left| \sqrt{\frac{\alpha^{(i)}\beta_\eta^{(i)}\rho^{(i)}(x)}{\varepsilon_n^{-d^{(i)}} \deg_{n,\varepsilon_n}(T_n(x))}} - 1 \right|^2 \frac{|u(x)|^2}{\alpha^{(i)}\beta_\eta^{(i)}\rho^{(i)}(x)} \,\mathrm{dVol}^{(i)}(x) = 0
\end{aligned}$$

and hence the $TL^2(\mathcal{M})$-convergence of v_n to v holds.

Applying the $\liminf$ inequality from Theorem A.5 to the manifolds $\mathcal{M}^{(i)}$ we obtain that

$$\liminf_{n\to\infty} E_{n,\varepsilon_n}(u_n) \geq \sum_{i=1}^{2} \frac{\alpha^{(i)}\sigma_\eta^{(i)}}{\beta_\eta^{(i)}} \int_{\mathcal{M}^{(i)}} \left| \nabla_{\mathcal{M}^{(i)}} \left(\frac{u}{\sqrt{\rho^{(i)}}} \right) \rho^{(i)} \right|^2 \mathrm{dVol}^{(i)} \tag{4.5}$$

holds almost surely. In particular, this inequality shows that $\frac{u}{\sqrt{\rho^{(i)}}} \in H^1(\mathcal{M}^{(i)})$ for $i = 1, 2$.

Case 1, $d^{(1)} \neq d^{(2)}$ or $d^{(1)} - d^{(12)} > 1$ In this case we are done since $u \in H^1_{\sqrt{\mu}}(\mathcal{M})$ (defined in Equation (2.10)) and hence (4.5) implies $\liminf_{n\to\infty} E_{n,\varepsilon_n}(u_n) \geq \mathcal{E}(u)$.

Case 2, $d^{(1)} = d^{(2)} =: d$ and $d^{(12)} = d - 1$ To show that $u \in H^1_{\sqrt{\mu}}(\mathcal{M})$ we have to show the trace condition of Equation (2.10). For this, we take a sequence of transport maps $T_n : \mathcal{M} \to \mathcal{M}$ satisfying $(T_n)_\sharp \mu = \mu_n$ and (2.12). Using their pushforward property it holds

$$E_{n,\varepsilon_n}(u_n) = \frac{1}{\varepsilon_n^{d+2}} \int_{\mathcal{M}} \int_{\mathcal{M}} \eta_{\varepsilon_n}(|T_n(x) - T_n(y)|) \, |v_n(T_n(x)) - v_n(T_n(y))|^2 \, \mathrm{d}\mu(x) \, \mathrm{d}\mu(y).$$

By Assumption 2 there exists $t_0 \in (0, 1]$ and $a_0 > 0$ such that $\eta(t) \geq a_0$ for all $0 \leq t \leq t_0$. Hence, as in [32, Section 5] we can assume without loss of generality that $\eta(t) = 1_{[0,1]}(t)$. Using the estimate

$$|T_n(x) - T_n(y)| \geq |x - y| - 2 \, \|T_n - \mathrm{id}\|_{L^\infty(\mathcal{M})}$$

and defining $\tilde{\varepsilon}_n := \varepsilon_n - 2 \, \|T_n - \mathrm{id}\|_{L^\infty(\mathcal{M})}$ we see that

$$E_{n,\varepsilon_n}(u_n) \geq \frac{\tilde{\varepsilon}_n^{d+2}}{\varepsilon_n^{d+2}} \frac{1}{\tilde{\varepsilon}_n^{d+2}} \int_{\mathcal{M}} \int_{\mathcal{M}} \eta_{\tilde{\varepsilon}_n}(|x - y|) \, |v_n(T_n(x)) - v_n(T_n(y))|^2 \, \mathrm{d}\mu(x) \, \mathrm{d}\mu(y).$$

By Assumption 3 and Remark 2.3 it holds $\lim_{n\to\infty} \frac{\tilde{\varepsilon}_n}{\varepsilon_n} = 1$ and $\lim_{n\to\infty} \tilde{\varepsilon}_n = 0$. Using also (4.4) we have

$$\liminf_{n\to\infty} \frac{1}{\tilde{\varepsilon}_n^{d+2}} \int_{\mathcal{M}} \int_{\mathcal{M}} \eta_{\tilde{\varepsilon}_n}(|x - y|) \, |v_n(T_n(x)) - v_n(T_n(y))|^2 \, \mathrm{d}\mu(x) \, \mathrm{d}\mu(y) < \infty. \tag{4.6}$$

Above, we have already proved that $v_n \circ T_n \to v$ in $L^2(\mathcal{M})$. Using this together with (4.6) and applying Theorem A.6 we obtain that $\mathrm{Tr}^{(1)}\left(u/\sqrt{\alpha^{(1)}\rho^{(1)}}\right) = \mathrm{Tr}^{(2)}\left(u/\sqrt{\alpha^{(2)}\rho^{(2)}}\right)$ on $\mathcal{M}^{(12)}$ which implies $u \in H^1_{\sqrt{\mu}}(\mathcal{M})$. Combining this with (4.5) we have that $\liminf_{n\to\infty} E_{n,\varepsilon_n}(u_n) \geq \mathcal{E}(u)$, which concludes the proof. □

4.3 Γ-limsup inequality

To conclude the proof of Theorem 2.2 we need to prove the $\limsup$ inequality, which requires distinguishing three cases. If the codimension of the intersection of the two manifolds within the one with the smaller dimension is bigger than two, the problem decouples completely and we show that recovery sequences on each of the manifolds do not interact in a significant way. Otherwise, one has to modify the recovery sequence on the manifold with larger dimension locally around the intersection so that the interaction with the other manifold does not contribute substantially to the energy. Lastly, if the two manifolds have equal dimension and their intersection has codimension one, we have to ensure the regularity of recovery sequence accross the intersection to control the interaction across the intersection. This requires us to prove an approximation result for Sobolev functions on the two manifolds whose traces on the intersection coincide with smooth functions that enjoy the same property, see Theorem A.7 in the appendix.

Before we state and prove the $\limsup$ inequality, we prove two useful lemmas which allow us to crudely bound the nonlocal energy from above by the local one. We shall use these estimates for bounding certain error terms. To prove them, we recall the Cheeger–Colding segment inequality which is satisfied for all Riemannian manifolds with a lower Ricci curvature bound and hence unconditionally for compact manifolds.

Lemma 4.7 (Cheeger–Colding segment inequality [19])**.** *Let $\mathcal{M}$ be a d-dimensional compact, connected Riemannian manifold equipped with the metric $d_{\mathcal{M}}$. For a function $f : \mathcal{M} \to [0, \infty)$ and $x, y \in \mathcal{M}$ let*

$$\mathcal{F}_f(x, y) = \inf_\gamma \int_0^\ell f(\gamma(t)) \, \mathrm{d}t, \qquad x, y \in \mathcal{M},$$

where the infimum is taken over all unit-speed geodesics from x to y, where $\ell = d_{\mathcal{M}}(x, y)$. There exists $r_0 = r_0(\mathcal{M}) > 0$ and a constant $C = C(\mathcal{M}) > 0$ such that for all $0 < r < r_0$, and all open sets $A, B \subset B_{\mathcal{M}}(x_0, r)$ it holds that

$$\iint_{A\times B} \mathcal{F}_f \,\mathrm{dVol}_{\mathcal{M}\times\mathcal{M}} \leq Cr\,(\mathrm{Vol}(A) + \mathrm{Vol}(B)) \int_{B_{\mathcal{M}}(x_0,2r)} f \,\mathrm{dVol}.$$

Lemma 4.8. *Let $\mathcal{M} \subset \mathbb{R}^N$ be a d-dimensional smooth, connected, compact manifold without boundary, $A \subset \mathcal{M}$ be an open subset, and $u \in \mathrm{Lip}(\mathcal{M})$. Under Assumption 2 there exist $C, \varepsilon_0 > 0$, independent of u, such that for all $0 < \varepsilon < \varepsilon_0$ it holds*

$$\int_A \int_{\mathcal{M}} \eta_\varepsilon(|x-y|)\,|u(x) - u(y)|^2 \,\mathrm{dVol}(y)\,\mathrm{dVol}(x) \leq C\varepsilon^{2+d} \int_{A^\varepsilon} |\nabla_{\mathcal{M}} u(y)|^2 \,\mathrm{dVol}(x),$$

where $A^\varepsilon = \{x \in \mathcal{M} \,:\, d_{\mathcal{M}}(x, A) < \varepsilon\}$ is the ε-neighborhood of A in $\mathcal{M}$.

Proof. We use the notation $a \lesssim b$ to denote that $a \leq Cb$ for some constant $C > 0$ independent of a and b. Furthermore, when we write ε sufficiently small, this upper bound is independent of u. For $x, y \in \mathcal{M}$ we let $\gamma_{xy} : [0, d_{\mathcal{M}}(x, y)] \to \mathcal{M}$ denote a unit-speed geodesic between x and y. Note that for $\varepsilon > 0$ sufficiently small and $d_{\mathcal{M}}(x, y) < \varepsilon$ there exists a unique unit-speed geodesic connecting them. Using the fundamental theorem of calculus, Jensen's inequality and Assumption 2 we get

$$\begin{aligned}
&\int_A \int_{\mathcal{M}} \eta_\varepsilon(|x-y|)\,|u(x) - u(y)|^2 \,\mathrm{dVol}(y)\,\mathrm{dVol}(x)\\
&\qquad= \int_A \int_{\mathcal{M}} \eta_\varepsilon(|x-y|) \left| \int_0^{d_{\mathcal{M}}(x,y)} \frac{\mathrm{d}}{\mathrm{d}t} u(\gamma_{xy}(t))\,\mathrm{d}t \right|^2 \mathrm{dVol}(y)\,\mathrm{dVol}(x)\\
&\qquad\leq \int_A \int_{\mathcal{M}} \eta_\varepsilon(|x-y|) d_{\mathcal{M}}(x,y) \underbrace{\int_0^{d_{\mathcal{M}}(x,y)} |\nabla_{\mathcal{M}} u(\gamma_{xy}(t))|^2 \,\mathrm{d}t}_{=:\mathcal{F}(x,y)} \mathrm{dVol}(y)\,\mathrm{dVol}(x)\\
&\qquad\lesssim \varepsilon \int_A \int_{B_{\mathcal{M}}(x,2\varepsilon)} \mathcal{F}(x,y)\,\mathrm{dVol}(y)\,\mathrm{dVol}(x).
\end{aligned}$$

To apply the Cheeger–Colding segment inequality from Theorem 4.7 to this integral we need to localize it. Since A is an open subset of a compact manifold we can cover it by finitely many geodesic balls of radius ε, i.e., $A \subset \bigcup_{k=1}^K B_{\mathcal{M}}(x_k, \varepsilon)$. Hence, we get that

$$\begin{aligned}
\int_A \int_{B_{\mathcal{M}}(x,2\varepsilon)} \mathcal{F}(x,y)\,\mathrm{dVol}(y)\,\mathrm{dVol}(x) &\leq \sum_{k=1}^K \int_{B_{\mathcal{M}}(x_k,\varepsilon)} \int_{B_{\mathcal{M}}(x,2\varepsilon)} \mathcal{F}(x,y)\,\mathrm{dVol}(y)\,\mathrm{dVol}(x)\\
&\leq \sum_{k=1}^K \int_{B_{\mathcal{M}}(x_k,\varepsilon)} \int_{B_{\mathcal{M}}(x_k,3\varepsilon)} \mathcal{F}\,\mathrm{dVol}_{\mathcal{M}\times\mathcal{M}}\\
&\lesssim \varepsilon^{1+d} \sum_{k=1}^K \int_{B_{\mathcal{M}}(x_k,3\varepsilon)} |\nabla_{\mathcal{M}} u|^2 \,\mathrm{dVol},
\end{aligned}$$

where we applied Theorem 4.7 in the last inequality. Note that we can pick a covering such that

$$\sum_{k=1}^K \int_{B_{\mathcal{M}}(x_k,3\varepsilon)} |\nabla_{\mathcal{M}} u|^2 \,\mathrm{dVol} \lesssim \int_{A^\varepsilon} |\nabla_{\mathcal{M}} u|^2 \,\mathrm{dVol}.$$

Combing the three inequalities of this proof we obtain the assertion. □

Lemma 4.9. *Let $\mathcal{M} \subset \mathbb{R}^N$ be a d-dimensional smooth, connected, compact manifold without boundary, $A \subset \mathcal{M}$ be an open subset, $u \in \operatorname{Lip}(\mathcal{M})$, and $T : \mathcal{M} \to \mathcal{M}$ a map which satisfies $\operatorname{ess\,sup}_{x\in\mathcal{M}} |T(x) - x| \leq \delta$ for some $\delta > 0$. Under Assumption 2 there exist $C, \varepsilon_0, \delta_0 > 0$, independent of u, such that for all $0 < \varepsilon < \varepsilon_0$ and all $0 < \delta < \delta_0$ it holds*

$$\int_A \int_{\mathcal{M}} \eta_\varepsilon(|x-y|)\, |u(T(y)) - u(y)|^2 \,\mathrm{dVol}(y)\,\mathrm{dVol}(x) \leq C\varepsilon^d \delta^2 \int_{A^{2\varepsilon}} \operatorname{ess\,sup}_{B_{\mathcal{M}}(x,2\delta)} |\nabla_{\mathcal{M}} u|^2 \,\mathrm{dVol}(x).$$

Proof. Using Fubini's theorem and the fact that $\mathrm{Vol}(B_{\mathcal{M}}(x,\varepsilon)) \lesssim \varepsilon^d$ for $\varepsilon > 0$ sufficiently small, we get

$$\begin{aligned}
&\int_A \int_{\mathcal{M}} \eta_\varepsilon(|x-y|)\, |u(T(y)) - u(y)|^2 \,\mathrm{dVol}(y)\,\mathrm{dVol}(x) \\
&\qquad = \int_{\mathcal{M}} |u(T(y)) - u(y)|^2 \int_A \eta_\varepsilon(|x-y|)\,\mathrm{dVol}(x)\,\mathrm{dVol}(y) \\
&\qquad \lesssim \varepsilon^d \int_{\{y\in\mathcal{M}\,:\,d(y,A)<\varepsilon\}} |u(T(y)) - u(y)|^2 \,\mathrm{dVol}(y).
\end{aligned}$$

For almost every $y \in \mathcal{M}$ we can pick a geodesic $\gamma_y : [0,1] \to \mathcal{M}$ with $\gamma_y(0) = y$ and $\gamma_y(1) = T(y)$. Using the fundamental theorem of calculus and Jensen's inequality we get

$$\begin{aligned}
&\int_A \int_{\mathcal{M}} \eta_\varepsilon(|x-y|)\, |u(T(y)) - u(y)|^2 \,\mathrm{dVol}(y)\,\mathrm{dVol}(x) \\
&\qquad \lesssim \varepsilon^d \int_{\{y\in\mathcal{M}\,:\,d(y,A)<\varepsilon\}} \int_0^1 \left| \frac{\mathrm{d}}{\mathrm{d}s} u(\gamma_y(s)) \right|^2 \mathrm{d}s\,\mathrm{dVol}(y) \\
&\qquad \lesssim \varepsilon^d \delta^2 \int_{\{y\in\mathcal{M}\,:\,d(y,A)<\varepsilon\}} \operatorname{ess\,sup}_{B_{\mathcal{M}}(y,2\delta)} |\nabla_{\mathcal{M}} u|^2 \,\mathrm{dVol}(y),
\end{aligned}$$

where we used that $d_{\mathcal{M}}(y, T(y)) \leq 2\,|y - T(y)| \leq 2\delta$ for $\delta > 0$ sufficiently small. We can conclude the proof by noting that $\{y \in \mathcal{M} : d(y,A) < \varepsilon\} \subset \{y \in \mathcal{M} : d_{\mathcal{M}}(y,A) < 2\varepsilon\} = A^{2\varepsilon}$ for $\varepsilon > 0$ sufficiently small. ☐

Now we continue with the proof of the limsup inequality for Theorem 2.2. For notational convenience, we let $G_{n^{(i)},\varepsilon_n}(\cdot; \mathcal{M}^{(i)})$ denote the standard graph Dirichlet energy from Theorem A.4, evaluated on the $n^{(i)}$ points in $V_n \cap \mathcal{M}^{(i)}$.

Furthermore, as before $\mathcal{G}(\cdot; \mathcal{M}^{(i)})$ denotes the standard local Dirichlet energy from the same theorem, evaluated on the manifold $\mathcal{M}^{(i)}$ using the probability measure $\mu^{(i)} = \rho^{(i)}\mathrm{Vol}^{(i)}$.

Proposition 4.10 (Limsup inequality)**.** *Under Assumptions 1 to 3 for all $u \in L^2(\mathcal{M})$ there exists a sequence $(u_n)_{n\in\mathbb{N}} \subset L^2(\mu_n)$ such that $u_n \xrightarrow{TL^2} u$ and*

$$\limsup_{n\to\infty} E_{n,\varepsilon_n}(u_n) \leq \mathcal{E}(u).$$

Proof. The proof is divided into several cases based on the codimension of the intersection $\mathcal{M}^{(12)}$. Recall that we assume that $d^{(1)} \leq d^{(2)}$.

Case 1, $d^{(1)} \leq d^{(2)}$ and $d^{(1)} - d^{(12)} > 2$. Since the energy $\mathcal{E}(u)$ is uncoupled in this case, one can assume without loss of generality that $u|_{\mathcal{M}^{(i)}}$ for $i \in \{1,2\}$ is smooth, as the general case follows by a standard density argument. Let us define

$$v^{(i)} = \frac{u}{\sqrt{\alpha^{(i)}\beta_\eta^{(i)}\rho^{(i)}}} \quad \text{on } \mathcal{M}^{(i)} \cap V_n.$$

We note

$$\lim_{n\to\infty}\left(\frac{n^{(i)}}{n}\right)^2 G_{n^{(i)},\varepsilon_n}(v^{(i)};\mathcal{M}^{(i)}) = \left(\alpha^{(i)}\right)^2 \mathcal{G}\left(v^{(i)};\mathcal{M}^{(i)}\right), \tag{4.7}$$

where these limits are known to exist almost surely by (2.3) as well as Theorem A.5, in the proof of which a constant recovery sequence is used.

We claim that the constant sequence $u|_{V_n}$ is a recovery sequence for $\mathcal{E}$. To see this, note that for $A_\varepsilon = \mathcal{M}^{(1)} \cup \{x \in \mathcal{M}^{(2)} \,:\, d(x,\mathcal{M}^{(1)}) > \varepsilon\}$ and $S_\varepsilon = \mathcal{M} \setminus A_\varepsilon$ we have

$$\begin{aligned}
E_{n,\varepsilon_n}(u) &= \frac{1}{\varepsilon_n^2}\int_{\mathcal{M}}\int_{\mathcal{M}} \eta_{\varepsilon_n}(|x-y|)\left|\frac{v(x)\sqrt{\alpha^{(i)}\beta^{(i)}\rho^{(i)}(x)}}{\sqrt{\eta_{\varepsilon_n}(|x-\cdot|)\star\mu_n}} - \frac{v(y)\sqrt{\alpha^{(i)}\beta^{(i)}\rho^{(i)}(y)}}{\sqrt{\eta_{\varepsilon_n}(|y-\cdot|)\star\mu_n}}\right|^2 \mathrm{d}\mu_n(x)\,\mathrm{d}\mu_n(y)\\
&\le \frac{1}{\varepsilon_n^2}\int_{A_{\varepsilon_n}}\int_{A_{\varepsilon_n}} \eta_{\varepsilon_n}(|x-y|)\left|\frac{v(x)\sqrt{\alpha^{(i)}\beta^{(i)}\rho^{(i)}(x)}}{\sqrt{\eta_{\varepsilon_n}(|x-\cdot|)\star\mu_n}} - \frac{v(y)\sqrt{\alpha^{(i)}\beta^{(i)}\rho^{(i)}(y)}}{\sqrt{\eta_{\varepsilon_n}(|y-\cdot|)\star\mu_n}}\right|^2 \mathrm{d}\mu_n(x)\,\mathrm{d}\mu_n(y)\\
&\quad + \frac{2}{\varepsilon_n^2}\int_{\mathcal{M}}\int_{S_{\varepsilon_n}} \eta_{\varepsilon_n}(|x-y|)\left|\frac{v(x)\sqrt{\alpha^{(i)}\beta^{(i)}\rho^{(i)}(x)}}{\sqrt{\eta_{\varepsilon_n}(|x-\cdot|)\star\mu_n}} - \frac{v(y)\sqrt{\alpha^{(i)}\beta^{(i)}\rho^{(i)}(y)}}{\sqrt{\eta_{\varepsilon_n}(|y-\cdot|)\star\mu_n}}\right|^2 \mathrm{d}\mu_n(x)\,\mathrm{d}\mu_n(y)\\
&=: E^{far}_{n,\varepsilon_n}(u) + E^{near}_{n,\varepsilon_n}(u).
\end{aligned}$$

Since, by Theorem 4.1, $\chi_{A_{\varepsilon_n}}\varepsilon_n^{-d^{(i)}}\eta_{\varepsilon_n}(|x-\cdot|)\star\mu_n - \chi_{A_{\varepsilon_n}}\alpha^{(i)}\beta^{(i)}\rho^{(i)}(x)$ converges uniformly to zero, and also using (4.7), we have that

$$\limsup_{n\to\infty} E^{far}_{n,\varepsilon_n}(u) \le \limsup_{n\to\infty}\sum_{i=1}^2 \left(\frac{n^{(i)}}{n}\right)^2 G_{n^{(i)},\varepsilon_n}(v^{(i)};\mathcal{M}^{(i)}) \le \sum_{i=1}^2\left(\alpha^{(i)}\right)^2\mathcal{G}\left(v^{(i)};\mathcal{M}^{(i)}\right) = \mathcal{E}(u).$$

So it remains to show that $\lim_{n\to\infty} E^{near}_{n,\varepsilon_n}(u) = 0$. By Remark 4.3, on $\mathcal{M}^{(i)}$ it holds $\varepsilon_n^{-d^{(i)}}\eta_{\varepsilon_n}(|x-\cdot|)\star\mu \ge C > 0$ for n large enough.

Note that by the first inequality of Theorem 4.2 we have $\eta_{\varepsilon_n}(|x-\cdot|)\star\mu_n \llcorner S_{2\varepsilon_n} \le C\varepsilon_n^{(2)}$ and, furthermore, that by a concentration inequality analogous to one of Theorem 4.2 we have that $\mu_n^{(2)}(S_{2\varepsilon_n}) \le C\varepsilon_n^{d^{(2)}-d^{(12)}}$ and $\mu_n^{(1)}(\{x \in \mathcal{M}^{(1)} \,:\, d(x,\mathcal{M}^{(2)}) < \varepsilon_n\}) \le C\varepsilon_n^{d^{(2)}-d^{(12)}}$. Thus, the reminder of the proof is practically identical to the estimate in the proof of limsup inequality in Section 3.3.

Case 2, $d^{(1)} \le d^{(2)}$ and $d^{(1)} - d^{(12)} \le 2 \le d^{(2)} - d^{(12)}$. In this case the previous argument breaks down and the cross term does not vanish in the limit. Thus we have to construct a non-trivial recovery sequence which makes sure that the interaction term between both manifolds converges to zero as $n\to\infty$. As before, the energy $\mathcal{E}(u)$ is uncoupled and one can assume without loss of generality that the functions $u^{(i)} := u|_{\mathcal{M}^{(i)}}$ for $i \in \{1,2\}$ are smooth and apply a diagonal argument to build a recovery sequence for general functions.

For $i \in \{1,2\}$, let us define the normalized Dirichlet energy $\mathcal{E}(\cdot;\mathcal{M}^{(i)})$ on $\mathcal{M}^{(i)}$ as

$$\mathcal{E}(u;\mathcal{M}^{(i)}) := \frac{\sigma_\eta^{(i)}}{\beta_\eta^{(i)}}\int_{\mathcal{M}^{(i)}}\left|\nabla_{\mathcal{M}^{(i)}}\left(\frac{u^{(i)}}{\sqrt{\rho^{(i)}}}\right)\rho^{(i)}\right|^2 \mathrm{dVol}^{(i)}.$$

The recovery sequence which we shall construct is denoted by u_n and for vertices lying in the lower-dimensional manifold $\mathcal{M}^{(1)}$ it is defined by

$$u_n(x) := u^{(1)}(x)\frac{\sqrt{\varepsilon_n^{-d^{(1)}}\deg_{n,\varepsilon_n}(x)}}{\sqrt{\alpha^{(1)}\beta_\eta^{(1)}\rho^{(1)}(x)}}, \quad x \in V_n \cap \mathcal{M}^{(1)}. \tag{4.8}$$

Note that from Theorem 4.3 it follows that u_n converges uniformly to $u^{(1)}$ on $\mathcal{M}^{(1)}$.

On the higher-dimensional one, we need a more subtle modification of the function $u^{(2)}$. For this let us take $\varepsilon_n > 0$ so small that the geodesic projection $\pi : \mathcal{M}^{(2)} \to \mathcal{M}^{(12)}$ onto the intersection is uniquely defined in an ε_n-neighborhood of $\mathcal{M}^{(12)}$. We abbreviate the distance of $x \in \mathcal{M}^{(2)}$ to $\mathcal{M}^{(12)}$ as $d(x) := \operatorname{dist}(x, \mathcal{M}^{(12)})$ and subdivide the higher-dimensional manifold $\mathcal{M}^{(2)}$ into a near, a mid, and a far region with respect to the distance to the intersection $\mathcal{M}^{(12)}$:

$$
\begin{aligned}
\mathcal{M}^{(2)}_{\text{near}} &:= \{x \in \mathcal{M}^{(2)} \,:\, d(x) < C\varepsilon_n\},\\
\mathcal{M}^{(2)}_{\text{mid}} &:= \{x \in \mathcal{M}^{(2)} \,:\, C\varepsilon_n \le d(x) < C\sqrt{\varepsilon_n}\},\\
\mathcal{M}^{(2)}_{\text{far}} &:= \{x \in \mathcal{M}^{(2)} \,:\, d(x) > C\sqrt{\varepsilon_n}\}.
\end{aligned}
$$

Here, the constant $C > 0$ is chosen such that for $n \in \mathbb{N}$ large enough (and hence $\varepsilon_n > 0$ small enough) we have that $B(x, \varepsilon) \cap \mathcal{M}^{(1)} = \emptyset$ for all $x \in \mathcal{M}^{(2)} \setminus \mathcal{M}^{(2)}_{\text{near}}$ which is possible thanks to Theorem A.1.

With the partitioning of $\mathcal{M}^{(2)}$ into a near, a middle, and a far region at hand, we can now define the recovery sequence as a logarithmic interpolation between the value of $u^{(1)}$ on the intersection $\mathcal{M}^{(12)}$ and $u^{(2)}$ in the middle region $\mathcal{M}^{(2)}_{\text{mid}}$. More precisely, it is defined as

$$
u_n^{(2)}(x) := \begin{cases}
u^{(1)}(\pi(x))\sqrt{\dfrac{\varepsilon_n^{-d^{(1)}}\deg_{n,\varepsilon_n}(x)}{\alpha^{(1)}\beta_\eta^{(1)}\rho^{(1)}(\pi(x))}}, & \text{if } x \in \mathcal{M}^{(2)}_{\text{near}},\\[2ex]
u^{(2)}(x)\sqrt{\dfrac{\varepsilon_n^{-d^{(2)}}\deg_{n,\varepsilon_n}(x)}{\alpha^{(2)}\beta_\eta^{(2)}\rho^{(2)}(x)}} + \dfrac{\log\sqrt{\varepsilon_n} - \log d(x)}{\log\sqrt{\varepsilon_n} - \log\varepsilon_n}\times & \\[2ex]
\times\left(u^{(1)}(\pi(x))\sqrt{\dfrac{\varepsilon_n^{-d^{(1)}}\deg_{n,\varepsilon_n}(x)}{\alpha^{(1)}\beta_\eta^{(1)}\rho^{(1)}(\pi(x))}} - u^{(2)}(x)\sqrt{\dfrac{\varepsilon_n^{-d^{(2)}}\deg_{n,\varepsilon_n}(x)}{\alpha^{(2)}\beta_\eta^{(2)}\rho^{(2)}(x)}}\right), & \text{if } x \in \mathcal{M}^{(2)}_{\text{mid}},\\[2ex]
u^{(2)}(x)\sqrt{\dfrac{\varepsilon_n^{-d^{(2)}}\deg_{n,\varepsilon_n}(x)}{\alpha^{(2)}\beta_\eta^{(2)}\rho^{(2)}(x)}}, & \text{if } x \in \mathcal{M}^{(2)}_{\text{far}},
\end{cases}
\tag{4.9}
$$

Our goal is to show $u_n \xrightarrow{TL^2} u$ as well as the upper bound $\limsup_{n\to\infty} E_{n,\varepsilon_n}(u_n) \le \mathcal{E}(u)$ which would make u_n a recovery sequence. First note that $u_n \to u$ in the TL^2-sense follows from similar arguments as Theorem 4.4 using $d^{(1)} = d^{(2)}$, the Lipschitz continuity of $u^{(i)}$ for $i \in \{1, 2\}$, and Assumption 1.

For proving the desired upper bound, we note that we can split the energy $E_{n,\varepsilon_n}(u_n)$ into a sum over vertices in $\mathcal{M}^{(1)} \times \mathcal{M}^{(1)}$, two over $\mathcal{M}^{(1)} \times \mathcal{M}^{(2)}$, and one over $\mathcal{M}^{(2)} \times \mathcal{M}^{(2)}$ which we now estimate. For the sum over $\mathcal{M}^{(1)} \times \mathcal{M}^{(1)}$ we observe that by the definition of u_n it holds that

$$
\begin{aligned}
&\limsup_{n\to\infty} \frac{1}{n^2\varepsilon_n^2} \sum_{x,y\in V_n\cap\mathcal{M}^{(1)}} \eta_{\varepsilon_n}(|x-y|)\left|\frac{u_n(x)}{\sqrt{\deg_{n,\varepsilon_n}(x)}} - \frac{u_n(y)}{\sqrt{\deg_{n,\varepsilon_n}(y)}}\right|^2\\
&= \frac{1}{\alpha^{(1)}\beta_\eta^{(1)}} \limsup_{n\to\infty}\left(\frac{n^{(1)}}{n}\right)^2 \frac{1}{(n^{(1)})^2\varepsilon_n^2} \sum_{x,y\in V_n\cap\mathcal{M}^{(1)}} \eta_{\varepsilon_n}(|x-y|)\left|\frac{u^{(1)}(x)}{\sqrt{\rho^{(1)}(x)}} - \frac{u^{(1)}(y)}{\sqrt{\rho^{(1)}(y)}}\right|^2\\
&= \alpha^{(1)}\mathcal{E}(u^{(1)}; \mathcal{M}^{(1)})
\end{aligned}
\tag{4.10}
$$

since the Lipschitz function $\frac{u^{(1)}}{\sqrt{\rho^{(1)}}}$ is a recovery sequence for $\mathcal{E}(u^{(1)}; \mathcal{M}^{(1)})$.

We continue with the sum over vertices $\mathcal{M}^{(1)} \times \mathcal{M}^{(2)}$. We note that $u_n^{(2)}$ is defined in such a way that these cross-interaction terms are small.

$$
\begin{aligned}
&\frac{1}{n^2\varepsilon_n^2} \sum_{\substack{y\in V_n\cap\mathcal{M}^{(1)}\\ x\in V_n\cap\mathcal{M}^{(2)}}} \eta_{\varepsilon_n}(|x-y|) \left| \frac{u_n(x)}{\sqrt{\deg_{n,\varepsilon_n}(x)}} - \frac{u_n(y)}{\sqrt{\deg_{n,\varepsilon_n}(y)}} \right|^2 \\
&= \frac{1}{n^2\varepsilon_n^{2+d^{(1)}}} \frac{1}{\alpha^{(1)}\beta_\eta^{(1)}} \sum_{\substack{y\in V_n\cap\mathcal{M}^{(1)}\\ x\in V_n\cap\mathcal{M}^{(2)}}} \eta_{\varepsilon_n}(|x-y|) \left| \frac{u^{(1)}(\pi(x))}{\sqrt{\rho^{(1)}(\pi(x))}} - \frac{u^{(1)}(y)}{\sqrt{\rho^{(1)}(y)}} \right|^2 \\
&\leq \frac{C}{\varepsilon_n^{2+d^{(1)}}} \int_{\mathcal{M}^{(1)}} \int_{\mathcal{M}^{(2)}} \eta_{\varepsilon_n}\left(\left| T_n^{(2)}(x) - T_n^{(1)}(y) \right| \right) \left| \frac{u^{(1)}(\pi(T_n^{(2)}(x)))}{\sqrt{\rho^{(1)}(\pi(T_n^{(2)}(x)))}} - \frac{u^{(1)}(T_n^{(1)}(y))}{\sqrt{\rho^{(1)}(T_n^{(1)}(y))}} \right|^2 \\
&\qquad\qquad \mathrm{d}\mu^{(2)}(x)\,\mathrm{d}\mu^{(1)}(y) \\
&\leq C \frac{\tilde{\varepsilon}_n^2}{\varepsilon_n^{2+d^{(1)}}} \int_{\mathcal{M}^{(1)}} \int_{\mathcal{M}^{(2)}} \eta_{\tilde{\varepsilon}_n}(|x-y|)\,\mathrm{d}\mu^{(2)}(x)\,\mathrm{d}\mu^{(1)}(y) \\
&= O\left(\varepsilon_n^{-d^{(1)}} \tilde{\varepsilon}_n^{d^{(1)}+d^{(2)}-d^{(12)}} \right) = O(\varepsilon_n^2) \quad \text{as } n \to \infty. \qquad (4.11)
\end{aligned}
$$

where we used that $u^{(1)}/\sqrt{\rho^{(1)}}$ is Lipschitz continuous and that the codimension $d^{(2)} - d^{(12)}$ is at least two. As in Case 1, we assumed without loss of generality that $\eta(t) = 1_{[0,1]}(t)$, used the quantity $\tilde{\varepsilon}_n$ which satisfies $\tilde{\varepsilon}_n > \varepsilon_n$ and $\lim_{n\to\infty} \frac{\varepsilon_n}{\tilde{\varepsilon}_n} = 1$, and utilized the same transport maps $T_n^{(i)} : \mathcal{M}^{(i)} \to \mathcal{M}^{(i)}$. Furthermore, the constant C changes its value betweem the lines.

It remains to estimate the sum over $\mathcal{M}^{(2)} \times \mathcal{M}^{(2)}$ for which we introduce the following abbreviation

$$
S(n,\varepsilon_n) := \left\{ x \in V_n \cap \mathcal{M}^{(2)} \,:\, x \in \mathcal{M}^{(2)}_{\text{near}} \cup \mathcal{M}^{(2)}_{\text{mid}} \text{ or } \operatorname{dist}(x, \mathcal{M}^{(2)}_{\text{mid}}) \leq \varepsilon_n \right\}
$$

for the set of graph points that fall into the near or mid region, or fall into the far region but have Euclidean distance at most ε to the mid region. In particular we have that

$$
\left(V_n \cap \mathcal{M}^{(2)} \right) \setminus S(n,\varepsilon_n) = \left\{ x \in V_n \cap \mathcal{M}^{(2)} \,:\, x \in \mathcal{M}^{(2)}_{\text{far}} \text{ and } \operatorname{dist}(x, \mathcal{M}^{(2)}_{\text{mid}}) > \varepsilon_n \right\},
$$

so points in that set do not interact with points in the near or the middle region. Using this definition and the fact that $u_n = u^{(2)} \sqrt{\frac{\varepsilon_n^{-d^{(2)}} \deg_{n,\varepsilon_n}}{\alpha^{(2)}\beta^{(2)}\rho^{(2)}}}$ on $V_n \cap \mathcal{M}^{(2)}_{\text{far}}$ we split the sum into the terms that account for interaction only in the far region, that accounts for the bulk of the energy on $\mathcal{M}^{(2)}$, and a term that accounts for the transition layers and their interactions within $\mathcal{M}^{(2)}$, which we show vanish in the limit

$$
\begin{aligned}
&\frac{1}{n^2\varepsilon_n^2}\sum_{x,y\in V_n\cap\mathcal{M}^{(2)}}\eta_{\varepsilon_n}(|x-y|)\left|\frac{u_n(x)}{\sqrt{\deg_{n,\varepsilon_n}(x)}}-\frac{u_n(y)}{\sqrt{\deg_{n,\varepsilon_n}(y)}}\right|^2\\
&\quad=\frac{1}{n^2\varepsilon_n^2\alpha^{(2)}\beta_\eta^{(2)}}\sum_{\substack{x\in V_n\cap\mathcal{M}^{(2)}_{\mathrm{far}}\\ x\notin S(n,\varepsilon_n)}}\sum_{y\in V_n\cap\mathcal{M}^{(2)}}\eta_{\varepsilon_n}(|x-y|)\left|\frac{u^{(2)}(x)}{\sqrt{\rho^{(2)}(x)}}-\frac{u^{(2)}(y)}{\sqrt{\rho^{(2)}(y)}}\right|^2\\
&\qquad+\frac{1}{n^2\varepsilon_n^2}\sum_{x\in S(n,\varepsilon_n)}\sum_{y\in V_n\cap\mathcal{M}^{(2)}}\eta_{\varepsilon_n}(|x-y|)\,|\tilde u_n(x)-\tilde u_n(y)|^2\\
&\quad\le\frac{1}{n^2\varepsilon_n^2\alpha^{(2)}\beta_\eta^{(2)}}\sum_{x,y\in V_n\cap\mathcal{M}^{(2)}}\eta_{\varepsilon_n}(|x-y|)\left|\frac{u^{(2)}(x)}{\sqrt{\rho^{(2)}(x)}}-\frac{u^{(2)}(y)}{\sqrt{\rho^{(2)}(y)}}\right|^2\\
&\qquad+\frac{1}{n^2\varepsilon_n^2}\sum_{x\in S(n,\varepsilon_n)}\sum_{y\in V_n\cap\mathcal{M}^{(2)}}\eta_{\varepsilon_n}(|x-y|)\,|\tilde u_n(x)-\tilde u_n(y)|^2 \qquad (4.12)
\end{aligned}
$$

where the function $\tilde u_n$ is the recovery sequence divided by the square root of the degrees and equals

$$
\tilde u_n(x)=\begin{cases}
\dfrac{u^{(1)}(\pi(x))}{\sqrt{\varepsilon_n^{d^{(1)}}\alpha^{(1)}\beta^{(1)}\rho^{(1)}(\pi(x))}}, & x\in\mathcal{M}^{(2)}_{\mathrm{near}},\\
\dfrac{u^{(2)}(x)}{\sqrt{\varepsilon_n^{d^{(2)}}\alpha^{(2)}\beta_\eta^{(2)}\rho^{(2)}(x)}}+\dfrac{\log\sqrt{\varepsilon_n}-\log d(x)}{\log\sqrt{\varepsilon_n}-\log\varepsilon_n}\times & \\
\times\left(\dfrac{u^{(1)}(\pi(x))}{\sqrt{\varepsilon_n^{d^{(1)}}\alpha^{(1)}\beta^{(1)}\rho^{(1)}(\pi(x))}}-\dfrac{u^{(2)}(x)}{\sqrt{\varepsilon_n^{d^{(2)}}\alpha^{(2)}\beta^{(2)}\rho^{(2)}(x)}}\right), & x\in\mathcal{M}^{(2)}_{\mathrm{mid}},\\
\dfrac{u^{(2)}(x)}{\sqrt{\varepsilon_n^{d^{(2)}}\alpha^{(2)}\beta^{(2)}\rho^{(2)}(x)}}, & x\in\mathcal{M}^{(2)}_{\mathrm{far}}.
\end{cases} \qquad (4.13)
$$

Since the function $\frac{u^{(2)}}{\sqrt{\rho^{(2)}}}$ is Lipschitz, similar to before we have for the first term in (4.12) that

$$
\limsup_{n\to\infty}\frac{1}{n^2\varepsilon_n^2\beta_\eta^{(2)}}\sum_{x,y\in V_n\cap\mathcal{M}^{(2)}}\eta_{\varepsilon_n}(|x-y|)\left|\frac{u^{(2)}(x)}{\sqrt{\rho^{(2)}(x)}}-\frac{u^{(2)}(y)}{\sqrt{\rho^{(2)}(y)}}\right|^2\le\alpha^{(2)}\mathcal{E}(u^{(2)};\mathcal{M}^{(2)}). \qquad (4.14)
$$

We claim that the second term in (4.12) goes to zero as $n\to\infty$ and we estimate it as follows

$$
\begin{aligned}
&\frac{1}{n^2\varepsilon_n^2}\sum_{x\in S(n,\varepsilon_n)}\sum_{y\in V_n\cap\mathcal{M}^{(2)}}\eta_{\varepsilon_n}(|x-y|)\,|\tilde u_n(x)-\tilde u_n(y)|^2\\
&=\left(\frac{n^{(2)}}{n}\right)^2\frac{1}{\varepsilon_n^2}\int_{D(n,\varepsilon_n)}\int_{\mathcal{M}^{(2)}}\eta_{\varepsilon_n}\left(\left|T_n^{(2)}(x)-T_n^{(2)}(y)\right|\right)\left|\tilde u_n(T_n^{(2)}(x))-\tilde u_n(T_n^{(2)}(y))\right|^2\,\mathrm{d}\mu^{(2)}(x)\,\mathrm{d}\mu^{(2)}(y)\\
&\le\left(\frac{n^{(2)}}{n}\right)^2\frac{1}{\varepsilon_n^2}\int_{D(n,\varepsilon_n)}\int_{\mathcal{M}^{(2)}}\eta_{\tilde\varepsilon_n}(|x-y|)\left|\tilde u_n\circ T_n^{(2)}(x)-\tilde u_n\circ T_n^{(2)}(y)\right|^2\,\mathrm{d}\mu^{(2)}(x)\,\mathrm{d}\mu^{(2)}(y),
\end{aligned}
$$

where $D(n,\varepsilon_n) = \{x \in \mathcal{M}^{(2)} : T_n^{(2)}(x) \in S(n,\varepsilon_n)\}$. We claim that in the limit $n \to \infty$ we can replace $\tilde{u}_n \circ T_n^{(2)}$ by $\tilde{u}_n$. To show this, similarly as in [27, p.267], we define

$$A_n := \frac{1}{\varepsilon_n^2} \int_{D(n,\varepsilon_n)} \int_{\mathcal{M}^{(2)}} \eta_{\tilde{\varepsilon}_n} (|x-y|) \left|\tilde{u}_n(x) - \tilde{u}_n(y)\right|^2 \, \mathrm{d}\mu^{(2)}(x) \, \mathrm{d}\mu^{(2)}(y)$$
$$B_n := \frac{1}{\varepsilon_n^2} \int_{D(n,\varepsilon_n)} \int_{\mathcal{M}^{(2)}} \eta_{\tilde{\varepsilon}_n} (|x-y|) \left|\tilde{u}_n \circ T_n^{(2)}(x) - \tilde{u}_n \circ T_n^{(2)}(y)\right|^2 \, \mathrm{d}\mu^{(2)}(x) \, \mathrm{d}\mu^{(2)}(y).$$

Applying Theorem 4.9 we obtain that

$$\begin{aligned}
&(\sqrt{A_n} - \sqrt{B_n})^2 \\
&\leq \frac{1}{\varepsilon_n^2} \int_{D(n,\varepsilon_n)} \int_{\mathcal{M}^{(2)}} \eta_{\tilde{\varepsilon}_n} (|x-y|) \left|\tilde{u}_n \circ T_n^{(2)}(x) - \tilde{u}_n(x) + \tilde{u}_n(y) - \tilde{u}_n \circ T_n^{(2)}(y)\right|^2 \, \mathrm{d}\mu^{(2)}(x) \, \mathrm{d}\mu^{(2)}(y) \\
&\leq \frac{2}{\varepsilon_n^2} \int_{D(n,\varepsilon_n)} \int_{\mathcal{M}^{(2)}} \eta_{\tilde{\varepsilon}_n} (|x-y|) \left|\tilde{u}_n \circ T_n^{(2)}(x) - \tilde{u}_n(x)\right|^2 + \left|\tilde{u}_n(y) - \tilde{u}_n \circ T_n^{(2)}(y)\right|^2 \, \mathrm{d}\mu^{(2)}(x) \, \mathrm{d}\mu^{(2)}(y) \\
&\lesssim \left(\frac{\delta_n}{\varepsilon_n}\right)^2 \tilde{\varepsilon}_n^{d^{(2)}} \int_{D(n,\varepsilon_n)^{2\varepsilon_n}} \sup_{B_{\mathcal{M}}(x,2\delta_n)} \left|\nabla_{\mathcal{M}^{(2)}} \tilde{u}_n\right|^2 \, \mathrm{dVol}^{(2)}(x),
\end{aligned} \tag{4.15}$$

where $\delta_n := \max_{i\in\{1,2\}} \left\|T_n^{(i)} - \mathrm{id}\right\|_{L^\infty(\mathcal{M}^{(i)})}$. Applying Theorem 4.8 we also get that

$$A_n \lesssim \varepsilon_n^{d^{(2)}} \int_{D(n,\varepsilon_n)^{\varepsilon_n}} \left|\nabla_{\mathcal{M}^{(2)}} \tilde{u}_n\right|^2 \, \mathrm{dVol}^{(2)}. \tag{4.16}$$

Let us assume for now that we already know that $(\sqrt{A_n} - \sqrt{B_n})^2$ in (4.15) goes to zero as $n \to \infty$ and that also A_n goes to zero by (4.16). This implies that $|A_n - B_n| \to 0$ as $n \to \infty$ and, by applying Theorem 4.8 again, we can estimate the limies superior of the second term in (4.12) as follows:

$$\begin{aligned}
&\limsup_{n\to\infty} \frac{1}{n^2\varepsilon_n^2} \sum_{x\in S(n,\varepsilon_n)} \sum_{y \in V_n \cap \mathcal{M}^{(2)}} \eta_{\varepsilon_n}(|x-y|) \left|\tilde{u}_n(x) - \tilde{u}_n(y)\right|^2 \\
&\qquad \leq \left(\alpha^{(2)}\right)^2 \limsup_{n\to\infty} A_n \leq \left(\alpha^{(2)}\right)^2 \limsup_{n\to\infty} \varepsilon_n^{d^{(2)}} \int_{D(n,\varepsilon_n)^{\varepsilon_n}} \left|\nabla_{\mathcal{M}^{(2)}} \tilde{u}_n\right|^2 \, \mathrm{dVol}^{(2)} = 0
\end{aligned}$$

and therefore as a consequence of (4.12) and (4.14) we get

$$\limsup_{n\to\infty} \frac{1}{n^2\varepsilon_n^2} \sum_{x,y\in V_n \cap \mathcal{M}^{(2)}} \eta_{\varepsilon_n}(|x-y|) \left|\frac{u_n(x)}{\sqrt{\deg_{n,\varepsilon_n}(x)}} - \frac{u_n(y)}{\sqrt{\deg_{n,\varepsilon_n}(y)}}\right|^2 \leq \alpha^{(2)} \mathcal{E}(u^{(2)}; \mathcal{M}^{(2)}).$$

In combination with (4.10) and (4.11) and the definition of $\mathcal{E}$ in (2.9) this would conclude the proof.

For showing that the right hand sides in (4.15) and (4.16) indeed converge to zero, and taking into account Assumption 3 and Remark 2.3, it suffices to prove that

$$\lim_{n\to\infty} \varepsilon_n^{d^{(2)}} \int_{D(n,\varepsilon_n)^{2\varepsilon_n}} \sup_{B_{\mathcal{M}}(x,2\delta_n)} \left|\nabla_{\mathcal{M}^{(2)}} \tilde{u}_n\right|^2 \, \mathrm{dVol}^{(2)}(x) = 0.$$

From the definition of $\tilde{u}_n$ in (4.13) it is immediately clear that

$$\lim_{n\to\infty} \varepsilon_n^{d^{(2)}} \int_{A} \sup_{B_{\mathcal{M}}(x,\delta_n)} \left|\nabla_{\mathcal{M}^{(2)}} \tilde{u}_n\right|^2 \, \mathrm{dVol}^{(2)}(x) = 0$$

if $A \subset \{x \in \mathcal{M}^{(2)}_{\text{near}} \,:\, \operatorname{dist}(x, \mathcal{M}^{(2)}_{\text{mid}}) > 2\delta_n\}$ or if $A \subset \{x \in \mathcal{M}^{(2)}_{\text{far}} \,:\, \operatorname{dist}(x, \mathcal{M}^{(2)}_{\text{mid}}) > 2\delta_n\}$. Therefore, it suffices to compute the integral over the following neighborhood of the middle region $A := \{x \in \mathcal{M}^{(2)} \,:\, \operatorname{dist}(x, \mathcal{M}^{(2)}_{\text{mid}}) < 2\delta_n\}$. Taking into account the definition of $\tilde{u}_n$ in (4.13) and using the product rule we see that on $\mathcal{M}^{(2)}_{\text{mid}}$ the gradient $\nabla_{\mathcal{M}^{(2)}} \tilde{u}_n$ consists of three terms: The first one is the gradient of $\frac{u^{(1)}\circ\pi}{\sqrt{\varepsilon_n^{d^{(1)}}\alpha^{(1)}\beta^{(1)}\rho^{(1)}\circ\rho}}$, the second one is the the gradient of the logarithmic interpolation times the round brackets, and the third one is the logarithmic interpolation factor (which lies between 0 and 1) times the gradient of the round brackets. We note that the terms whose gradients are bounded can be neglected since the size of the domain vanishes as $n \to \infty$ and thus their contribution vanishes in the limit. Since the functions $u^{(i)}/\sqrt{\rho^{(i)}}$ for $i = 1, 2$ are Lipschitz continuous it hence suffices to consider the integral over the second of these terms. Using that

$$\left|\nabla_{\mathcal{M}^{(2)}} \frac{\log\sqrt{\varepsilon_n} - \log d(x)}{\log\sqrt{\varepsilon_n} - \log \varepsilon_n}\right| = \frac{2}{|\log \varepsilon_n|\, d(x)}$$

we have

$$\sup_{B_{\mathcal{M}}(x,2\delta_n)} \left|\nabla_{\mathcal{M}^{(2)}} \frac{\log\sqrt{\varepsilon_n} - \log d(x)}{\log\sqrt{\varepsilon_n} - \log \varepsilon_n}\right|^2 = \frac{4}{|\log \varepsilon_n|^2\,(d(x) - 2\delta_n)^2}.$$

Hence, using also the coarea formula we have

$$\begin{aligned}
&\varepsilon_n^{d^{(2)}} \left(\left\| \frac{u^{(1)}\circ\pi}{\sqrt{\varepsilon^{d^{(1)}}\alpha^{(1)}\beta^{(1)}\rho^{(1)}\circ\pi}} \right\|_\infty^2 + \left\| \frac{u^{(2)}}{\sqrt{\varepsilon^{d^{(2)}}\alpha^{(2)}\beta^{(2)}\rho^{(2)}}} \right\|_\infty^2 \right) \times \\
&\times \int_A \sup_{B_{\mathcal{M}}(x,2\delta_n)} \left|\nabla_{\mathcal{M}^{(2)}} \frac{\log\sqrt{\varepsilon_n} - \log d(x)}{\log\sqrt{\varepsilon_n} - \log \varepsilon_n}\right|^2 \mathrm{dVol}^{(2)}(x) \\
&\quad \lesssim \frac{1}{|\log\varepsilon_n|^2} \int_A \frac{1}{(d(x) - 2\delta_n)^2}\,\mathrm{dVol}^{(2)}(x) \\
&\quad \lesssim \frac{1}{|\log\varepsilon_n|^2} \int_{C\varepsilon_n - 2\delta_n}^{C\sqrt{\varepsilon_n}+2\delta_n} \mathcal{H}^{d^{(2)}-1}\left(\{y \in \mathcal{M}^{(2)} \,:\, d(y) = t - 2\delta_n\}\right) \frac{1}{(t-2\delta_n)^2}\,\mathrm{d}t \\
&\quad \lesssim \frac{1}{|\log\varepsilon_n|^2} \int_{C\varepsilon_n - 4\delta_n}^{C\sqrt{\varepsilon_n}} \mathcal{H}^{d^{(2)}-1}\left(\{y \in \mathcal{M}^{(2)} \,:\, d(y) = s\}\right) \frac{1}{s^2}\,\mathrm{d}s \\
&\quad \lesssim \frac{1}{|\log\varepsilon_n|^2} \int_{C\varepsilon_n - 4\delta_n}^{C\sqrt{\varepsilon_n}} s^{d^{(2)} - d^{(12)} - 3}\,\mathrm{d}s = O\left(\frac{1}{|\log\varepsilon_n|}\right)
\end{aligned}$$

where we used $d^{(2)} - d^{(12)} \geq 2$ and inferred $\lim_{n\to\infty} \frac{\delta_n}{\varepsilon_n} = 0$ from Assumption 3 and Remark 2.3.

Case 3, $d^{(1)} = d^{(2)} =: d$ **and** $d - d^{(12)} = 1$ Using Theorem A.6 as in the proof of the nonlocal limsup inequality for Theorem 3.2 in Section 3.3, we can assume that $v := u/\sqrt{\rho}$ is Lipschitz continuous on $\mathcal{M}$. Similar to the argument there, we consider the sequence

$$u_n(x) := u(x)\sqrt{\frac{\varepsilon_n^{-d}\deg_{n,\varepsilon_n}(x)}{\alpha^{(i)}\beta_\eta\rho^{(i)}(x)}}, \qquad x \in \mathcal{M}^{(i)} \cap V_n,$$

where we dropped the index i on β_η since $d^{(1)} = d^{(2)} = d$. Again, the idea is that the square-root term converges to one and its numerator cancels once plugged into the normalized graph Dirichlet energy E_{n,ε_n}.

Using the usual definition $v := \frac{u}{\sqrt{\alpha^{(i)}\beta_\eta \rho^{(i)}}}$ we have

$$E_{n,\varepsilon_n}(u_n) = \sum_{i=1}^{2} \left(\frac{n^{(i)}}{n}\right)^2 G_{n,\varepsilon_n}\left(v; \mathcal{M}^{(i)}\right) + \frac{2}{n^2 \varepsilon_n^{2+d}} \sum_{\substack{x\in V_n\cap\mathcal{M}^{(1)} \\ y\in V_n\cap\mathcal{M}^{(2)}}} \eta_{\varepsilon_n}(|x-y|)\,|v(x)-v(y)|^2 .$$

Note that v is globally Lipschitz on $\mathcal{M}$. In particular, by Theorem A.5 (in the proof of which a constant recovery is used for Lipschitz functions) and the definition of $\mathcal{E}$ in (2.9) we have almost surely

$$\limsup_{n\to\infty} \sum_{i=1}^{2} \left(\frac{n^{(i)}}{n}\right)^2 G_{\varepsilon_n}\left(v; \mathcal{M}^{(i)}\right) = \sum_{i=1}^{2} \left(\alpha^{(i)}\right)^2 \mathcal{G}\left(v; \mathcal{M}^{(i)}\right) = \mathcal{E}(u).$$

It remains to show that the cross term in the above energy decomposition of $E_{n,\varepsilon_n}(u_n)$ tends to zero for which we shall use Lipschitzness of v:

$$\begin{aligned}
&\limsup_{n\to\infty} \frac{2}{n^2\varepsilon_n^{2+d}} \sum_{\substack{x\in V_n\cap\mathcal{M}^{(1)} \\ y\in V_n\cap\mathcal{M}^{(2)}}} \eta_{\varepsilon_n}(|x-y|)\,|v(x)-v(y)|^2 \\
&\leq \limsup_{n\to\infty} \frac{\operatorname{Lip}(v)^2}{n^2\varepsilon_n^{d}} \sum_{\substack{x\in V_n\cap\mathcal{M}^{(1)} \\ y\in V_n\cap\mathcal{M}^{(2)}}} \eta_{\varepsilon_n}(|x-y|) \\
&= \limsup_{n\to\infty} \frac{n^{(1)}n^{(2)}}{n^2} \frac{\operatorname{Lip}(v)^2}{\varepsilon_n^{d}} \int_{\mathcal{M}^{(1)}} \int_{\mathcal{M}^{(2)}} \eta_{\varepsilon_n}\left(\left|T_n^{(1)}(x) - T_n^{(2)}(y)\right|\right) \,\mathrm{d}\mu^{(2)}(x)\,\mathrm{d}\mu^{(1)}(y) \\
&\leq \limsup_{n\to\infty} \alpha^{(1)}\alpha^{(2)} \frac{\operatorname{Lip}(v)^2}{\tilde\varepsilon_n^{d}} \int_{\mathcal{M}^{(1)}} \int_{\mathcal{M}^{(2)}} \eta_{\tilde\varepsilon_n}\left(|x-y|\right) \,\mathrm{d}\mu^{(2)}(x)\,\mathrm{d}\mu^{(1)}(y) \\
&\leq \limsup_{n\to\infty} \frac{C\operatorname{Lip}(v)^2}{\tilde\varepsilon_n^{d}} \tilde\varepsilon_n^{d-d^{(12)}} \tilde\varepsilon_n^{d} = 0,
\end{aligned}$$

where $C > 0$ is a constant and we used that $d - d^{(12)} = 1$.

Case 4, $d^{(1)} - d^{(12)} < 2$ **and** $d^{(2)} - d^{(12)} < 2$ We have the standing assumption that $d^{(12)} < d^{(1)}$ which in this case is only possible if $d^{(12)} = 0$ and $d^{(1)} = 1$. Then, however, we also have that $d^{(2)} = 1 = d^{(1)}$ which takes us back to Case 3.

□

Acknowledgments

The authors would like to thank Simone Di Marino and for enlightening discussions, Tim Laux for pointing them to the Cheeger–Colding segment inequality, and Anton Ullrich for careful reading of the manuscript and valuable comments. This work was initiated while the authors were visiting the Simons Institute for the Theory of Computing to participate in the program "Geometric Methods in Optimization and Sampling" during the Fall of 2021, and they are very grateful for the hospitality of the institute. Part of this work was done while LB was affiliated with the Hausdorff Center for Mathematics at the University of Bonn. LB acknowledges funding by the German Ministry of Science and Technology (BMBF) under grant agreement No. 01IS24072A (COMFORT) and by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – project number 544579844 (GeoMAR). DS is grateful to NSF for support via NSF RTG grant 2342349 and NSF grant DMS-2511684. The authors are grateful to CMU's Center for Nonlinear Analysis for its support.

References

[1] Michele Allegra, Elena Facco, Francesco Denti, Alessandro Laio, and Antonietta Mira. "Data segmentation based on the local intrinsic dimension". *Scientific Reports* 10.1 (2020), p. 16449. `doi.org/10.1038/s41598-020-72222-0`.

[2] Patricia Alonso Ruiz and Fabrice Baudoin. *Dirichlet forms on metric measure spaces as Mosco limits of Korevaar-Schoen energies*. 2023. `arxiv.org/abs/2301.08273`.

[3] Ery Arias-Castro, Guangliang Chen, and Gilad Lerman. "Spectral clustering based on local linear approximations". *Electronic Journal of Statistics* 5.none (Jan. 2011). `dx.doi.org/10.1214/11-EJS651`.

[4] Scott Armstrong and Raghavendra Venkatraman. "Optimal convergence rates for the spectrum of the graph Laplacian on Poisson point clouds". *Foundations of Computational Mathematics* (2025), pp. 1–26. `doi.org/10.1007/s10208-025-09696-9`.

[5] Amir Babaeian, Alireza Bayestehtashk, and Mojtaba Bandarabadi. "Multiple manifold clustering using curvature constrained path". *PloS one* 10.9 (2015), e0137986. `doi.org/10.1371/journal.pone.0137986`.

[6] Mikhail Belkin and Partha Niyogi. "Laplacian eigenmaps and spectral techniques for embedding and clustering". *Advances in Neural Information Processing Systems* 14 (2001). `doi.org/10.7551/mitpress/1120.003.0080`.

[7] Mikhail Belkin and Partha Niyogi. "Laplacian eigenmaps for dimensionality reduction and data representation". *Neural computation* 15.6 (2003), pp. 1373–1396. `dl.acm.org/doi/10.1162/089976603321780317`.

[8] Mikhail Belkin and Partha Niyogi. "Convergence of Laplacian eigenmaps". *Advances in Neural Information Processing Systems* 19 (2007), p. 129. `doi.org/10.7551/mitpress/7503.003.0021`.

[9] Bradley CA Brown, Anthony L. Caterini, Brendan Leigh Ross, Jesse C Cresswell, and Gabriel Loaiza-Ganem. "Verifying the Union of Manifolds Hypothesis for Image Data". *The Eleventh International Conference on Learning Representations*. 2023. `openreview.net/forum?id=Rvee9CAX4fi`.

[10] Joan Bruna, Wojciech Zaremba, Arthur Szlam, and Yann LeCun. "Spectral networks and locally connected networks on graphs". *arXiv preprint arXiv:1312.6203* (2013). `arxiv.org/abs/1312.6203`.

[11] Leon Bungert, Jeff Calder, Max Mihailescu, Kodjo Houssou, and Amber Yuan. *Convergence rates for Poisson learning to a Poisson equation with measure data*. 2024. `arxiv.org/abs/2407.06783`.

[12] Leon Bungert, Jeff Calder, and Tim Roith. "Ratio convergence rates for Euclidean first-passage percolation: applications to the graph infinity Laplacian". *The Annals of Applied Probability* 34.4 (2024), pp. 3870–3910. `doi.org/10.1214/24-AAP2052`.

[13] Jeff Calder. "The game theoretic p-Laplacian and semi-supervised learning with few labels". *Nonlinearity* 32.1 (2018), p. 301. `doi.org/10.1088/1361-6544/aae949`.

[14] Jeff Calder and Nicolás García Trillos. "Improved spectral convergence rates for graph Laplacians on ε-graphs and k-NN graphs". *Appl. Comput. Harmon. Anal.* 60 (2022), pp. 123–175. `doi.org/10.1016/j.acha.2022.02.004`.

[15] Jeff Calder, Nicolás García Trillos, and Marta Lewicka. “Lipschitz regularity of graph Laplacians on random data clouds”. *SIAM J. Math. Anal.* 54.1 (2022), pp. 1169–1222. doi.org/10.1137/20M1356610.

[16] Jeff Calder, Dejan Slepčev, and Matthew Thorpe. “Rates of convergence for Laplacian semi-supervised learning with low labeling rates”. *Research in the Mathematical Sciences* 10.1 (2023), p. 10. doi.org/10.1007/s40687-022-00371-x.

[17] Marco Caroccia, Antonin Chambolle, and Dejan Slepčev. “Mumford-Shah functionals on graphs and their asymptotics”. *Nonlinearity* 33.8 (2020), pp. 3846–3888. doi.org/10.1088/1361-6544/ab81ee.

[18] Kevin M. Carter, Raviv Raich, and Alfred O. Hero III. “On Local Intrinsic Dimension Estimation and Its Applications”. *IEEE Transactions on Signal Processing* 58.2 (2010), pp. 650–663. doi.org/10.1109/tsp.2009.2031722.

[19] Jeff Cheeger and Tobias H Colding. “Lower bounds on Ricci curvature and the almost rigidity of warped products”. *Annals of Mathematics* 144.1 (1996), pp. 189–237. doi.org/10.2307/2118589.

[20] Haoyu Chen, Anna Little, and Akil Narayan. “Largest Angle Path Distance for Multi-Manifold Clustering”. *2023 International Conference on Sampling Theory and Applications (SampTA)*. IEEE. 2023, pp. 1–7. doi.org/10.1109/sampta59647.2023.10301401.

[21] Haoyu Chen, Anna Little, and Akin Narayan. *Robust Multi-Manifold Clustering via Simplex Paths*. 2025. arxiv.org/abs/2507.10710.

[22] Xiuyuan Cheng and Nan Wu. “Eigen-convergence of Gaussian kernelized graph Laplacian by manifold heat interpolation”. *Appl. Comput. Harmon. Anal.* 61 (2022), pp. 132–190. doi.org/10.1016/j.acha.2022.06.003.

[23] Ronald R. Coifman and Stéphane Lafon. “Diffusion maps”. *Appl. Comput. Harmon. Anal.* 21.1 (2006), pp. 5–30. doi.org/10.1016/j.acha.2006.04.006.

[24] Michaël Defferrard, Xavier Bresson, and Pierre Vandergheynst. “Convolutional neural networks on graphs with fast localized spectral filtering”. *Advances in Neural Information Processing Systems* 29 (2016). doi.org/10.48550/arXiv.1606.09375.

[25] Lawrence C. Evans. *Partial differential equations*. Second. Vol. 19. Graduate Studies in Mathematics. American Mathematical Society, Providence, RI, 2010, pp. xxii+749. doi.org/10.1090/gsm/019.

[26] Masatoshi Fukushima, Yoichi Oshima, and Masayoshi Takeda. *Dirichlet forms and symmetric Markov processes*. extended. Vol. 19. De Gruyter Studies in Mathematics. Walter de Gruyter & Co., Berlin, 2011, pp. x+489.

[27] Nicolás García Trillos. “Variational limits of k-NN graph-based functionals on data clouds”. *SIAM Journal on Mathematics of Data Sci.* 1.1 (2019), pp. 93–120. doi.org/10.1137/18m1188999.

[28] Nicolás García Trillos, Moritz Gerlach, Matthias Hein, and Dejan Slepčev. “Error estimates for spectral convergence of the graph Laplacian on random geometric graphs toward the Laplace–Beltrami operator”. *Foundations of Computational Mathematics* 20.4 (2020), pp. 827–887. doi.org/10.1007/s10208-019-09436-w.

[29] Nicolás García Trillos, Pengfei He, and Chenghui Li. “Large sample spectral analysis of graph-based multi-manifold clustering”. *Journal of Machine Learning Research* 24.143 (2023), pp. 1–71. jmlr.org/papers/v24/21-1254.html.

[30] Nicolás García Trillos, Chenghui Li, and Raghavendra Venkatraman. *Minimax Rates for the Estimation of Eigenpairs of Weighted Laplace-Beltrami Operators on Manifolds*. 2025. arxiv.org/abs/2506.00171.

[31] Nicolás García Trillos and Dejan Slepčev. "On the rate of convergence of empirical measures in ∞-transportation distance". *Canadian Journal of Mathematics* 67.6 (2015), pp. 1358–1383. doi.org/10.4153/CJM-2014-044-6.

[32] Nicolás García Trillos and Dejan Slepčev. "Continuum limit of total variation on point clouds". *Archive for Rational Mechanics and Analysis* 220 (2016), pp. 193–241. doi.org/10.1007/s00205-015-0929-z.

[33] Nicolás García Trillos and Dejan Slepčev. "A variational approach to the consistency of spectral clustering". *Appl. Comput. Harmon. Anal.* 45.2 (2018), pp. 239–281. doi.org/10.1016/j.acha.2016.09.003.

[34] Andrew Goldberg, Xiaojin Zhu, Aarti Singh, Zhiting Xu, and Robert Nowak. "Multi-Manifold Semi-Supervised Learning". *Proceedings of the Twelfth International Conference on Artificial Intelligence and Statistics*. Proceedings of Machine Learning Research. 2009, pp. 169–176. proceedings.mlr.press/v5/goldberg09a.html.

[35] Dian Gong, Xuemei Zhao, and Gérard Medioni. "Robust multiple manifolds structure learning". *Proceedings of the 29th International Coference on International Conference on Machine Learning*. 2012, pp. 25–32. icml.cc/2012/papers/191.pdf.

[36] Matthias Hein, Jean-Yves Audibert, and Ulrike von Luxburg. "Graph laplacians and their convergence on random neighborhood graphs". *Journal of Machine Learning Research* 8.6 (2007). www.jmlr.org/papers/volume8/hein07a/hein07a.pdf.

[37] Nicholas J. Korevaar and Richard M. Schoen. "Sobolev spaces and harmonic maps for metric space targets". *Comm. Anal. Geom.* 1.3-4 (1993), pp. 561–659. doi.org/10.4310/CAG.1993.v1.n4.a4.

[38] Takashi Kumagai and Karl-Theodor Sturm. "Construction of diffusion processes on fractals, d-sets, and general metric measure spaces". *J. Math. Kyoto Univ.* 45.2 (2005), pp. 307–327. doi.org/10.1215/kjm/1250281992.

[39] Tim Laux and Jona Lelmi. "Large data limit of the MBO scheme for data clustering: convergence of the dynamics". *Journal of Machine Learning Research* 24.344 (2023), pp. 1–49. jmlr.org/papers/v24/22-1089.html.

[40] Tim Laux and Jona Lelmi. "Large data limit of the MBO scheme for data clustering: Γ-convergence of the thresholding energies". *Applied and Computational Harmonic Analysis* (2025), p. 101800. doi.org/10.1016/j.acha.2025.101800.

[41] Ulrike von Luxburg. "A tutorial on spectral clustering". *Statistics and Computing* 17.4 (2007), pp. 395–416. doi.org/10.1007/s11222-007-9033-z.

[42] Ulrike von Luxburg, Mikhail Belkin, and Olivier Bousquet. "Consistency of spectral clustering". *Ann. Statist.* 36.2 (2008), pp. 555–586. .

[43] F Patricia Medina, Linda Ness, Melanie Weber, and Karamatou Yacoubou Djima. "Heuristic framework for multiscale testing of the multi-manifold hypothesis". *Research in Data Science*. Springer, 2019, pp. 47–80. link.springer.com/chapter/10.1007/978-3-030-11566-1_3.

[44] Andrew Ng, Michael Jordan, and Yair Weiss. "On spectral clustering: Analysis and an algorithm". *Advances in Neural Information Processing Systems* 14 (2001). papers.nips.cc/paper_files/paper/2001/file/801272ee79cfde7fa5960571fee36b9b-Paper.pdf.

[45] Amit Singer. "From graph to manifold Laplacian: The convergence rate". *Applied and Computational Harmonic Analysis* 21.1 (2006), pp. 128–134. doi.org/10.1016/j.acha.2006.03.004.

[46] Amit Singer and Hau-Tieng Wu. "Spectral convergence of the connection Laplacian from random samples". *Information and Inference: A Journal of the IMA* 6.1 (2017), pp. 58–123. doi.org/10.1093/imaiai/iaw016.

[47] Dejan Slepčev and Matthew Thorpe. "Analysis of p-Laplacian regularization in semisupervised learning". *SIAM Journal on Mathematical Analysis* 51.3 (2019), pp. 2085–2120. doi.org/10.1137/17M115222X.

[48] Karl-Theodor Sturm. "Diffusion processes and heat kernels on metric spaces". *Ann. Probab.* 26.1 (1998), pp. 1–55. doi.org/10.1214/aop/1022855410.

[49] Karl-Theodor Sturm. "How to construct diffusion processes on metric spaces". *Potential Anal.* 8.2 (1998), pp. 149–161. doi.org/10.1023/A:1008667129215.

[50] Anton Ullrich. *The heat flow on glued spaces with varying dimension*. 2024. arxiv.org/abs/2406.09996.

[51] Cédric Villani. *Optimal transport: old and new*. Vol. 338. Springer, 2008.

[52] Xu Wang, Konstantinos Slavakis, and Gilad Lerman. "Multi-manifold modeling in non-Euclidean spaces". *Artificial Intelligence and Statistics*. PMLR. 2015, pp. 1023–1032. proceedings.mlr.press/v38/wang15b.html.

[53] Yong Wang, Yuan Jiang, Yi Wu, and Zhi-Hua Zhou. "Multi-manifold clustering". *Pacific Rim International Conference on Artificial Intelligence*. Springer. 2010, pp. 280–291. link.springer.com/chapter/10.1007/978-3-642-15246-7_27.

[54] Yong Wang, Yuan Jiang, Yi Wu, and Zhi-Hua Zhou. "Spectral clustering on multiple manifolds". *IEEE Transactions on Neural Networks* 22.7 (2011), pp. 1149–1161. doi.org/10.1109/tnn.2011.2147798.

[55] Caroline L. Wormell and Sebastian Reich. "Spectral convergence of diffusion maps: improved error bounds and an alternative normalization". *SIAM J. Numer. Anal.* 59.3 (2021), pp. 1687–1734. doi.org/10.1137/20M1344093.

[56] Hau-Tieng Wu and Nan Wu. "Strong uniform consistency with rates for kernel density estimators with general kernels on manifolds". *Information and Inference: A Journal of the IMA* 11.2 (2022), pp. 781–799. doi.org/10.1093/imaiai/iaab014.

[57] Jie Zhou, Ganqu Cui, Shengding Hu, Zhengyan Zhang, Cheng Yang, Zhiyuan Liu, Lifeng Wang, Changcheng Li, and Maosong Sun. "Graph neural networks: A review of methods and applications". *AI Open* 1 (2020), pp. 57–81. www.sciencedirect.com/science/article/pii/S2666651021000012.

A Appendix

A.1 The principal angle

The positivity of the principal angle in (2.1) implies the following lemma which we need for some of our proofs.

Lemma A.1. *If $\theta > 0$, where θ is defined in* (2.1)*, then there exist constants $C, \varepsilon_0 > 0$ such that for all $0 < \varepsilon < \varepsilon_0$ whenever $x \in \mathcal{M}^{(i)}$ with $\operatorname{dist}(x, \mathcal{M}^{(12)}) > C\varepsilon$, then $B(x, \varepsilon) \cap \mathcal{M}^{(j)} = \emptyset$ for $i \neq j$.*

Proof. Take $C = 2/\sin\theta$ and consider ε_0 to be smaller than the injectivity radii of both manifolds. If the claim does not hold, then there exist sequences $x_n \in \mathcal{M}^{(1)}$ and $y_n \in \mathcal{M}^{(2)}$ such that $\operatorname{dist}(x_n, \mathcal{M}^{(12)}) \to 0$ and $|x_n - y_n| < \frac{\sin\theta}{2} \operatorname{dist}(x_n, \mathcal{M}^{(12)})$. Let $z_n \in \mathcal{M}^{(12)}$ be the closest point to x_n with respect to the Euclidean distance, and let $\log^{(i)}_{z_n}$ be the inverse of the exponential map from the ball of radius ε_0 in $T_{\mathcal{M}^{(i)}}$ to $\mathcal{M}^{(i)}$ for $i = 1, 2$.

First, we note that combining the law of sines with our assumption that $|x_n - y_n| \leq \frac{\sin\theta}{2} |x_n - z_n|$, the angle $\beta_n := \angle(x_n - z_n, y_n - z_n)$ satisfies:

$$2R_n \sin\beta_n = |x_n - y_n| \leq |x_n - z_n| \frac{\sin\theta}{2} \leq R_n \sin\theta,$$

where R_n is the radius of the circumcircle of the triangle with vertices x_n, y_n, and z_n. Hence, we obtain $\sin\beta_n \leq \frac{\sin\theta}{2}$ which implies $\beta_n \leq \frac{\theta}{2}$. Using this together with the definition of θ in (2.1) and a simple triangle inequality for vectors we have

$$\begin{aligned}
\theta &\leq \angle\left(\log^{(1)}_{z_n}(x_n), \log^{(2)}_{z_n}(y_n)\right) \\
&\leq \angle\left(\log^{(1)}_{z_n}(x_n), x_n - z_n\right) + \angle(x_n - z_n, y_n - z_n) + \angle\left(\log^{(2)}_{z_n}(x_n), y_n - z_n\right) \\
&\leq \angle\left(\log^{(1)}_{z_n}(x_n), x_n - z_n\right) + \frac{\theta}{2} + \angle\left(\log^{(2)}_{z_n}(x_n), y_n - z_n\right).
\end{aligned}$$

Since for large n we have $\angle\left(\log^{(1)}_{z_n}(x_n), x_n - z_n\right) < \frac{\theta}{8}$ and $\angle\left(\log^{(2)}_{z_n}(y_n), y_n - z_n\right) < \frac{\theta}{8}$ we arrive at a contradiction. □

A.2 The TL^p convergence

To have a unified domain for functionals defined on graphs and continuum manifolds, we rely on the TL^p-topology, introduced in [32]. There the TL^p-spaces for $p \in [1, \infty)$ where defined as spaces of pairs of probability measures and L^p-functions on a metric space M:

$$TL^p(M) := \{(u, \mu) : \mu \in \mathcal{P}(M),\ u \in L^p(M; \mu)\}.$$

These spaces can be equipped with an optimal transport type metric by setting

$$d_{TL^p}((u,\mu),(v,\nu)) := \inf_{\pi \in \Gamma(\mu,\nu)} \left(\iint_{M \times M} \left(d(x,y)^p + |u(x) - v(y)|\right)^p \mathrm{d}\pi(x,y) \right)^{\frac{1}{p}},$$

where the infimum is taken over couplings $\pi \in \mathcal{P}(M \times M)$ of the two probability measures μ and ν. In [32] these definitions were used for M being an open subset of Euclidean space. In the case where $M = \mathcal{M}$ is a compact Riemannian manifold without boundary, equipped with a probability measure μ that is absolutely continuous with respect to the volume measure on $\mathcal{M}$, we refer to [28, 40] for definitions and properties of TL^p-distances on such spaces. In our situation, where $\mathcal{M} = \mathcal{M}^{(1)} \cup \mathcal{M}^{(2)}$ is a union of manifolds embedded in $\mathbb{R}^N$ as defined in Section 2.1, we note that $M = \mathcal{M}$ is a metric space when equipped with the ambient Euclidean metric of $\mathbb{R}^N$.

It turns out that one can characterize the convergence $(u_n, \mu_n) \to (u, \mu)$ in $TL^p(M)$ by means of so-called stagnating transport plans which are couplings $\pi_n \in \Pi(\mu, \mu_n) := \{\pi \in \mathcal{P}(M \times M) : \pi(\cdot \times M) = \mu,\ \pi(M \times \cdot) = \mu_n\}$ of μ and μ_n that satisfy

$$\lim_{n\to\infty} \iint_{M \times M} d(x,y)\, \mathrm{d}\pi_n(x,y) = 0.$$

A special case of stagnating transport plans are of the form $\pi_n = (\mathrm{id} \times T_n)_\sharp \mu$ where $T_n : M \to M$ is a stagnating transport map, meaning that it satisfies $(T_n)_\sharp \mu = \mu_n$ and

$$\lim_{n\to\infty} \int_M d(x, T_n(x)) \,\mathrm{d}\mu(x) = 0.$$

Note, however, that such maps T_n do not necessarily exist. We have the following characterization of TL^p-convergence:

Proposition A.2 ([32, Proposition 3.12])**.** *Let M be a Polish space, $\mu \in \mathcal{P}(M)$ be a Borel probability measure, and $p \in [1, \infty)$. Then the following statements are equivalent:*

(i) $(u_n, \mu_n) \to (u, \mu)$ *in* $TL^p(M)$ *as* $n \to \infty$*;*

(ii) $\mu_n \overset{*}{\rightharpoonup} \mu$ *weakly as measures and for every stagnating sequence of transport plans* $\pi_n \in \Pi(\mu, \mu_n)$ *it holds*

$$\lim_{n\to\infty} \iint_{M\times M} |u(x) - u_n(y)|^p \,\mathrm{d}\pi_n(x, y) = 0; \tag{A.1}$$

(iii) $\mu_n \overset{*}{\rightharpoonup} \mu$ *weakly as measures and there exists a stagnating sequence of transport plans* $\pi_n \in \Pi(\mu, \mu_n)$ *such that* (A.1) *holds.*

Moreover, if $M = \mathcal{M}$ is a union of Riemannian manifolds without boundary with volume form Vol *and $\mu \in \mathcal{P}(\mathcal{M})$ is absolutely continuous with respect to* Vol*, then 1.-3. are also equivalent to*

(iv) $\mu_n \overset{*}{\rightharpoonup} \mu$ *weakly as measures and for every stagnating sequence of transport maps* $T_n : \mathcal{M} \to \mathcal{M}$ *(with* $(T_n)_\sharp \mu = \mu_n$*) it holds*

$$\lim_{n\to\infty} \int_{\mathcal{M}} |u(x) - u_n(T_n(x))|^p \,\mathrm{d}\mu(x) = 0; \tag{A.2}$$

(v) $\mu_n \overset{*}{\rightharpoonup} \mu$ *weakly as measures and there exists a stagnating sequence of transport maps* $T_n : \mathcal{M} \to \mathcal{M}$ *(with* $(T_n)_\sharp \mu = \mu_n$*) such that* (A.2) *holds.*

Remark A.1. Note that (i)-(iii) in Theorem A.2 are, in particular, applicable to the union of manifolds $\mathcal{M}$ equipped with the probability measure $\mu \in \mathcal{P}(\mathcal{M})$.

Remark A.2. Note that [32, Proposition 3.12] is stated in a less general form where M is an open and bounded domain in $\mathbb{R}^d$. However, an inspection of the proof shows that Theorem A.2 holds. Indeed, the equivalence of (ii) and (iii) follows from the assumption that μ is a Borel probability measure and hence regular, which implies that bounded Lipschitz functions are dense in $L^1(M; \mu)$, cf. [32, Lemma 3.10]. Furthermore, the equivalence of (i) and (iii) is proved as in [32, Proposition 3.12], using that convergence in the Wasserstein-p distance for $p \in [1, \infty)$ metrizes weak convergence of measures in a Polish space [51, Theorem 6.9].

Finally, we use the same abuse of notation which was introduced in [32, Definition 3.13]. It accounts for the fact that in graph-based learning the measure part of a sequence in TL^p is typically prescribed and therefore less important to deal with than the function part.

Definition A.1 (TL^p-convergence and compactness of functions)**.** Let $\mu_n \overset{*}{\rightharpoonup} \mu$ weakly as measures as $n \to \infty$ and let $(u_n)_{n\in\mathbb{N}}$ with $u_n \in L^p(M, \mu_n)$ be a sequence. We say that u_n converges to u in $TL^p(M)$ if and only if $(u_n, \mu_n) \to (u, \mu)$ in $TL^p(M)$. Similarly, we say that $(u_n)_{n\in\mathbb{N}}$ is relatively compact in $TL^p(M)$ if $(u_n, \mu_n)_{n\in\mathbb{N}}$ is relatively compact in $TL^p(M)$.

A.3 Γ-convergence of unnormalized Dirichlet energies

Theorem A.3 ([40, Theorem 11])**.** *Let $\mathcal{M}$ be a d-dimensional compact Riemannian manifold embedded in $\mathbb{R}^N$, let $\eta : [0,\infty) \to [0,\infty)$ satisfy Assumption 2, let $\rho \in C(\mathcal{M})$ be a continuous function satisfying $\frac{1}{C} \leq \rho \leq C$ on $\mathcal{M}$ with some $C > 0$, and let $(\varepsilon_n)_{n\in\mathbb{N}} \subset (0,\infty)$ be a sequence converging to zero. Then as $n \to \infty$ the functionals*

$$\begin{aligned}
&\mathcal{G}_{\varepsilon_n} : L^2(\mathcal{M}) \to [0,\infty],\\
&u \mapsto \frac{1}{\varepsilon_n^{d+2}} \int_{\mathcal{M}} \int_{\mathcal{M}} \eta_{\varepsilon_n}(|x-y|)\, |u(x) - u(y)|^2\, \rho(x)\rho(y)\, \mathrm{dVol}(x)\, \mathrm{dVol}(y),
\end{aligned}$$

Γ-converge in $L^2(\mathcal{M})$ to

$$\begin{aligned}
&\mathcal{G} : L^2(\mathcal{M}) \to [0,\infty],\\
&u \mapsto \begin{cases} \sigma_\eta \int_{\mathcal{M}} |\nabla_{\mathcal{M}} u|^2\, \rho^2\, \mathrm{dVol} & \text{if } u \in H^1(\mathcal{M}),\\ \infty & \text{else}, \end{cases}
\end{aligned}$$

where $\sigma_\eta := \int_{\mathbb{R}^d} \eta(|x|)\, |x_1|^2\, \mathrm{d}x$. Furthermore, any sequence $u_n \subset L^2(\mathcal{M})$ such that

$$\sup_{n\in\mathbb{N}} \mathcal{G}_{\varepsilon_n}(u_n) < \infty, \qquad \sup_{n\in\mathbb{N}} \|u_n\|_{L^2(\mathcal{M})} < \infty,$$

is precompact in $L^2(\mathcal{M})$.

Theorem A.4. *Let $\mathcal{M}$ be a d-dimensional compact Riemannian manifold embedded in a Euclidean space. Let $\eta : [0,\infty) \to [0,\infty)$ satisfy Assumption 2, and let $\rho \in C(\mathcal{M})$ be a continuous function satisfying $\frac{1}{C} \leq \rho \leq C$ on $\mathcal{M}$ with some $C > 0$. Letting $\{x_1, \ldots, x_n\} \subset \mathcal{M}$ for $n \in \mathbb{N}$ be an i.i.d. sample of the measure $\mathrm{d}\mu := \rho\, \mathrm{dVol}$, define $\mu_n := \frac{1}{n}\sum_{i=1}^n \delta_{x_i}$ and let ε_n satisfy Assumption 3. Then as $n \to \infty$ almost surely the functionals*

$$\begin{aligned}
&G_{n,\varepsilon_n} : TL^2(\mathcal{M}_n; \mu_n) \to [0,\infty],\\
&u \mapsto \frac{1}{\varepsilon_n^{d+2}} \int_{\mathcal{M}} \int_{\mathcal{M}} \eta_{\varepsilon_n}(|x-y|)\, |u(x) - u(y)|^2\, \mathrm{d}\mu_n(x)\, \mathrm{d}\mu_n(y),
\end{aligned}$$

Γ-converge in $TL^2(\mathcal{M})$ to

$$\begin{aligned}
&\mathcal{G} : L^2(\mathcal{M}) \to [0,\infty],\\
&u \mapsto \begin{cases} \sigma_\eta \int_{\mathcal{M}} |\nabla_{\mathcal{M}} u|^2\, \rho^2\, \mathrm{dVol} & \text{if } u \in H^1(\mathcal{M}),\\ \infty & \text{else}, \end{cases}
\end{aligned}$$

where $\sigma_\eta := \int_{\mathbb{R}^d} \eta(|x|)\, |x_1|^2\, \mathrm{d}x$. Furthermore, any sequence $(u_n, \mu_n) \subset TL^2(\mathcal{M})$ such that

$$\sup_{n\in\mathbb{N}} G_{n,\varepsilon_n}(u_n) < \infty, \qquad \sup_{n\in\mathbb{N}} \|u_n\|_{L^2(\mu_n)} < \infty,$$

is precompact in $TL^2(\mathcal{M})$.

Proof. The theorem is essentially the one stated in [40, Theorem 9], where it was assumed that ρ is smooth and $d \geq 2$, although the proof allows for more general ρ and for $d = 1$. The proof relies on the corresponding Γ-convergence results in the Euclidean setting, see [33] for $d \geq 2$ and [47] for $d = 1$, as well as the argument of [14, Proposition 2.12] (or [17, Lemmas 3.1, 3.2]) to get the sharp dependence of ε_n on n for $d = 2$. □

We note that the theorem can be easily extended to the situation that the number of points n, is not deterministic, but is a random variable with controlled variance. We state and prove the corollary for n which has a binomial distribution as this is the situation we face when dealing with samples from union of two manifolds.

Corollary A.5. *In the setting of Theorem A.4, let $N \in \mathbb{N}$, $p \in (0,1)$, $n \sim \mathrm{Bin}(N,p)$, and $\{x_1, \ldots, x_n\} \subset \mathcal{M}$ be an i.i.d. sample of the measure $\mathrm{d}\mu := \rho\,\mathrm{dVol}$. If Assumption 3 holds, then as $N \to \infty$ the assertions of Theorem A.4 remain true for the functional G_{n,ε_N}.*

Proof. A simple application of Hoeffding's inequality to the binomial random variable n shows

$$\mathbb{P}\left[|n - Np| \geq t\right] \leq 2\exp\left(-\frac{2t^2}{N}\right)$$

and taking $t := Np/2$ we obtain

$$\mathbb{P}\left[n \geq \frac{Np}{2}\right] \geq 1 - 2\exp\left(-\frac{Np^2}{2}\right).$$

At the same time, by [28, Theorem 2] there exist constants $C_1, C_2 > 0$ such that for $m \in \mathbb{N}$ with high probability there exists a transport map T_m pushing μ onto μ_m with infinity transport distance at most $C_1\delta_m$

$$\mathbb{P}\left[\|\mathrm{id} - T_m\|_{L^\infty(\mathcal{M})} > C_1\delta_m\right] \leq \frac{C_2}{m^2},$$

where

$$\delta_n := \begin{cases} \sqrt{\dfrac{\log\log n}{n}} & \text{if } d = 1, \\ \dfrac{(\log n)^{3/4}}{n^{1/2}} & \text{if } d = 2, \\ \left(\dfrac{\log n}{n}\right)^{\frac{1}{d}} & \text{if } d \geq 2. \end{cases}$$

With a union bound we obtain for any $m \in \mathbb{N}$:

$$\mathbb{P}\left[n \geq \frac{Np}{2} \text{ and } \|\mathrm{id} - T_m\|_{L^\infty(\mathcal{M})} \leq C_1\delta_m\right] \geq 1 - \frac{C_2}{m^2} - 2\exp\left(-\frac{Np^2}{2}\right).$$

Now we can choose m as the smallest integer larger or equal to $\frac{Np}{2}$ (implying that $n \geq m$ in the above probability) and use the law of total probability to obtain

$$\begin{aligned} \mathbb{P}\left[\|\mathrm{id} - T_n\|_{L^\infty(\mathcal{M})} \leq C_1\delta_{Np/2}\right] &\geq \mathbb{P}\left[\|\mathrm{id} - T_n\|_{L^\infty(\mathcal{M})} \leq C_1\delta_{Np/2} \,\middle|\, n \geq \frac{Np}{2}\right]\mathbb{P}\left[n \geq \frac{Np}{2}\right] \\ &= \mathbb{P}\left[n \geq \frac{Np}{2} \text{ and } \|\mathrm{id} - T_n\|_{L^\infty(\mathcal{M})} \leq C_1\delta_{Np/2}\right] \\ &\geq 1 - \frac{4C_2}{p^2N^2} - 2\exp\left(-\frac{Np^2}{2}\right). \end{aligned}$$

Combining this estimate with the Borel–Cantelli lemma and using that $\limsup_{n\to\infty} \frac{\delta_N}{\delta_{Np/2}} \in (0,\infty)$ implies that almost surely we have

$$\limsup_{n\to\infty} \frac{\|\mathrm{id} - T_n\|_{L^\infty(\mathcal{M})}}{\delta_N} < \infty.$$

Since this is the only estimate on the distribution of points used in the proof of Theorem A.4 and since $\left(\frac{\log N}{N}\right)^{\frac{1}{d}} \ll \varepsilon_N \ll 1$ we assume is equivalent to $\left(\frac{\log(pN/2)}{pN/2}\right)^{\frac{1}{d}} \ll \varepsilon_N \ll 1$, we conclude that the corollary holds. $\square$

A.4 Establishing the trace condition

The challenging part in proving the liminf inequality is to establish the trace condition encoded in (2.9) and (2.10) in the case that the manifolds have equal dimension and the codimension of their intersection equals one. Towards that goal, we first prove the following auxiliary lemma.

Lemma A.6. *In the multi-manifold setting of Section 2.1, in particular, under Assumptions 1 and 2 let $d^{(1)} = d^{(2)} =: d$ and $d^{(12)} = d - 1$. Assume that $\varepsilon_n \to 0$ and $v_n \to v$ in $L^2(\mathcal{M})$ as $n \to \infty$ where $v|_{\mathcal{M}^{(i)}} \in H^1(\mathcal{M}^{(i)})$, and that*

$$\liminf_{n\to\infty} \frac{1}{\varepsilon_n^{d+2}} \int_{\mathcal{M}} \int_{\mathcal{M}} \eta_{\varepsilon_n}(|x-y|)\, |v_n(x) - v_n(y)|^2 \,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y) < \infty.$$

Then $\mathrm{Tr}^{(1)}(v) = \mathrm{Tr}^{(2)}(v)$ *on* $\mathcal{M}^{(12)}$.

Proof. Let us fix a point $z \in \mathcal{M}^{(12)}$ and a radius $r > 0$. Since $\mathcal{M}^{(12)}$ has codimension 1 in $\mathcal{M}^{(1)}$ and $\mathcal{M}^{(2)}$, for r small, $\mathcal{M}^{(12)}$ divides them locally into two pieces. Namely let $\mathcal{M}_{\pm}^{(i)}$ for $i = 1, 2$ be connected components of $B(z, r) \cap \mathcal{M}^{(i)} \setminus \mathcal{M}^{(12)}$. We define $U := B(z, r) \cap (\mathcal{M}_+^{(1)} \cup \mathcal{M}_+^{(2)} \cup \mathcal{M}^{(12)})$. Since the intersection of the two manifolds is assumed nondegenerate and they have equal dimension, there exists open set $V \subset \mathbb{R}^d$ and a bi-Lipschitz map $\Psi : U \to V$, i.e., there exists $L > 0$ such that

$$\frac{1}{L} |\Psi(x) - \Psi(y)| \le |x - y| \le L\, |\Psi(x) - \Psi(y)| \quad \forall x, y \in U.$$

Without loss of generality we can assume that Ψ is such that

$$\Psi(U \cap \mathcal{M}_+^{(1)}) \subset \{x \in \mathbb{R}^d \,:\, x_1 \ge 0\}, \qquad \Psi(U \cap \mathcal{M}_+^{(2)}) \subset \{x \in \mathbb{R}^d \,:\, x_1 \le 0\}$$

and hence

$$\Psi(B(z, r) \cap \mathcal{M}^{(12)}) \subset \{x \in \mathbb{R}^d \,:\, x_1 = 0\}.$$

Let $w_n := v_n \circ \Psi^{-1}$. Using the regularity of Ψ we obtain

$$\begin{aligned}
&\frac{1}{\varepsilon_n^{d+2}} \int_{\mathcal{M}} \int_{\mathcal{M}} \eta_{\varepsilon_n}(|x-y|)\, |v_n(x) - v_n(y)|^2 \,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y) \\
&\qquad \ge \frac{1}{\varepsilon_n^{d+2}} \int_U \int_U \eta_{\varepsilon_n}(|x-y|)\, |v_n(x) - v_n(y)|^2 \,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y) \\
&\qquad = \frac{1}{\varepsilon_n^{d+2}} \int_V \int_V \eta_{\varepsilon_n}(\left|\Psi^{-1}(\tilde{x}) - \Psi^{-1}(\tilde{y})\right|)\, |w_n(\tilde{x}) - w_n(\tilde{y})|^2 \,\mathrm{d}\Phi_\sharp\mu(\tilde{x})\,\mathrm{d}\Phi_\sharp\mu(\tilde{y}) \\
&\qquad \ge \frac{1}{\varepsilon_n^{d+2}} \int_V \int_V \eta_{\varepsilon_n}(L\, |\tilde{x} - \tilde{y}|)\, |w_n(\tilde{x}) - w_n(\tilde{y})|^2 \,\mathrm{d}\Phi_\sharp\mu(\tilde{x})\,\mathrm{d}\Phi_\sharp\mu(\tilde{y}) \\
&\qquad \ge \frac{1}{L^{d+2}} \frac{1}{\tilde{\varepsilon}_n^{d+2}} \int_V \int_V \eta_{\tilde{\varepsilon}_n}(|\tilde{x} - \tilde{y}|)\, |w_n(\tilde{x}) - w_n(\tilde{y})|^2 \,\mathrm{d}\Phi_\sharp\mu(\tilde{x})\,\mathrm{d}\Phi_\sharp\mu(\tilde{y}) \\
&\qquad \ge \frac{c}{L^{d+2}} \frac{1}{\tilde{\varepsilon}_n^{d+2}} \int_V \int_V \eta_{\tilde{\varepsilon}_n}(|\tilde{x} - \tilde{y}|)\, |w_n(\tilde{x}) - w_n(\tilde{y})|^2 \,\mathrm{d}\tilde{x}\,\mathrm{d}\tilde{y}
\end{aligned}$$

where $\tilde{\varepsilon}_n := \varepsilon_n/L$ and $c > 0$ depends on L and the lower bound of ρ (the density of μ). Using that w_n (being a composition of a Lipschitz function and a convergent sequence of functions v_n) converges to $w := v \circ \Psi^{-1} \in L^2(V)$ and the $\liminf$ inequality from Theorem A.4 inequality in the Euclidean setting we get

$$\int_V |\nabla w|^2 \, \mathrm{d}x < \infty.$$

Therefore, the trace of w on the $d-1$-dimensional set $\mathcal{S} := V \cap \{x \in \mathbb{R}^d \,:\, x_1 = 0\}$ is well-defined. Thus, the trace of v on $\Psi^{-1}(\mathcal{S})$ is well defined and hence

$$\mathrm{Tr}^{(1)}(v|_U) = \mathrm{Tr}^{(2)}(v|_U).$$

□

A.5 Smooth approximation with trace constraint

We show that smooth functions on a union of manifolds of equal dimension are dense in the space of H^1-functions which have identical traces on the intersection $\mathcal{M}^{(12)}$ in case the latter has codimension one. This is nontrivial since smoothing the functions separately on each of the manifolds is not guaranteed to lead to traces values which agree. We were not able to find such result in the literature and think it is of independent interest.

Lemma A.7. *Let $d^{(1)} = d^{(2)} =: d$, assume that $d^{(12)} = d - 1$, and define*

$$\mathcal{S} := \left\{ u \in L^2(\mathcal{M}) \,:\, u|_{\mathcal{M}^{(i)}} \in C^\infty(\mathcal{M}^{(i)}) \text{ for } i = 1, 2 \text{ and } \mathrm{Tr}^{(1)}(u) = \mathrm{Tr}^{(2)}(u) \right\}.$$

Then $\mathcal{S}$ is dense in

$$H^1(\mathcal{M}) := \left\{ u \in L^2(\mathcal{M}) \,:\, u|_{\mathcal{M}^{(i)}} \in H^1(\mathcal{M}^{(i)}) \text{ for } i \in \{1, 2\} \text{ and } \mathrm{Tr}^{(1)}(u) = \mathrm{Tr}^{(2)}(u) \right\}.$$

Proof. We let $u \in H^1(\mathcal{M})$ and abbreviate $u^{(i)} := u|_{\mathcal{M}^{(i)}}$ for $i \in \{1, 2\}$.

Step 1, local tubular neighborhood theorem. For any point $x \in \mathcal{M}^{(12)}$ let us consider an open neighborhood $V \subset \mathbb{R}^N$ of x in $\mathbb{R}^N$. Let us furthermore consider an open neighborhood O of x in $\mathcal{M}^{(12)}$, compactly contained in V. We choose the neighborhood O sufficiently small such that we can apply tubular neighborhood theorem to obtain $\varepsilon > 0$ and open neighborhoods $U^{(i)}$ of x in $\mathcal{M}^{(i)}$ as well as diffeomorphisms $\Psi^{(i)} : U^{(i)} \to O \times (-\varepsilon, \varepsilon)$ with $\Psi^{(i)}(p, 0) = p$ for all $p \in O$. Consequently, we can define the diffeomorphism $\Psi := (\Psi^{(1)})^{-1} \circ \Psi^{(2)} : U^{(2)} \to U^{(1)}$. Note that by construction the intersection $\mathcal{M}^{(12)}$ divides $U^{(i)}$ into two parts, $U^{(i)} \setminus \mathcal{M}^{(12)} = U^{(i)}_+ \cup U^{(i)}_-$ where $U^{(i)}_+ := (\Psi^{(i)})^{-1}(O \times (0, \varepsilon))$ and $U^{(i)}_- := (\Psi^{(i)})^{-1}(O \times (-\varepsilon, 0))$. Let $W \subset \mathbb{R}^N$ be an open neighborhood of x such that $W \cap \mathcal{M}^{(i)} \subset U^{(i)}$ for $i = 1, 2$.

Step 2, partition of unity. By compactness of $\mathcal{M}^{(12)}$ we can cover $\mathcal{M}^{(12)}$ by finitely many such open sets $(O_k)_{k \in K}$ corresponding to points on $\mathcal{M}^{(12)}$, as constructed above. By $(U^{(i)}_k)_{k \in K}$, $(W_k)_{k \in K} \subset \mathbb{R}^N$, and $(\Psi^{(i)}_k)_{k \in K}$ for $i \in \{1, 2\}$ we denote the associated open covers and diffeomorphisms constructed in Step 1. We let $(\xi_k)_{k \in K \cup \{\infty\}} \subset C^\infty(\mathbb{R}^N)$ be a smooth partition of unity subordinate to the open cover $(W_k)_{k \in K \cup \{\infty\}}$ of $\mathbb{R}^N$ where we let $W_\infty := \mathbb{R}^N \setminus \mathcal{M}^{(12)}$ with $\operatorname{supp} \xi_k \subset W_k$ for all $k \in K \cup \{\infty\}$.

Step 3, local reduction to zero trace. Using the partition of unity we can now write $u^{(i)} = \sum_{k\in K} \xi_k u + \xi_\infty u$ and note that $\xi_\infty u \equiv 0$ in a neighborhood of the intersection $\mathcal{M}^{(12)}$. Hence, the last term can be easily smoothed without changing its trace values. So the key step is to smooth each of the functions $\xi_k u$, which have compact support in $U_k^{(i)}$.

We now fix the index $k \in K$ and omit it in our notation of the open sets $U^{(i)}$ and the diffeomorphism $\Psi : U^{(2)} \to U^{(1)}$. We also use the abbreviation $\bar{u}^{(i)} := \xi_k u^{(i)}$. We define the functions $\hat{u}^{(2)}, u_0 : U^{(1)} \to \mathbb{R}$ via $\hat{u}^{(2)} := \bar{u}^{(2)} \circ \Psi^{-1}$ and $u_0 := \bar{u}^{(1)} - \hat{u}^{(2)}$ and note that

$$\mathrm{Tr}^{(1)}(u_0) = \mathrm{Tr}^{(1)}(\bar{u}^{(1)}) - \mathrm{Tr}^{(1)}(\bar{u}^{(2)} \circ \Psi^{-1}) = \mathrm{Tr}^{(1)}(\bar{u}^{(1)}) - \mathrm{Tr}^{(2)}(\bar{u}^{(2)}) = 0,$$

since Ψ equals the identity map on $\mathcal{M}^{(12)}$ and $\mathrm{Tr}^{(i)}(\bar{u}^{(i)}) = \mathrm{Tr}^{(i)}(\xi_k u^{(i)})$ is independent of $i \in \{1, 2\}$. We also note that, by its definition, u_0 has compact support in $U^{(1)}$ We can find a smooth approximation $v^{(1)}$ of $\bar{u}^{(1)}$ with compact support in $U^{(1)}$, not caring about the trace values. Provided we manage to find a smooth approximation $v_0 \in C^\infty(\mathcal{M}^{(1)})$ of u_0 with $\mathrm{Tr}^{(1)}(v_0) = 0$, we claim that the function $v^{(2)} := (v^{(1)} - v_0) \circ \Psi : U^{(2)} \to \mathbb{R}$ is a smooth approximation of $\bar{u}^{(2)}$ with the same trace as $v^{(1)}$. Indeed, since Ψ is the identity on $\mathcal{M}^{(12)}$ and $\mathrm{Tr}^{(1)}(v_0) = 0$ it holds $\mathrm{Tr}^{(2)}(v^{(2)}) = \mathrm{Tr}^{(1)}(v^{(1)})$. Furthermore, we have

$$\begin{aligned}
\left\| v^{(2)} - \bar{u}^{(2)} \right\|_{H^1(U^{(2)})} &= \left\| (v^{(1)} - v_0) \circ \Psi - \bar{u}^{(2)} \right\|_{H^1(U^{(2)})} \\
&= \left\| (v^{(1)} - \bar{u}^{(1)} + \bar{u}^{(1)} - u_0 + u_0 - v_0) \circ \Psi - \bar{u}^{(2)} \right\|_{H^1(U^{(2)})} \\
&= \left\| (v^{(1)} - \bar{u}^{(1)} + \hat{u}^{(2)} + u_0 - v_0) \circ \Psi - \bar{u}^{(2)} \right\|_{H^1(U^{(2)})} \\
&= \left\| (v^{(1)} - \bar{u}^{(1)} + u_0 - v_0) \circ \Psi \right\|_{H^1(U^{(2)})} \\
&\le \left\| (v^{(1)} - \bar{u}^{(1)}) \circ \Psi \right\|_{H^1(U^{(2)})} + \| (u_0 - v_0) \circ \Psi \|_{H^1(U^{(2)})} \\
&\lesssim \left\| \nabla \Psi^{-1} \right\|_{L^\infty(U^{(1)})} \left[\left\| v^{(1)} - \bar{u}^{(1)} \right\|_{H^1(U^{(1)})} + \| u_0 - v_0 \|_{H^1(U^{(1)})} \right]
\end{aligned}$$

using that Ψ is a diffeomorphism. Hence, $v^{(2)}$ is a smooth approximation of $\bar{u}^{(2)}$ as claimed.

Step 4, smoothing with zero trace. To conclude the proof we have to prove that we can approximate the function $u_0 \in H^1(\mathcal{M}^{(1)})$, defined above, which is compactly supported in $U^{(1)}$ and has zero trace on the intersection, by a smooth function v_0 which too is supported in $U^{(1)}$ and has zero trace on the intersection. Furthermore, as noted in Step 1 the intersection $\mathcal{M}^{(12)}$ divides $U_k^{(1)} \setminus \mathcal{M}^{(12)}$ into two components $U_\pm^{(1)}$. Since u_0 has zero trace in each of the parts and it can be approximated by smooth functions, v_0^+ and v_0^- compactly supported in open sets $U_+^{(1)}$ and $U_-^{(1)}$ respectively. The function v defined to be equal to v_0^+ on $U_+^{(1)}$, v_0^- on $U_-^{(1)}$, and zero on $\mathcal{M}^{(12)} \cap U_k^{(1)}$ is the desired smooth approximation of u_0 on $U_k^{(1)}$. □

A.6 Density of Lipschitz functions in $H^1(\mathcal{M})$ when $d^{(2)} - d^{(12)} \ge 2$

When $d^{(2)} - d^{(12)} \ge 2$ the Sobolev space $H^1(\mathcal{M})$ is the space of functions whose restriction to $\mathcal{M}^{(i)}$ belong to $H^1(\mathcal{M}^{(i)})$ for $i = 1, 2$. Here we show that the Lipschitz functions on the union, $\mathcal{M}$, are still dense in $H^1(\mathcal{M})$. This helps bridge the gap to the literature on metric measure spaces where the Dirichlet energy is typically defined on the closure of the space of Lipschitz functions.

Lemma A.8. *In the setting of Section 2.1, assume $d^{(2)} - d^{(12)} \ge 2$. We claim that Lipschitz continuous functions on $\mathcal{M}$ are dense in $H^1(\mathcal{M})$ which we previously characterized in Equation* (2.20) *as $H^1(\mathcal{M}^{(1)}) \times H^1(\mathcal{M}^{(2)})$.*

Proof. Since functions $u \in H^1(\mathcal{M})$ whose restrictions to $\mathcal{M}^{(1)}$ and to $\mathcal{M}^{(2)}$ are smooth are dense in $H^1(\mathcal{M})$ it suffices to show that such function u can be approximated in $H^1(\mathcal{M})$ by Lipschitz continuous functions. Let u_i be the smooth representative of the restriction of u to $\mathcal{M}^{(i)}$ for $i = 1, 2$. The construction is similar to the one in the proof of limsup results in Sections 3.3 and 4.3.

For $n > 1$ let $\varepsilon_n = \frac{1}{n}$. Let us consider n large enough so that the orthogonal geodesic projection $\pi : \mathcal{M}^{(2)} \to \mathcal{M}^{(12)}$ onto the intersection is uniquely defined in an ε_n-neighborhood of $\mathcal{M}^{(12)}$. We denote the distance of $x \in \mathcal{M}^{(2)}$ to $\mathcal{M}^{(12)}$ by $d(x) := \operatorname{dist}(x, \mathcal{M}^{(12)})$. Let

$$u_n(x) = \begin{cases} u^{(1)}(x) & \text{if } x \in \mathcal{M}^{(1)}, \\ u^{(1)}(\pi(x)) & \text{if } x \in \mathcal{M}^{(2)} \text{ and } d(x) \leq \varepsilon_n, \\ u^{(2)}(x) + \frac{\ln\sqrt{\varepsilon_n} - \ln d(x)}{\ln\sqrt{\varepsilon_n} - \ln\varepsilon_n}\left(u^{(1)}(\pi(x)) - u^{(2)}(x)\right) & \text{if } x \in \mathcal{M}^{(2)} \text{ and } \varepsilon_n < d(x) < \sqrt{\varepsilon_n} \\ u^{(2)}(x) & \text{if } x \in \mathcal{M}^{(2)} \text{ and } d(x) \geq \sqrt{\varepsilon_n}. \end{cases} \tag{A.3}$$

We note that u_n is Lipschitz by construction, since both π and d are Lipschitz functions on the regions where they are used. It remains to show that $u_n - u^{(2)} \to 0$ in $H^1(\mathcal{M}^{(2)})$. Let $A_n = \{x \in \mathcal{M}^{(2)} : d(x) < \sqrt{\varepsilon_n}\}$. Note that volume of A_n vanishes as $n \to \infty$ while u_n and $u^{(2)}$ are uniformly bounded on $\mathcal{M}^{(2)}$. Thus $u_n - u^{(2)} \to 0$ in $L^2(\mathcal{M}^{(2)})$. Note that for $x \in \mathcal{M}^{(2)}$

$$\nabla u_n(x) = \begin{cases} D\pi(x)^T \nabla u^{(1)}(\pi(x)) & \text{if } d(x) \leq \varepsilon_n, \\ \nabla u^{(2)}(x) + \frac{2\nabla d(x)}{d(x)\ln\varepsilon_n}\left(u^{(1)}(\pi(x)) - u^{(2)}(x)\right) & \\ +\frac{\ln\sqrt{\varepsilon_n} - \ln d(x)}{\ln\sqrt{\varepsilon_n} - \ln\varepsilon_n}\left(D\pi(x)^T \nabla u^{(1)}(\pi(x)) - \nabla u^{(2)}(x)\right) & \text{if } \varepsilon_n < d(x) < \sqrt{\varepsilon_n} \\ \nabla u^{(2)}(x) & \text{if } d(x) \geq \sqrt{\varepsilon_n}. \end{cases} \tag{A.4}$$

Showing that $\nabla u_n - \nabla u^{(2)}$ converges to zero in $L^2(\mathcal{M}^{(2)})$ boils down, due to uniform in x boundedness of $D\pi$, $\nabla u^{(1)}$, and ∇d, to showing that

$$\frac{1}{(\ln\varepsilon_n)^2} \int_{\{\varepsilon_n < d(x) < \sqrt{\varepsilon_n}\}} \frac{1}{d(x)^2}\, \mathrm{d}x \to 0 \qquad \text{as } n \to \infty.$$

This follows by disintegrating over perpendicular slices, and arguing analogously to (3.8) in each slice, to obtain

$$\int_{\{\varepsilon_n < d(x) < \sqrt{\varepsilon_n}\}} \frac{1}{d(x)^2}\, \mathrm{d}x = O(|\ln\varepsilon_n|).$$

$\square$