



# Dynamic Isoperimetry on Graphs and Weighted Riemannian manifolds

**Author:**

Kwok, Eric

**Publication Date:**

2018

**DOI:**

<https://doi.org/10.26190/unsworks/20274>

**License:**

<https://creativecommons.org/licenses/by-nc-nd/3.0/au/>

Link to license to see what you are allowed to do with this resource.

Downloaded from <http://hdl.handle.net/1959.4/59708> in <https://unsworks.unsw.edu.au> on 2024-05-06

# DYNAMIC ISOPERIMETRY ON GRAPHS AND WEIGHTED RIEMANNIAN MANIFOLDS

**Eric Kwok**

A thesis in fulfilment of the requirements for the degree of

Doctor of Philosophy



School of Mathematics and Statistics

Faculty of Science

February 20, 2018

PLEASE TYPE

THE UNIVERSITY OF NEW SOUTH WALES  
Thesis/Dissertation Sheet

Surname or Family name: Kwok

First name: Eric

Other name/s:

Abbreviation for degree as given in the University calendar: PhD

School: Mathematics and Statistics

Faculty: Science

Title: Dynamic isoperimetry on graphs and weighted Riemannian manifolds

Transport and mixing in dynamical systems are important properties for many physical, chemical, biological, and engineering processes. The detection of transport barriers for dynamics with general time dependence is a difficult, but important problem, because such barriers control how rapidly different parts of phase space (which might correspond to different chemical or biological agents) interact. The key factor is the growth of interfaces that partition phase space into separate regions. In a recent paper, Froyland introduced the notion of dynamic isoperimetry: the study of sets with persistently small boundary size (the interface) relative to enclosed volume, when evolved by the dynamics. Sets with this minimal boundary size to volume ratio were identified as level sets of dominant eigenfunctions of a dynamic Laplace operator. In this dissertation, we develop a data-driven approach for transport barrier detection, by extending and generalising dynamic isoperimetry to graphs and weighted Riemannian manifolds.

First we model trajectory data as dynamics of graphs. We use minimum disconnecting cuts to search for coherent structure in dynamic graphs, where the graph dynamic arises from a general sequence of vertex permutations. We develop a dynamic spectral partitioning method via a new dynamic Laplacian matrix. We prove a dynamic Cheeger inequality for graphs, and demonstrate the effectiveness of this dynamic spectral partitioning method on both structured and unstructured graphs.

We then generalise the dynamic isoperimetric problem on manifolds to situations where the dynamics (i) is not necessarily volume-preserving, (ii) acts on initial agent concentrations different from uniform concentrations, and (iii) occurs on a possibly curved phase space. Our main results include generalised versions of the dynamic isoperimetric problem, the dynamic Laplacian, the dynamic Cheeger's inequality, and the Federer-Fleming theorem. We illustrate the computational approach with some simple numerical examples.

Finally, we form a connection between the weighted graph version of our dynamic Laplacian matrix and the manifold dynamic Laplace operator. We then form a dynamic Laplacian-based manifold learning algorithm, which is designed to approximate solutions of our generalised dynamic isoperimetric problem from trajectory data. We highlight the robustness of our dynamic manifold learning method through numerical experiments.

Declaration relating to disposition of project thesis/dissertation

I hereby grant to the University of New South Wales or its agents the right to archive and to make available my thesis or dissertation in whole or in part in the University libraries in all forms of media, now or here after known, subject to the provisions of the Copyright Act 1968. I retain all property rights, such as patent rights. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

I also authorise University Microfilms to use the 350 word abstract of my thesis in Dissertation Abstracts International (this is applicable to doctoral theses only).

Signature

Date

19/9/2017

The University recognises that there may be exceptional circumstances requiring restrictions on copying or conditions on use. Requests for restriction for a period of up to 2 years must be made in writing. Requests for a longer period of restriction may be considered in exceptional circumstances and require the approval of the Dean of Graduate Research.

FOR OFFICE USE ONLY

Date of completion of requirements for Award:

**ORIGINALITY STATEMENT**

'I hereby declare that this submission is my own work and to the best of my knowledge it contains no materials previously published or written by another person, or substantial proportions of material which have been accepted for the award of any other degree or diploma at UNSW or any other educational institution, except where due acknowledgement is made in the thesis. Any contribution made to the research by others, with whom I have worked at UNSW or elsewhere, is explicitly acknowledged in the thesis. I also declare that the intellectual content of this thesis is the product of my own work, except to the extent that assistance from others in the project's design and conception or in style, presentation and linguistic expression is acknowledged.'

Signed .....

Date .....

**COPYRIGHT STATEMENT**

'I hereby grant the University of New South Wales or its agents the right to archive and to make available my thesis or dissertation in whole or part in the University libraries in all forms of media, now or here after known, subject to the provisions of the Copyright Act 1968. I retain all proprietary rights, such as patent rights. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

I also authorise University Microfilms to use the 350 word abstract of my thesis in Dissertation Abstract International (this is applicable to doctoral theses only).

I have either used no substantial portions of copyright material in my thesis or I have obtained permission to use copyright material; where permission has not been granted I have applied/will apply for a partial restriction of the digital copy of my thesis or dissertation.'

Signed .....

Date .....

**AUTHENTICITY STATEMENT**

'I certify that the Library deposit digital copy is a direct equivalent of the final officially approved version of my thesis. No emendation of content has occurred and if there are any minor variations in formatting, they are the result of the conversion to digital format.'

Signed .....

Date .....

## Abstract

Transport and mixing in dynamical systems are important properties for many physical, chemical, biological, and engineering processes. The detection of transport barriers for dynamics with general time dependence is a difficult, but important problem, because such barriers control how rapidly different parts of phase space (which might correspond to different chemical or biological agents) interact. The key factor is the growth of interfaces that partition phase space into separate regions. In a recent paper, Froyland introduced the notion of dynamic isoperimetry: the study of sets with persistently small boundary size (the interface) relative to enclosed volume, when evolved by the dynamics. Sets with this minimal boundary size to volume ratio were identified as level sets of dominant eigenfunctions of a dynamic Laplace operator. In this dissertation, we develop a data-driven approach for transport barrier detection, by extending and generalising dynamic isoperimetry to graphs and weighted Riemannian manifolds.

First we model trajectory data as dynamics of graphs. We use minimum disconnecting cuts to search for coherent structure in dynamic graphs, where the graph dynamic arises from a general sequence of vertex permutations. We develop a dynamic spectral partitioning method via a new dynamic Laplacian matrix. We prove a dynamic Cheeger inequality for graphs, and demonstrate the effectiveness of this dynamic spectral partitioning method on both structured and unstructured graphs.

We then generalise the dynamic isoperimetric problem on manifolds to situations where the dynamics (i) is not necessarily volume-preserving, (ii) acts on initial agent concentrations different from uniform concentrations, and (iii) occurs on a possibly curved phase space. Our main results include generalised versions of the dynamic isoperimetric problem, the dynamic Laplacian, the dynamic Cheeger's inequality, and the Federer-Fleming theorem. We illustrate the computational approach with some simple numerical examples.

Finally, we form a connection between the weighted graph version of our dynamic Laplacian matrix and the manifold dynamic Laplace operator. We then form a dynamic Laplacian-based manifold learning algorithm, which is designed to approximate solutions of our generalised dynamic isoperimetric problem from trajectory data. We highlight the robustness of our dynamic manifold learning method through numerical experiments.



# Acknowledgements

I would like to express my sincere gratitude to my supervisor Prof. Gary Froyland for the continuous support of my PhD study. His guidance, insightful comments, and attention to detail had helped me tremendously in the last 3.5 years. I would also like to thank the School of Mathematics and Statistics at the University of New South Wales. Finally, I would like to thank my family and friends.



# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Dynamic isoperimetric problem on graphs . . . . .	2
1.1.1	Cheeger inequality on static graphs . . . . .	2
1.1.2	Vertex permutation dynamics of graphs . . . . .	3
1.1.3	Cheeger inequality for graphs . . . . .	4
1.1.4	Original contributions of Chapter 2 . . . . .	5
1.2	Dynamic isoperimetry problem on weighted Riemannian manifolds . .	5
1.2.1	Cheeger ratio for weighted Riemannian manifolds . . . . .	6
1.2.2	Dynamics on Manifolds . . . . .	7
1.2.3	Cheeger inequality for unweighted Riemannian manifolds . . .	9
1.2.4	Original contributions in Chapter 3 . . . . .	9
1.3	Manifold learning for dynamic isoperimetric problems . . . . .	10
1.3.1	Point-cloud data as weighted graph . . . . .	10
1.3.2	Trajectory data as dynamics of graphs . . . . .	11
1.3.3	Laplacian-based manifold learning . . . . .	12
1.3.4	Original contributions of Chapter 4 . . . . .	13
<b>2</b>	<b>Partitions of graphs that are robust to vertex permutation dynam- ics</b>	<b>15</b>
2.1	A graph Cheeger constant for dynamic graphs . . . . .	16
2.2	A spectral method for dynamic graphs . . . . .	18
2.3	Dynamics over $\tau$ time steps . . . . .	21
2.4	Numerical method and experiments . . . . .	22
2.4.1	Example 1: A structured graph . . . . .	23
2.4.2	Example 2: Randomly generated graph. . . . .	29
2.5	Conclusion to the chapter . . . . .	30
<b>3</b>	<b>A dynamic Laplacian for identifying Lagrangian coherent struc- tures on weighted Riemannian manifolds.</b>	<b>33</b>

3.1	Primer on differential geometry . . . . .	36
3.1.1	Shear on a two-dimensional cylinder . . . . .	38
3.2	A dynamic isoperimetric problem on weighted manifolds . . . . .	40
3.2.1	Multiple discrete time steps and continuous time . . . . .	41
3.2.2	Dynamic Federer-Fleming theorem on weighted manifolds . . . . .	42
3.3	The dynamic Laplace operator on weighted manifolds . . . . .	44
3.3.1	The dynamic Laplace-Beltrami operator . . . . .	44
3.3.2	Continuous time . . . . .	47
3.3.3	Spectral theory and a dynamic Cheeger inequality on weighted manifolds . . . . .	47
3.4	Geometry and probability: linking finite-time coherent sets with dy- namic isoperimetry . . . . .	51
3.5	Numerical experiments . . . . .	54
3.5.1	Numerical approximation for $\mathcal{H}$ and $\mathcal{H}^*$ . . . . .	54
3.5.2	Finite-difference estimate for $\Delta^{dyn}$ . . . . .	55
3.5.3	Case study 1: dynamics on a cylinder . . . . .	57
3.5.4	Case study 2: dynamics on a torus . . . . .	61
3.6	Conclusions to the chapter . . . . .	63
<b>4</b>	<b>A dynamic manifold learning method for approximating Lagrangian coherent structures.</b>	<b>65</b>
4.1	A robust Laplacian-based manifold learning method for weighted Rie- mannian manifolds . . . . .	67
4.1.1	Graph Laplacian for scalar weighted graphs . . . . .	68
4.1.2	From graph Laplacian to Laplace-Beltrami operator . . . . .	70
4.1.3	Algorithm for weighted Laplacian eigenmap . . . . .	71
4.1.4	Numerical experiments for weighted Laplacian eigenmap . . . . .	72
4.2	A Dynamical manifold learning for dynamics on weighted Riemannian manifolds . . . . .	79
4.2.1	From dynamic graph Laplacian to dynamic Laplacian . . . . .	81
4.2.2	Trajectory data from multiple time step dynamics . . . . .	82
4.2.3	Algorithm for dynamic Laplacian eigenmap . . . . .	84
4.2.4	Numerical experiments for dynamic Laplacian eigenmap . . . . .	85
4.3	Conclusion to the chapter . . . . .	94
<b>A</b>	<b>Appendix for Chapter 2</b>	<b>97</b>
A.1	The proof of Theorem 2.2.2 . . . . .	97
A.2	The proof of Theorem 2.2.3(2) . . . . .	98

<b>B</b>	<b>Appendix for Chapter 3</b>	<b>111</b>
B.1	Muckenhoupt weights $A_p$ . . . . .	111
B.2	Additional notes on differential geometry . . . . .	113
B.2.1	Differential forms . . . . .	113
B.2.2	Differential operators on weighted manifolds . . . . .	115
B.2.3	Properties of $T_*$ and $T^*$ . . . . .	116
B.2.4	Local properties of charts . . . . .	122
B.3	Weighted Sobolev spaces . . . . .	124
B.4	The proof of Theorem 3.2.4 . . . . .	134
B.5	The proof of Theorem 3.3.3 . . . . .	138
B.5.1	Weak formulation of the $\Delta^{dyn}$ eigenproblem . . . . .	139
B.5.2	Existence of weak solution and variational characterisation of eigenvalues . . . . .	142
B.5.3	Ellipticity and global regularity of weak solutions . . . . .	148
B.6	The proof of Theorem 3.3.4 . . . . .	151
B.6.1	Time-discrete and time-continuous case . . . . .	154
B.7	The proof of Theorem 3.4.1 . . . . .	155
<b>C</b>	<b>Appendix for Chapter 4</b>	<b>163</b>
C.1	The proof of Theorem 4.1.2 . . . . .	163
C.2	The proof of Theorem 4.1.4 . . . . .	163
C.3	The proof of Theorem 4.2.2 . . . . .	166



# List of Figures

1.1	Partition of 5-vertex graph . . . . .	3
1.2	Vertex permutation dynamics on 5-vertex graph . . . . .	3
1.3	Disconnecting curve on 2-dimensional cylinder . . . . .	6
1.4	Linear shear on 2-dimensional cylinder . . . . .	8
1.5	Random samples drawn from a 2-dimensional torus. . . . .	11
1.6	Laplacian eigenmap on the random samples in Figure 1.5. . . . .	13
2.1	Pullback of partition boundary and total degree function on graphs .	17
2.2	Ellingham-Horton 54-graph static spectral partition . . . . .	24
2.3	Spectra of static and dynamic graph Laplacians . . . . .	25
2.4	Ellingham-Horton 54-graph dynamic partition . . . . .	27
2.5	Ellingham-Horton 54-graph dynamic partition (multiple time-steps) .	28
3.1	Nonlinear shear on weighted 2-dimensional cylinder . . . . .	39
3.2	Dynamic spectral partition of 2-dimensional cylinder (nonlinear shear)	59
3.3	Dynamic spectral partition of 2-dimensional cylinder (non-volume- preserving shear) . . . . .	60
3.4	Dynamic spectral partition of 2-dimensional torus (standard map) . .	62
4.1	Improved Laplacian eigenmap on nonuniform point-cloud data . . . .	75
4.2	Spectra comparison of graph and manifold Laplacians . . . . .	76
4.3	Manifold learning on weighted 2-dimensional torus . . . . .	78
4.4	Nonuniform trajectory data on 2-dimensional torus . . . . .	86
4.5	Eigenfunctions of space-time diffusion matrix . . . . .	87
4.6	Eigenfunctions of dynamic Laplacian matrix . . . . .	88
4.7	Approximation of transport barrier from trajectory data (standard map) . . . . .	90
4.8	Eigenfunction approximation of dynamic Laplace operator . . . . .	91
4.9	Dynamic Laplacian eigenmap on ocean drifter data . . . . .	93
4.10	Approximation of transport barrier from ocean drifter data . . . . .	93



# List of Tables

2.1	Spectral partition Ellinham-Horton 54 graph . . . . .	26
2.2	Spectral partition of Ellingham-Horton 54 graph (multiple time-steps)	29
2.3	Spectral partition randomly generated graph . . . . .	30

# Chapter 1

## Introduction

The classical isoperimetric problem on a plane can be stated as follows: among all simple 2-dimensional closed curves  $\mathcal{C}$  of fixed length  $L$ , find the set of curves for which the enclosed area  $A$  is maximum. The intricate geometry of this question has attracted substantial interest from many mathematicians, with the earliest work on the problem dated as far back as antiquity. The ancient Greeks proposed that the unique solution to the isoperimetric problem is a circle, when Zenodorus attempted to prove that a circle has greater enclosed area than any polygon of the same perimeter. Due to Zenodorus's work, the isoperimetric problem has often been expressed as the *isoperimetric inequality* [95]: for any curve in  $\mathcal{C}$  show that

$$L^2 \geq 4\pi A, \tag{1.1}$$

with equality if and only if the curve is a circle. The proof<sup>1</sup> of the isoperimetric inequality arrived in the 19<sup>th</sup> century by Steiner [115], solving the isoperimetric problem on a plane.

Inspired by the elegance of the inequality (1.1), many generalisations and extensions of the isoperimetric problem have been formulated in areas such as differential geometry [23], mathematical physics [18] and dynamical systems [49], as well as non-geometric analogies in partial differential equation [97] and graph theory [25]. The purpose of this thesis is the development of a *dynamic isoperimetric problem* in two settings: graphs and weighted Riemannian manifolds. The motives, formulations, previous works and connections of the two dynamic isoperimetric problems are elaborated in Sections 1.1, 1.2 and 1.3.

---

<sup>1</sup>This proof was incomplete as it did not show the existence of the solution. The more vigorous treatment using Steiner's techniques were developed in [19].

## 1.1 Dynamic isoperimetric problem on graphs

Many spatio-temporal systems arising from physical processes can be modelled as dynamics on graphs, or dynamics of graphs (see [33] for a discussion of the distinction). The motivation for a dynamic isoperimetry problem on a graph is an attempt to understand the complex combination of dynamics and graph structure in terms of graph connectivity. The strength of graph connectivity concerns the number of edges that needs to be removed in order to disconnect the graph, and is a fundamental characteriser of graph structure. Efficient algorithms for graph partitioning and the detection of community structures have led to applications in image segmentation [110, 64], parallel computing [72], social graph analysis [108], dynamical systems [50], image and video synthesis [81], nonlinear fluid flow [58], and route planning [30] (see [16] for a recent review on several numerical algorithms and applications).

### 1.1.1 Cheeger inequality on static graphs

In order to quantitatively measure how interconnected a graph is, we make use of the well-known *graph Cheeger ratio* [25, 64, 8]: denote a simple, connected graph by  $G = G(V, E)$ , where  $V = \{v_1, v_2, \dots, v_k\}$  is the vertex set and  $E \subseteq V \times V$  is the set of (undirected) edges. We define a *disconnection*  $G' = G'(V, E')$  of  $G$  by partitioning  $V = V_1 \cup V_2$  into two disjoint vertex sets  $V_1, V_2$  and forming the reduced edge set  $E' = E \setminus \{[v_i, v_j] \in E : v_i \in V_1, v_j \in V_2\}$ , where  $[v_i, v_j] \in E$  is an undirected edge. The *balanced graph bisection problem* for a connected graph  $G(V, E)$  asks for a disconnection  $G'(V, E')$ , where the set of removed edges  $E \setminus E'$  is minimised, while maintaining a similar number of vertices (counting multiplicity of degree) between  $V_1$  and  $V_2$ . We define the *partition boundary*  $C(V_1, V_2)$  between the partitions  $V_1$  and  $V_2$  as the set of edges removed to disconnect  $G$ ; that is,  $C(V_1, V_2) = \{[v_i, v_j] \in E : v_i \in V_1, v_j \in V_2\}$ . The total degree of the vertex set  $V' \subseteq V$  is denoted by  $D(V') := \sum_{v_i \in V'} d(v_i)$ , where  $d(v_i)$  is the degree<sup>2</sup> of the vertex  $v_i$ . For any nontrivial partition  $V = V_1 \cup V_2$ , the graph Cheeger ratio for  $G$  is defined by the number

$$H_G(V_1, V_2) = \frac{|C(V_1, V_2)|}{\min\{D(V_1), D(V_2)\}}. \quad (1.2)$$

A partition  $\{V_1, V_2\}$  that achieves a low graph Cheeger ratio  $H_G(V_1, V_2)$  has high internal connectivity within each component corresponding to vertices  $V_i, i = 1, 2$ , and low connectivity between the two components. Moreover, neither component is small in terms of total degree.

<sup>2</sup>The degree of the vertex  $v_i$  is defined as the cardinality of the set  $\{v_j \in V : [v_i, v_j] \in E\}$ .

**Example 1.1.1.** In Figure 1.1, setting  $V_1 = \{v_4, v_5\}$ ,  $V_2 = \{v_1, v_2, v_3\}$ , we have  $|C(V_1, V_2)| = |\{[v_2, v_4]\}| = 1$ ,  $D(V_1) = 2 + 1 = 3$ ,  $D(V_2) = 1 + 3 + 1 = 5$ . Thus,  $H_G(V_1, V_2) = 1/\min\{3, 5\} = 1/3$ .

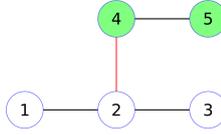


Figure 1.1: Graph with 5 vertices, coloured vertices =  $V_1$ , coloured edges =  $E \setminus E'$  and  $H_G(V_1, V_2) = 1/3$ .

### 1.1.2 Vertex permutation dynamics of graphs

The connective structure on an evolving graph can be transformed as time progresses. Examples of graph dynamics include transmission of diseases in populations [83], transmission of happiness in social graphs [46], and synchronisation of community structures [83]. We consider the situation where the vertex labels of  $G$  are subjected to permutation; dynamics of graphs. Abstractly, we have a permutation  $\pi_v : V \rightarrow V$ , which induces an action  $\pi_e : E \rightarrow \hat{E}$  on edges via  $\pi_e([v_i, v_j]) = [\pi_v(v_i), \pi_v(v_j)]$ ,  $[v_i, v_j] \in E$ . In this way, the entire graph  $G$  is transformed by  $\pi : G \rightarrow \hat{G}$ , where  $\pi(G(V, E)) = G(\pi_v(V), \pi_e(E))$ . The transformation  $\pi$  is a graph isomorphism: clearly edges  $\pi_e([v_i, v_j]), \pi_e([v_j, v_i])$  are adjacent in  $\hat{G}$  if edges  $[v_i, v_j], [v_j, v_i]$  edges are adjacent in  $G$ .

**Example 1.1.2.** For example, in Figure 1.2, we see the image of the graph of Figure 1.1 under the cyclic permutation  $\pi_v(v_i) = v_{i+1} \pmod{5}$ ,  $i = 1, 2, 3, 4, 5$ . One has  $H_{\pi(G)}(V_1, V_2) = 1$ .

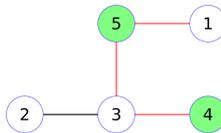


Figure 1.2: The graph of Figure 1.2 under cyclic permutation.

One can ask the very natural question: how well does a *fixed* partition  $\{V_1, V_2\}$  represent a minimal disconnection of *both*  $G$  and  $\pi(G)$ , according to the edge sets  $E$  and  $\pi_e(E)$ , respectively; that is, we ask for a nontrivial partition  $V = V_1 \cup V_2$ , so that the graph Cheeger ratios  $H_G(V_1, V_2)$  for  $G$  and  $H_{\pi(G)}(V_1, V_2)$  for  $\pi(G)$  are *both*

small. Persistently highly interconnected subregions on a dynamic vertex-labelled graph can highlight important physical properties of the underlying process, such as the stability of subprocesses and community structures over time. Partitions that are robust to specific vertex permutations offer a method to add extra pressure to keep the permuted vertices in the same community, and are required in situations in which there is uncertainty about a graph such as vertex identity. Persistently interconnected regions in a network of human contacts throughout the day could arise from individual's professional and social status, and are vital for modelling disease spread in human population. Similarly, the transmission of happiness in a social network could be dependent on the history of friendships.

It is well-known that the complexity of graph partition problems are NP-complete (see [59]), however, the importance of these problems have generated an extensive collection of heuristic algorithms that can produce good solutions [45, 16]. A very popular graph partition method for studying graph connective structures is known as *spectral graph partition*. This approach was initiated by Fiedler [43] (also [5]), and has been developed by several authors (e.g. [3, 91]). The essence of spectral graph partition is the Cheeger inequality [25], which we now review.

### 1.1.3 Cheeger inequality for graphs

Recall that the graph Cheeger ratio (1.2) measures the connectedness of a graph. To generate minimum-cut balanced partitions of a vertex-labelled graphs, one considers the *graph Cheeger constant* defined by

$$\mathbf{H}_G = \min_{V_1, V_2 \text{ partition } V} H_G(V_1, V_2). \quad (1.3)$$

Introducing the graph Laplacian [25] for  $G(V, E)$ . Define the *adjacency* matrix  $\mathbf{A}$  by

$$\mathbf{A}_{ij} = \begin{cases} 1 & \text{if } [v_i, v_j] \in E, i \neq j \\ 0 & \text{otherwise} \end{cases}, \quad (1.4)$$

and the *degree* matrix  $\mathbf{D}$  by a  $k \times k$  diagonal matrix with entries  $\mathbf{D}_{ii} = d(v_i)$ , for  $1 \leq i, j \leq k$ . The *graph Laplacian* is defined by

$$\mathbf{L} = \mathbf{A} - \mathbf{D}. \quad (1.5)$$

The *normalised graph Laplacian* is defined by  $\mathcal{L} = \mathbf{D}^{-1/2} \mathbf{L} \mathbf{D}^{-1/2}$ ; i.e.  $\mathcal{L}$  is the  $k \times k$  symmetric matrix

$$\mathcal{L}_{ij} = \begin{cases} \frac{1}{\sqrt{d(v_i)d(v_j)}} & \text{if } [v_i, v_j] \in E, i \neq j \\ -1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}. \quad (1.6)$$

Standard results concerning  $\mathcal{L}$  are: (i) the eigenvalues  $\lambda_1 \geq \lambda_2 \geq \dots$  of  $\mathcal{L}$  are non-positive and real, and (ii) the eigenvalue  $\lambda_1 = 0$  and has unit multiplicity (so that  $\lambda_2 < 0$ ) [44, 5]. The eigenvector corresponding to  $\lambda_2$  is commonly used to construct a balanced bisection  $V_1, V_2$  of  $G$  with a small number of edges connecting  $V_1$  to  $V_2$ . One computes  $\mathbf{g}_2$ , the eigenvector of  $\mathcal{L}$  corresponding to  $\lambda_2$ , and sets  $\mathbf{f} = \mathbf{D}^{-1/2}\mathbf{g}_2$ . For each  $\beta \in \{f_i\}_{i=1}^{k-1}$ , one defines the sets  $V_1^\beta = \{v_i \in V : f_i \leq \beta\}$  and  $V_2^\beta = \{v_i \in V : f_i > \beta\}$ . The sets  $V_1^\beta, V_2^\beta$  partition  $V$  and there are at most  $k - 1$  nontrivial partitions of this form. One evaluates  $H_G(V_1^\beta, V_2^\beta)$  for these at most  $k - 1$  distinct partitions and selects the partition that minimises  $H_G(V_1^\beta, V_2^\beta)$ . This approach was described in [38]; see [64] for a modern treatment. To create partitions of more than two components, further eigenvectors  $\mathbf{g}_3, \mathbf{g}_4, \dots$  can be used in an analogous way to existing algorithms in the static case; see [67, 21, 4, 110] for the use of multiple eigenvectors to partition static graphs. One has the celebrated Cheeger inequality [25]:

$$\mathbf{H}_G \leq \sqrt{-2\lambda_2} \leq 2\sqrt{\mathbf{H}_G}, \quad (1.7)$$

where  $\lambda_2$  is the smallest nonzero eigenvalue of  $\mathcal{L}$ . Moreover, the optimal partition  $\{V_1^\beta, V_2^\beta\}$  that minimises  $H_G(V_1^\beta, V_2^\beta)$  over  $\beta \in \{f_i\}_{i=1}^{k-1}$ , satisfies  $H_G(V_1^\beta, V_2^\beta) \leq \sqrt{-2\lambda_2}$ .

### 1.1.4 Original contributions of Chapter 2

The classical graph Cheeger inequality is limited to application on  $G$  or  $\pi(G)$ , and in general cannot be used to find a graph partition that represents a minimal disconnection on *both*  $G$  and  $\pi(G)$ . In Chapter 2, we consider a *dynamic minimum-cut balanced partition* on dynamics of graphs. In particular, we generalise (1.7) to a dynamic graph Cheeger inequality (see Theorem 2.2.3), and formulate an efficient algorithm designed to find good solutions to the dynamic isoperimetric problem on graphs (see Algorithm 2.1); the contents of Chapter 2 are published in [53].

## 1.2 Dynamic isoperimetry problem on weighted Riemannian manifolds

The mathematics of transport in nonlinear dynamical systems has received considerable attention for more than two decades, driven in part by applications in fluid dynamics, atmospheric and ocean dynamics, molecular dynamics, granular flow and other areas. We refer the reader to [96, 102, 89, 6, 123] for reviews of transport

and transport-related phenomena. Many transport processes can be modelled as dynamics on weighted Riemannian manifolds [66], because each of these manifolds is equipped with a measure to track the mass distribution of the physical quantity being transported (e.g. chemical concentrations in fluids, air mass in the atmosphere, salt in the ocean). We attempt to characterise transport barriers by identifying subsets of a weighted Riemannian manifold that have persistently small boundary size to enclosed mass ratio as the domain is transformed by a general dynamical system. Thus we introduce a dynamic isoperimetric problem on weighted Riemannian manifolds.

### 1.2.1 Cheeger ratio for weighted Riemannian manifolds

To provide a mathematical description of our dynamic isoperimetry problem on a weighted Riemannian manifold, we develop a weighted version of the well-known Cheeger ratio [24]: let  $M$  denote a compact, connected  $r$ -dimensional  $C^\infty$ -Riemannian manifold, and  $\Gamma$  denote a piecewise  $C^\infty$ -hypersurface that disconnects  $M$  into full dimensional submanifolds  $M_1, M_2$ ; that is  $\{M_1, M_2, \Gamma\}$  is a partition of  $M$ .

**Example 1.2.1.** For example, in Figure 1.3 the manifold  $M$  is a 2-dimensional cylinder  $[0, 4)/\sim \times [0, 1]$ , where  $\sim$  is identification at interval endpoints; that is,  $M$  is periodic in the first coordinate with period 4. A piecewise  $C^\infty$ -curve in  $M$  can either be a curve  $\Gamma$  from boundary points to boundary points as in Figure 1.3a, or a closed curve  $\Gamma'$  as in Figure 1.3b.

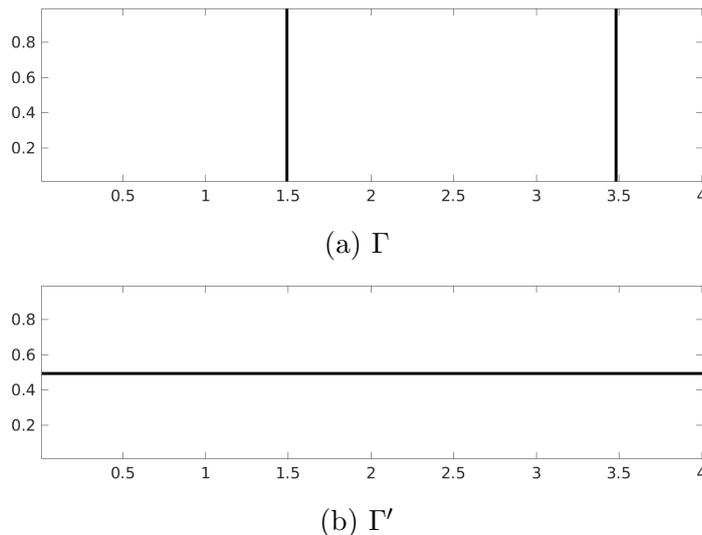


Figure 1.3: Disconnecting curves  $\Gamma$  and  $\Gamma'$  on a 2-dimensional cylinder.

Let  $\mu_r$  denote the measure equipped on  $M$ , and assume  $\mu_r$  is absolutely continuous. The size of a set  $M_1 \subset M$  is given by  $\mu_r(M_1)$  and by a process of inducing

explained in Chapter 3, we develop a measure  $\mu_{r-1}$  to determine the size of  $(r-1)$ -dimensional objects such as  $\Gamma$ . In order to track the transport of non-uniformly distributed passive tracers, we require the measure  $\mu_r$  to represent the initial distribution to be tracked. Similarly, in order to estimate the amount of material that would be ejected through the boundary at any given time by small amplitude isotropic diffusion, we require the measure  $\mu_{r-1}$  to compute initial boundary size. To find subsets of  $M$  with small boundary to interior size ratios with respect to the measure  $\mu_{r-1}$  and  $\mu_r$ , we define the *weighted* Cheeger ratio by

$$H_M(\Gamma) := \frac{\mu_{r-1}(\Gamma)}{\min\{\mu_r(M_1), \mu_r(M_2)\}}. \quad (1.8)$$

The numerator of (1.8) quantifies the boundary size of  $\Gamma$ , while denominator of (1.8) is a standard normalisation condition in isoperimetry problems to avoid trivial solutions and ensure that both  $M_1$  and  $M_2$  are of macroscopic size [23].

**Example 1.2.2.** *Suppose the 2-dimensional cylinder in Figure 1.3 is equipped with the Lebesgue measure  $\mu_2 = \ell_2$ , then the induced co-dimension 1 measure  $\mu_1$  is given by 1-dimensional Lebesgue  $\ell_1$ . Therefore in Figure 1.3a, the size of the disconnecting curve  $\Gamma$  is  $\ell_1(\Gamma) = 2$  and size of the interiors  $M_1, M_2$  are  $\ell_2(M_1) = \ell_2(M_2) = 2$ . Hence  $H_M(\Gamma) = 2/2 = 1$ . By similar calculations,  $H_M(\Gamma') = 2$  for the disconnection curve  $\Gamma'$  in Figure 1.3b.*

## 1.2.2 Dynamics on Manifolds

Let  $N$  denote another compact, connected  $r$ -dimensional  $C^\infty$ -Riemannian manifold, and suppose that the dynamics over a finite time duration is given by  $T : M \rightarrow N$ . The manifold  $N$  is equipped with a Riemannian metric  $n$ , and an absolutely continuous measure  $\nu_r$ . For many applications we may not want  $n$  to be related to the metric tensor  $m$  on  $M$ , however conservation of mass enforces  $\nu_r := \mu_r \circ T^{-1}$ ; i.e.  $n$  need not be the pushforward of  $m$  but  $\nu_r$  must be the pushforward of  $\mu_r$ . Importantly, the  $\mu_{r-1}$ -size of a hypersurface can be drastically different compared to the  $\nu_{r-1}$ -size of its the image under dynamics of  $T$ .

**Example 1.2.3.** *Returning to Example 1.2.2. Figure 1.4 shows the images of the disconnecting curves  $\Gamma$  and  $\Gamma'$  in Figure 1.3a and 1.3b under linear shear  $T(x, y) = ((x+y \bmod 4), y)$ . Clearly, the disconnecting curve  $\Gamma'$  is invariant under the action of  $T$  as shown in Figure 1.4b. However, in Figure 1.4b the size of the new disconnecting curve  $T\Gamma$  of  $N$  have increased by a factor of  $\sqrt{2}$ . By volume preservation of  $T$ , one has  $\ell_2(M_i) = \ell_2(TM_i) = 2$ , for  $i = 1, 2$ , where  $M_1, M_2$  are the partition*

elements of  $M$  separated by either  $\Gamma$  or  $\Gamma'$ . Therefore, computing the Cheeger ratio on  $N$ , one has  $H_N(T\Gamma) = \sqrt{2}$  and  $H_N(T\Gamma') = 2$ .

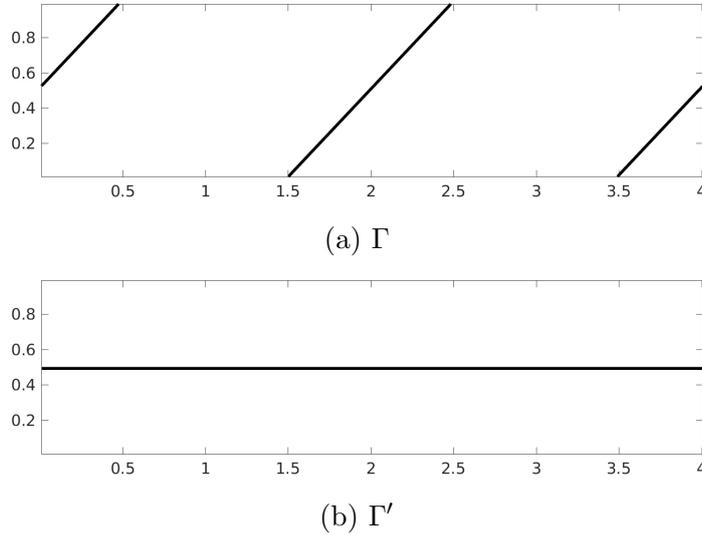


Figure 1.4: Disconnecting curves  $\Gamma$  and  $\Gamma'$  in Figure 1.3 transformed by a linear shear.

To identify transport barriers under a single iteration of  $T$ , we search over all piecewise  $C^\infty$ -disconnecting hypersurfaces of  $M$ , and find a hypersurface  $\Gamma$  that represents a balanced minimal disconnection on  $M$  according to the  $\mu_{r-1}$ -size of  $\Gamma$  and the  $\nu_{r-1}$ -size of  $T\Gamma$ ; that is, we attempt to find a *fixed* disconnecting hypersurface  $\Gamma$  such that  $H_M(\Gamma)$  and  $H_N(T\Gamma)$  are *both* small. In situations where  $T$  is a concatenation of several maps over several discrete time steps, or the flow map for a time-dependent vector field over some duration  $\tau$ , we measure the weighted Cheeger ratio on the disconnecting curve at each discrete time step, or we continuously check the boundary size under continuous time dynamics (see Section 3.2.1).

We have seen previously in Section 1.1.3, that the analytical inequality (1.7) formed a link between the linear operator  $\mathcal{L}$  (1.6) and the connectedness (1.2) of the associated graph. On a Riemannian manifold, the analogous operator is the Laplace-Beltrami operator. The Laplace-Beltrami operator has deep connections to the geometry of the Riemannian manifold on which it is defined [22, 13]. The two well-known classic results are: the Cheeger inequality on a Riemannian manifold [24] and the Federer-Fleming theorem [41]. We now give a brief review on these results.

### 1.2.3 Cheeger inequality for unweighted Riemannian manifolds

To partition a Riemannian manifold into  $\{M_1, M_2, \Gamma\}$ , such that the  $\mu_{r-1}$ -size of  $\Gamma$  is minimal, while maintaining a balanced  $\mu_r$ -size between  $M_1$  and  $M_2$ , one considers the *Cheeger constant* defined by

$$\mathbf{H}_M := \inf_{\Gamma} H_M(\Gamma), \quad (1.9)$$

where  $H_M$  is as in (1.8). Suppose the measure  $\mu_r$  equipped on the Riemannian manifold  $M$  is the full-dimensional volume measure<sup>3</sup>. The following spectral properties of the Laplace-Beltrami operator  $\Delta_m$  on  $M$  are well-known: (i) eigenvalues  $\lambda_1 \geq \lambda_2 \geq \dots$  of  $\Delta_m$  are non-positive and real, and (ii) the eigenvalue  $\lambda_1 = 0$  and has unit multiplicity (so that  $\lambda_2 > 0$ ) [22]. One has the *dynamic Cheeger inequality*

$$\mathbf{H}_M \leq 2\sqrt{-\lambda_2}, \quad (1.10)$$

where  $\lambda_2$  is the first non-trivial eigenvalue of  $\Delta_m$ . In addition to the inequality (1.10), one has the classical Federer-Fleming theorem (see e.g. p.131 [23]), which equates the geometric constant  $\mathbf{H}_M$  to the Sobolev constant (a functional representation of  $\mathbf{H}_M$ , where the disconnecting hypersurface  $\Gamma$  is generated by level surfaces of smooth functions).

To find a disconnecting hypersurface with small Cheeger constant  $H_M$ , we suggest a spectral manifold partitioning algorithm analogous to the graph partitioning method outlined in Section 1.1.3: let  $\phi_2$  be the first nontrivial eigenvector of  $\Delta_m$ . One searches over the level surfaces  $\Gamma^\beta = \{x \in M : \phi_2(x) = \beta\}$  to minimise the ratio  $H_M(\Gamma^\beta)$ , rather than all piecewise smooth  $C^\infty$ -disconnecting hypersurfaces  $\Gamma$  that partition  $M$  into  $\{M_1, M_2\}$ .

### 1.2.4 Original contributions in Chapter 3

The notion of combining dynamics and isoperimetry was first introduced by Froyland [49], whereby the optimisation problem (1.9) was modified to take into account of evolving size of the disconnecting curve  $\Gamma$  in a time-averaged sense under general time-dependent nonlinear dynamics. In particular, a dynamic generalisation of the Federer-Fleming theorem and Cheeger inequality in the setting of a flat, unweighted Riemannian manifold under volume-preserving dynamics was formulated in [49]. Moreover, Froyland [49] formed a link between his purely geometric approach of

---

<sup>3</sup>The volume measure is defined by the the nowhere-vanishing top-dimensional form associated with the metric tensor  $m$  on  $M$ , see Appendix B.2.1.

dynamic isoperimetry for finding transport barriers to the purely probabilistic constructions for finding almost invariant sets [48]. In Chapter 3, *We extend the results of [49] in three ways: (1) to dynamics that is not volume preserving, (2) to tracking the transport of non-uniformly distributed tracers, and (3) to dynamics operating on curved manifolds*; the contents of Chapter 3 are published in [54].

## 1.3 Manifold learning for dynamic isoperimetric problems

Manifold learning is a method in nonlinear dimensionality reduction, used for approximating low-dimensional features of a manifold from point-cloud data belonging to a high dimensional space [14, 10, 29]. Applications of manifold learning include data representation [11], pattern recognition [70] and image processing [122]. Manifold learning methods have also been used in dynamical systems for discovering slow manifolds [2, 93, 121]. The motivation of this work is the approximation of transport barriers from sparse trajectory data. In particular, following the geometric characterisation of transport barriers in Section 1.2, we attempt to approximate solutions to the dynamic isoperimetric problem on weighted Riemannian manifolds from non-uniformly distributed trajectory data.

### 1.3.1 Point-cloud data as weighted graph

The evolving structure of trajectory data can be modelled as dynamics of graphs. First we capture the local geometry of the initial points of the trajectory data via a weighted graph. Let  $M$  and  $N$  be compact, connected  $r$ -dimensional Riemannian manifolds, which are embedded in a possibly higher dimensional Euclidean space  $\mathbb{R}^d$ ;  $d \geq r$ . On  $M$  we place a Riemannian metric  $m$  and an absolutely continuous measure  $\mu_r$ ; the measure  $\mu_r$  describes the mass distribution of the physical quantity being transported by dynamics as in Section 1.2. Let  $S^k = \{x_1, x_2, \dots, x_k\}$  be a finite set of points randomly drawn from  $M$  in independently and identically distributed (i.i.d) fashion.

**Example 1.3.1.** *Suppose  $M$  is a 2-dimensional torus  $M := [0, 4) \times [0, 1) \setminus \sim$ . In Figure 1.5a, the set  $S^k$  is 1000 data points randomly sampled from  $M$  according to a uniform probability distribution. In Figure 1.5b, the set  $S^k$  is 1000 data points randomly sampled from  $M$  according to a nonuniform probability distribution.*

For fixed  $\epsilon > 0$ , one forms a weighted graph from  $S^k$  by taking  $x_1, x_2, \dots, x_k$  as vertices, and defining  $w_{ij}^{\mu, \epsilon}$  as the weight of the edge between the vertex pair  $x_i, x_j$  for

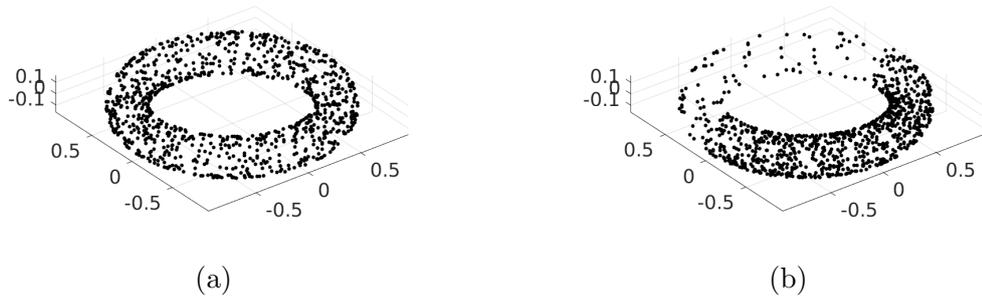


Figure 1.5: Random samples drawn from a 2-dimensional torus.

each  $1 \leq i, j \leq k$ . The edges  $w_{ij}^{\mu, \epsilon}$  are designed to capture the  $\mu_r$ -weighted geometry between the data points  $x_i$  and  $x_j$  via a process to be elaborated in Chapter 4, and the scalar  $\epsilon$  is a standard parameter in Laplacian-based manifold learning [10]. We denote the set of all edge weights by  $W^{\mu, \epsilon} = \{w_{ij}^{\mu, \epsilon}\}_{i,j=1}^k$ , and use the notation  $G(S^k, W^{\mu, \epsilon})$  for the weighted graph formed from  $S^k$  and  $W^{\mu, \epsilon}$  as above.

### 1.3.2 Trajectory data as dynamics of graphs

Recall that  $N$  is another compact, connected  $r$ -dimensional Riemannian manifold embedded in  $\mathbb{R}^d$ ;  $d \geq r$ . Suppose the dynamics over a finite time interval is given by  $T : M \rightarrow N$ . Under a single iteration of  $T$ , the measure  $\mu_r$  on  $M$  is pushed forward to the measure  $\nu_r := \mu_r \circ T^{-1}$  on  $N$ , and the point-cloud data  $S^k$  is transformed into  $\hat{S}^k = \{Tx_1, Tx_2, \dots, Tx_k\}$ , forming the trajectory data  $\{x_i, Tx_i\}$  for  $i = 1, 2, \dots, k$ . Similar to Section 1.3.1, one forms another weighted graph from  $\hat{S}^k$  by taking  $Tx_1, Tx_2, \dots, Tx_k$  as vertices, and setting  $w_{ij}^{\nu, \epsilon}$  as the edge weights between the vertex pairs  $Tx_i, Tx_j$  for each  $1 \leq i, j \leq k$ . We denote the weighted graph formed from  $\hat{S}^k$  by  $G(\hat{S}^k, W^{\nu, \epsilon})$ , where  $W^{\nu, \epsilon} = \{w_{ij}^{\nu, \epsilon}\}_{i,j=1}^k$ . In situations where the dynamics is over  $\tau$  time steps, by obvious modifications to the domain of  $Q_{\rho, \epsilon}$ , one can form a sequence of  $\tau$  weighted graphs from the input trajectory data of length  $\tau$  (see Section 4.2.2).

By construction the transformation  $T : M \rightarrow N$  induces the graph transformation  $T_G : G(S^k, W^{\mu, \epsilon}) \rightarrow G(\hat{S}^k, W^{\nu, \epsilon})$ . Now as a dynamic extension of manifold learning, we ask for the recovery of low-dimensional features from an unknown dynamical system, by studying the evolving connective structures of  $G(S^k, W^{\mu, \epsilon})$  under  $T_G$ . In particular, we search for a low-dimensional embedding map  $\varphi_{dyn} : S^k \rightarrow \mathbb{R}^s$  ( $s \ll k$ ) using the connective structures of *both*  $G(S^k, W^{\mu, \epsilon})$  and  $G(\hat{S}^k, W^{\nu, \epsilon})$ , such that low dimensional dynamical features of  $T : M \rightarrow N$  can be extract from the collection of points  $\varphi_{dyn}(S^k) \in \mathbb{R}^s$ . Moreover, point-cloud clusters of  $\lim_{k \rightarrow \infty} \varphi_{dyn} S^k$  correspond to submanifolds of  $M$  that have persistently small boundary size relative

to enclosed volume, when the geometry of  $M$  is evolved by  $T$ .

Our approach for finding  $\varphi_{dyn}$  builds on the idea of the Laplacian eigenmaps [10, 11] for static manifold learning, which we now review.

### 1.3.3 Laplacian-based manifold learning

Laplacian-based manifold learning is a popular method for dimensionality reduction, that utilises the spectral properties of the graph Laplacian to approximate low-dimensional geometric structures on  $M$  (see also [28] for the closely related diffusion maps). The spectral graph partitioning method outlined in Section 1.1.3 laid the groundwork for Laplacian-based manifold learning, where the second eigenvector  $\mathbf{g}_2$  of the normalised graph Laplacian  $\mathcal{L}$  (1.6) offered a 1-dimensional representation of a connective structure of an unweighted graph  $G(V, E)$ ; components of  $\mathbf{D}^{-1/2}\mathbf{g}_2$  rank the vertex set  $V$  by magnitude. More generally, the first  $s$  nontrivial eigenvectors  $\mathbf{g}_2, \mathbf{g}_3, \dots, \mathbf{g}_{s+1}$  of  $\mathcal{L}$  (1.5) provide an  $s$ -dimensional embedding of  $V$ , which are used to gain insight into the connective structures of  $G(V, E)$  [112]. Analogously, in [10, 11] the Laplacian eigenmaps are formed as follows: let  $\|\cdot\|_{\mathbb{R}^d}$  denote the  $d$ -dimensional Euclidean norm on  $\mathbb{R}^d$ . Set the edge weights of  $G(S^k, W^{\mu, \epsilon})$  as

$$w_{ij}^{\mu, \epsilon} = w_{ij}^{\epsilon} := \begin{cases} \exp\left(-\frac{\|x_i - x_j\|_{\mathbb{R}^d}^2}{\epsilon}\right) & \text{if } i \neq j \\ 0 & \text{if } i = j \end{cases},$$

for each  $1 \leq i, j \leq k$ . Define the *degree* of the vertex  $x_i \in S^k$  by  $d^\epsilon(x_i) := \sum_{j=1}^k w_{ij}^{\epsilon}$ , and *adjacency matrix*  $\mathbf{A}^\epsilon$  by the  $k \times k$  matrix with entries  $\mathbf{A}_{ij}^\epsilon = w_{ij}^{\epsilon}$ . Then analogous to (1.5), the *graph Laplacian* associated with  $G(S^k, W^{\mu, \epsilon})$  (see e.g. Chapter 2 in [15]) is given by  $\mathbf{L}^\epsilon := \mathbf{A}^\epsilon - \mathbf{D}^\epsilon$ , where  $\mathbf{D}^\epsilon$  is diagonal with entries  $\mathbf{D}_{ii}^\epsilon = d^\epsilon(x_i)$ ; that is, the matrix  $\mathbf{L}^\epsilon$  has entries

$$L_{ij}^\epsilon := \begin{cases} w_{ij}^\epsilon & \text{if } i \neq j \\ -\sum_{l=1}^k w_{il}^\epsilon & \text{if } i = j \end{cases}.$$

In addition, analogous to (1.6) the normalised graph Laplacian for  $G(S^k, W^{\mu, \epsilon})$  is defined by

$$\mathcal{L}^\epsilon := (\mathbf{D}^\epsilon)^{-\frac{1}{2}} \mathbf{L}^\epsilon (\mathbf{D}^\epsilon)^{-\frac{1}{2}}, \quad (1.11)$$

The crux of Laplacian eigenmaps is to use the first  $s$  nontrivial eigenvectors  $\mathbf{g}_2, \mathbf{g}_3, \dots, \mathbf{g}_{s+1}$  of  $\frac{1}{\epsilon}\mathcal{L}^\epsilon$ , to construct an  $s$ -dimensional embedding map  $\varphi(x_i) := \{(\mathbf{D}^\epsilon)^{-\frac{1}{2}}\mathbf{g}_2(x_i), (\mathbf{D}^\epsilon)^{-\frac{1}{2}}\mathbf{g}_3(x_i), \dots, (\mathbf{D}^\epsilon)^{-\frac{1}{2}}\mathbf{g}_{s+1}(x_i)\}$ ,  $1 \leq i \leq k$  that preserves certain low-dimensional geometric structures on  $M$ ; i.e. Low dimensional features of an unknown manifold is ‘learned’ from the new representation of  $\varphi(S^k)$  in  $\mathbb{R}^s$ .

**Example 1.3.2.** For example, applying the classical Laplacian eigenmaps to the random samples in Example 1.3.1. Figure 1.6a and Figure 1.6b are the images of Figure 1.5a and Figure 1.5b under the embedding map  $\varphi = \{(\mathbf{D}^\epsilon)^{-\frac{1}{2}}\mathbf{g}_2, (\mathbf{D}^\epsilon)^{-\frac{1}{2}}\mathbf{g}_3\}$ , where  $\mathbf{g}_2, \mathbf{g}_3$  are the first 2 nontrivial eigenvectors of  $\mathcal{L}^\epsilon$ .



Figure 1.6: Laplacian eigenmap on the random samples in Figure 1.5.

We note that the standard Laplacian-based manifold learning method [10, 11] is designed for analysis on uniformly distributed data. In Figure 1.6a the symmetrical structure of the 2-dimensional torus  $[0, 4] \setminus \sim \times [0, 1] \setminus \sim$  was recovered from a uniformly distributed sample. On the other hand, the nonuniform distribution of the random sample in Figure 1.5b is incorrectly captured as a geometric feature of  $[0, 4] \setminus \sim \times [0, 1] \setminus \sim$  by the standard Laplacian eigenmap  $\varphi$ ; as shown in Figure 1.6b.

### 1.3.4 Original contributions of Chapter 4

Since the Laplacian eigenmap  $\varphi$  does not depend on the density  $h_\mu$  of  $\mu_r$ , it is clear that the standard Laplacian-based manifold learning method described in Section 1.3.3 does not account for the measure  $\mu_r$  on  $M$ . Moreover, according to Theorem 5.2 in [12], the Laplacian eigenmap  $\varphi$  is dependent on the probability distribution of the random sample  $S^k$ . In particular, the Laplacian-based manifold learning method in [10, 11] is limited to applications on uniformly distributed random samples, and it is also limited to approximating the features of unweighted manifolds. In Chapter 4, we extend the work of [10, 11] in three ways: (1) application on input samples with nonuniform probability distributions, (2) approximating features of weighted Riemannian manifolds, and (3) a dynamic manifold learning method for approximating transport barriers from trajectory data.



## Chapter 2

# Partitions of graphs that are robust to vertex permutation dynamics

Spectral graph partitioning algorithms such as the method described in Section 1.1.3 are highly successful for finding highly interconnected subregions of static graphs. Prior work to extend spectral methods to application on dynamics of graphs includes the temporal network approach [42, 92], which encodes the additional dimension of time by making a copy of the graph at each time instance. A common technique for analysing evolving community structures on a temporal network is to reverse the time-extension by projecting the multiple temporal copies of the graph back to a single graph. However, such a technique involves making decisions on how to collapse vertices across time between one or more slices of a temporal network, which can produce an associated loss of temporal information. Extensive work has been done to improve the time-projection approach, e.g. [74] designed a dynamic quality function for community detection based on a time-projected temporal network. Changes in graph structures have also been considered by treating the changing graph as a series of static ones and evaluating the connective structures in each step (see [27, 39] for applications in image processing and social graphs, respectively). However, considering a series of static connectivity problems can only be used to find graph subregions that are community structures at a particular moment in time, rather than graph subregions that are persistently highly interconnected as the graph evolves.

In this chapter, we consider a *dynamic balanced graph bisection problem* designed for finding persistently highly interconnected subgraphs of dynamic graphs.

In particular, let  $G = G(V, E)$  be a connected graph<sup>1</sup> with  $V = \{v_1, v_2, \dots, v_k\}$ , and  $E = \{[v_i, v_j]\}_{i,j=1}^k$ . Let  $\pi = (\pi_v, \pi_e)$  be a graph isomorphism on  $G$ , where  $\pi_v : V \rightarrow V$  is a vertex label permutation and  $\pi_e : E \rightarrow \hat{E}$  satisfies  $\pi_e([v_i, v_j]) = [\pi_v(v_i), \pi_v(v_j)]$  for all  $[v_i, v_j] \in E$ . Under a single iteration of vertex permutation dynamics  $\pi_v : V \rightarrow V$  that transforms  $G$  into  $\pi(G)$ , we consider the dynamic optimisation problem:

$$\mathbf{H}_G^{dyn} := \min_{V_1, V_2 \text{ partition } V} H_G^{dyn}(V_1, V_2), \quad (2.1)$$

where

$$H_G^{dyn}(V_1, V_2) = \frac{|C(V_1, V_2)| + |C(\pi_v^{-1}(V_1), \pi_v^{-1}(V_2))|}{\min\{D(V_1), D(V_2)\} + \min\{D(\pi_v^{-1}(V_1)), D(\pi_v^{-1}(V_2))\}}, \quad (2.2)$$

where  $C(V_1, V_2)$  denotes the set of edges removed to disconnect  $G$ , and  $D(V_1)$  is the degree sum of all vertices in  $V_1$ . We show later in Section 2.1, that the quantities  $|C(\pi_v^{-1}(V_1), \pi_v^{-1}(V_2))|$ ,  $D(\pi_v^{-1}(V_1))$  and  $D(\pi_v^{-1}(V_2))$  appearing on the RHS of (2.2), are in fact related to the graph Cheeger ratio  $H_{\pi(G)}(V_1, V_2)$  (1.2) on the graph  $\pi(G)$ . Alternatives to (2.2) are: (1) minimising the sum of the Cheeger ratios of a fixed partition  $\{V_1, V_2\}$  on  $G$  and  $\pi(G)$ ; that is, find  $\{V_1, V_2\}$  such that  $H_G(V_1, V_2) + H_{\pi(G)}(V_1, V_2)$  is minimal, or (2) minimising the Cheeger ratio on the graph formed by ‘‘averaging’’ the edges of  $G$  and  $\pi(G)$ ; see (2.11) for details. To produce a good solution to (2.1), we develop a spectral method for dynamic graphs in Section 2.2; the proofs of the stated theorems are deferred to Section A of the appendix. A multiple time-step version of the dynamic balanced graph bisection problem is considered in Section 2.3. In Section 2.4, a numerical algorithm for dynamic spectral partitioning is applied to both structured and unstructured graphs, and the strength of the evolving community structures of the partitioned graphs are investigated.

## 2.1 A graph Cheeger constant for dynamic graphs

Note that the set of edges removed to disconnect  $G$  via the partition  $\{V_1, V_2\}$  is  $C(V_1, V_2) := \{[v_i, v_j] \in E : v_i \in V_1, v_j \in V_2\}$ . To describe the disconnection of the graph  $\pi(G)$  induced by  $\{V_1, V_2\}$ , we denote the reduced set of edges  $\pi_e(E)' = \pi_e(E) \setminus \{[v_i, v_j] \in \pi_e(E) : v_i \in V_1, v_j \in V_2\}$ . Let  $C_\pi$  denote the set of edges removed to disconnect  $\pi(G)$ ; that is

$$C_\pi(V_1, V_2) := \{[v_i, v_j] \in \pi_e(E) : v_i \in V_1, v_j \in V_2\}; \quad (2.3)$$

---

<sup>1</sup>For simplicity we assume there are no self-loops or multiple edges on  $G$ ; although the method we describe could be extended to cover these cases

in words, fix  $V_1$  and  $V_2$  and compare edges in  $\pi(G)$ . Equivalently,

$$\begin{aligned} |C_\pi(V_1, V_2)| &= |\{[\pi_v^{-1}(v_i), \pi_v^{-1}(v_j)] \in E : v_i \in V_1, v_j \in V_2\}| \\ &= |\{[v_i, v_j] \in E : v_i \in \pi_v^{-1}(V_1), v_j \in \pi_v^{-1}(V_2)\}|; \end{aligned} \quad (2.4)$$

that is, pullback the vertex sets  $V_1, V_2$  with  $\pi_v$  and compare edges in  $G$ . Thus,

$$|C_\pi(V_1, V_2)| = |C(\pi_v^{-1}(V_1), \pi_v^{-1}(V_2))|. \quad (2.5)$$

We now consider the computation of vertex degree in  $\pi(G)$ . For  $V' \subset V$ , define  $D_\pi(V') := \sum_{i \in V'} d_\pi(v_i)$ , where  $d_\pi(v_i)$  is the degree of  $v_i$  computed in the graph  $\pi(G)$ :

$$d_\pi(v_i) := |\{v_j \in V : [v_i, v_j] \in \pi_e(E)\}|. \quad (2.6)$$

Let  $d(v_i)$  denote the degree of the vertex  $v_i$  of the graph  $G$ . One can also do this degree computation in the original graph  $G$  by noticing that

$$\begin{aligned} |\{v_j \in V : [v_i, v_j] \in \pi_e(E)\}| &= |\{v_j \in V : [\pi_v^{-1}(v_i), \pi_v^{-1}(v_j)] \in E\}| \\ &= d(\pi_v^{-1}(v_i)). \end{aligned} \quad (2.7)$$

Thus,

$$D_\pi(V') = D(\pi_v^{-1}(V')). \quad (2.8)$$

**Example 2.1.1.** In Figure 2.1, setting  $V_1 = \{v_3, v_5\}$ ,  $V_2 = \{v_1, v_2, v_4\}$  and applying the cyclic permutation  $\pi_v(v_i) = v_{i+1}$ ,  $i = 1, 2, 3, 4, 5$  as in Example 1.1.2. We have:  $|C_\pi(V_1, V_2)| = |\{[v_3, v_2], [v_3, v_4], [v_5, v_1]\}| = 3$  and  $|C(\pi_v^{-1}(V_1), \pi_v^{-1}(V_2))| = |\{[v_2, v_1], [v_2, v_3], [v_4, v_5]\}| = 3$  (using  $\pi_v^{-1}(V_1) = \{v_2, v_4\}$ ,  $\pi_v^{-1}(V_2) = \{v_5, v_1, v_3\}$ ). Also,  $D_\pi(V_1) = 5 = D(\pi_v^{-1}(V_1))$  and  $D_\pi(V_2) = 3 = D(\pi_v^{-1}(V_2))$ .



Figure 2.1: Cyclic permutation on a graph with 5 vertices, colored vertices =  $V_1$ , colored edges =  $E \setminus E'$ .

Due to (2.5) and (2.8), the graph Cheeger ratio  $H_{\pi(G)}$  (1.2) on the partition  $\{V_1, V_2\}$  for  $\pi(G)$  can be expressed as

$$H_{\pi(G)}(V_1, V_2) := \frac{|C_\pi(V_1, V_2)|}{\min\{D_\pi(V_1), D_\pi(V_2)\}} = \frac{|C(\pi_v^{-1}(V_1), \pi_v^{-1}(V_2))|}{\min\{D(\pi_v^{-1}(V_1)), D(\pi_v^{-1}(V_2))\}}. \quad (2.9)$$

Moreover, (2.2) can be written as

$$H_G^{dyn}(V_1, V_2) = \frac{|C(V_1, V_2)| + |C_\pi(V_1, V_2)|}{\min\{D(V_1), D(V_2)\} + \min\{D_\pi(V_1), D_\pi(V_2)\}}. \quad (2.10)$$

Thus, the optimisation problem (2.1) is linked to the magnitudes of  $H_G(V_1, V_2)$  and  $H_{\pi(G)}(V_1, V_2)$  on a fixed partition  $\{V_1, V_2\}$ .

**Example 2.1.2.** *Returning to Example 2.1.1, we compute*

$$H_G^{dyn}(V_1, V_2) = \frac{2 + 3}{\min\{2, 6\} + \min\{5, 3\}} = 1.$$

If choose  $V'_1 = \{v_1, v_4, v_5\}$ ,  $V'_2 = \{v_2, v_3\}$ , we find

$$H_G^{dyn}(V'_1, V'_2) = \frac{2 + 2}{\min\{1 + 2 + 1, 3 + 1\} + \min\{1 + 1 + 2, 1 + 3\}} = \frac{1}{2};$$

in fact, this is the unique partition achieving this minimal value  $\mathbf{H}_G^{dyn}$ .

*Remark 2.1.3.* One could alternatively define a *dynamic balanced graph bisection problem*:

$$\hat{\mathbf{H}}_G^{dyn} := \min_{V_1, V_2 \text{ partition } V} \hat{H}_G^{dyn}(V_1, V_2), \quad (2.11)$$

where

$$\hat{H}_G^{dyn}(V_1, V_2) = \frac{|C(V_1, V_2)| + |C_\pi(V_1, V_2)|}{\min\{D(V_1) + D_\pi(V_1), D(V_2) + D_\pi(V_2)\}}. \quad (2.12)$$

Such a definition places less emphasis on producing a balanced partition both before and after the application of  $\pi$  because low degrees sums can be “averaged away”.

Clearly,  $\hat{\mathbf{H}}_G^{dyn} \leq \mathbf{H}_G^{dyn}$ , however, in general, one cannot conclude that the connectivity of a partition optimizing (2.12) is greater or less than a partition optimizing (2.1).

Thus, in our numerical experiments we report results for both quantities.

**Example 2.1.4.** *Returning to Example 2.1.1, we compute*

$$\hat{H}_G^{dyn}(V_1, V_2) = \frac{2 + 3}{\min\{2 + 5, 6 + 3\}} = \frac{5}{7}.$$

*Because the partition  $V = V_1 \cup V_2$  is highly unbalanced in vertex degrees on both  $G$  and  $\pi(G)$ , but  $|D(V_1) - D_\pi(V_2)| = 1$  and  $|D(V_2) - D_\pi(V_1)| = 1$ . The ratio  $H_G^{dyn}(V_1, V_2) = 1$  incurs a larger penalty compare to  $\hat{H}_G^{dyn}(V_1, V_2)$  due to degree unbalance.*

## 2.2 A spectral method for dynamic graphs

We now introduce a dynamic graph Laplacian to provide good solutions to the dynamic balanced graph bisection problem (2.1). Define the square permutation matrix

$$P_{ij} := \begin{cases} 1 & \text{if } \pi_v(v_i) = v_j \\ 0 & \text{otherwise} \end{cases}. \quad (2.13)$$

Note that the matrix  $\mathbf{P}$  is obtained by permuting the rows of the identity matrix, hence  $\mathbf{P}$  is invertible. Motivated by the properties (2.5) and (2.8), we define the *dynamic graph Laplacian*

$$\mathbf{L}^{dyn} := \frac{\mathbf{L} + \mathbf{P}^{-1}\mathbf{L}\mathbf{P}}{2}, \quad (2.14)$$

where  $\mathbf{L}$  is as in (1.5). The first term in (2.14) acts on  $G$ , while the second term transforms from  $G$  to  $\pi(G)$  using  $\mathbf{P}$ , then applies  $\mathbf{L}$  to  $\pi(G)$ , and finally pulls the result back to  $G$  with  $\mathbf{P}^{-1}$ . If one defines  $\mathbf{L}^\pi$  to be the Laplacian matrix for the graph  $\pi(G)$ , then by (2.7),

$$\mathbf{L}_{ij}^\pi := \begin{cases} d(\pi_v^{-1}(v_i)) & \text{if } i = j \\ -1 & \text{if } [\pi_v^{-1}(v_i), \pi_v^{-1}(v_j)] \in E, i \neq j \\ 0 & \text{otherwise} \end{cases} \quad (2.15)$$

$$= \mathbf{L}_{\pi_p^{-1}(i)\pi_p^{-1}(j)}, \quad (2.16)$$

where  $\pi_p$  is the vertex label permutation associated with  $\pi_v$ ; i.e.  $v_{\pi_p(i)} = \pi_v(v_i)$ . From the definition of  $\mathbf{P}$ , it is straightforward to show that  $\mathbf{L}^\pi = \mathbf{P}^{-1}\mathbf{L}\mathbf{P}$ . Let  $\mathbf{D}$  be  $k \times k$  diagonal with entries

$$\mathbf{D}_{ii} = d(v_i). \quad (2.17)$$

We apply a degree normalisation to  $\mathbf{L}^{dyn}$  and define the *normalised dynamic graph Laplacian* by

$$\mathcal{L}^{dyn} := \mathbf{D}^{-1/2}\mathbf{L}^{dyn}\mathbf{D}^{-1/2} = \frac{\mathcal{L} + \mathcal{L}^\pi}{2}, \quad (2.18)$$

where  $\mathcal{L}^\pi = \mathbf{D}^{-1/2}\mathbf{L}^\pi\mathbf{D}^{-1/2}$ .

*Remark 2.2.1.* The connection between the above normalised dynamic graph Laplacian  $\mathcal{L}^{dyn}$  and the directed graph Laplacian studied in [26] is as follows: On a directed graph, the dynamics of vertex transformation is captured by directed edges as a transition probability matrix. Therefore, the static isoperimetry problem on a directed graph concerns clusters of vertices that are more likely to be transferred within a cluster than between clusters. While Chung [26] can only see almost-invariant structures, there is no true time-dependence in [26], in contrast the dynamic Laplacian allows the coherent structures to be time-dependent.

**Theorem 2.2.2.** *Let  $G = G(V, E)$  be a simple, connected graph with  $|V| = k$ . Let  $\pi = (\pi_v, \pi_e)$  be a graph isomorphism as in Section 1.1, and denote by  $\pi_p$  the vertex label permutation associated with  $\pi_v$  as above. Define  $\mathbf{P}$  and  $\mathcal{L}^{dyn}$  as in (2.13) and (2.18), respectively. One has*

1. *The eigenvalues  $\lambda_1 \geq \lambda_2 \geq \dots$  of  $\mathcal{L}^{dyn}$  are nonpositive and real.*

2. The eigenvalue  $\lambda_1 = 0$ , and is of unit multiplicity.
3. Let  $\mathbf{D}$  be the degree matrix of  $G$  given by (2.17), and  $\mathbf{g}_1$  an eigenvector corresponding to  $\lambda_1$ . Then  $\mathbf{D}^{-1/2}\mathbf{g}_1$  is constant on a connected component<sup>2</sup> of  $G$ .
4. Let  $\mathbf{1} \in \mathbb{R}^k$  be a constant vector with unit components, and let  $\langle \cdot, \cdot \rangle$  be the inner-product on  $\mathbb{R}^k$ . Denote by  $\sum_{i \sim j}$  the ordered summation over all pairs of vertices such that  $[v_i, v_j] \in E$ . One has

$$\lambda_2 = - \min_{\mathbf{f} \in \mathbb{R}^k: \langle \mathbf{f}, \mathbf{D} \cdot \mathbf{1} \rangle = 0} \frac{\sum_{i \sim j} (f_i - f_j)^2 + (f_{\pi_p(i)} - f_{\pi_p(j)})^2}{\sum_{i=1}^k d(v_i) f_i^2}. \quad (2.19)$$

The minimum of (2.19) is attained when  $\mathbf{f} = \mathbf{D}^{-1/2}\mathbf{g}_2$ , where  $\mathbf{g}_2$  is the eigenvector of  $\mathcal{L}^{dyn}$  corresponding to  $\lambda_2$ .

*Proof.* See Appendix A. □

The connection between  $\mathcal{L}^{dyn}$  and  $\mathbf{H}_G^{dyn}$  is given by the following theorem.

**Theorem 2.2.3 (Dynamic graph Cheeger inequality).** *Let  $G = G(V, E)$  be a simple, connected graph, and  $\mathcal{L}^{dyn}$ ,  $\mathbf{H}_G^{dyn}$ ,  $\hat{\mathbf{H}}_G^{dyn}$  be defined by (2.18), (2.1), (2.11) respectively. If  $\lambda_2$  is the second smallest eigenvalue (by magnitude) of  $\mathcal{L}^{dyn}$ , then*

1.

$$\hat{\mathbf{H}}_G^{dyn} \leq \sqrt{-2\lambda_2}. \quad (2.20)$$

2. If in addition,  $d(v_i) \leq D(V)/4$  for each  $i$ , then

$$\mathbf{H}_G^{dyn} \leq 2\sqrt{-\lambda_2}. \quad (2.21)$$

*Proof.* The proof of part 1 is simple to deduce from (1.7) by noting that  $\mathcal{L}^{dyn}$  is the normalised Laplacian matrix for the edge-weighted graph  $G(V, E \cup \pi_e(E))$ , where each edge contributes 1/2 in the partition boundaries  $C(V_1, V_2)$  and  $C_\pi(V_1, V_2)$ . The stronger degree balancing in the definition of  $\mathbf{H}_G^{dyn}$  (part 2) requires a more detailed proof, which is elaborated in Appendix A. □

*Remark 2.2.4.* The additional condition for Theorem 2.2.3(2) can be made more precise, and only requires two particular vertices to have degree less than or equal to  $D(V)/4$ . Specifically, let  $\mathbf{g}_2 = \mathbf{D}^{1/2}\mathbf{f}$  be the eigenfunction of  $\mathcal{L}^{dyn}$  corresponding to  $\lambda_2$ , and order the vertices of  $G$  according to  $\mathbf{f}$  by  $\mathbf{f}_i \leq \mathbf{f}_{i+1}$ . Define  $S^i =$

<sup>2</sup>If  $V'$  is a connected component of  $G$ , then either  $(\mathbf{D}^{-1/2}\mathbf{g}_1)_i = 1$  or  $(\mathbf{D}^{-1/2}\mathbf{g}_1)_i = 0$  for all  $v_i \in V'$ .

$\{v_1, v_2, \dots, v_i\}$ , and let  $r$  and  $q$  denote the largest integer such that  $D(\pi_v^{-1}(S^{r-1})) < D(V)/2$  and  $D(\pi_v^{-1}(S^q)) < D(V)/2$ . It is sufficient for the degree of  $v_{r-1}$  and  $v_q$  to have degree less than or equal to  $D(V)/4$ .

The Laplacian matrix for a given graph is constructed from the graph's adjacency matrix; all information regarding the graph's connectivity is encoded within the graph's Laplacian. Since our new dynamic graph Laplacian was constructed from both the Laplacian for  $G$  and  $\pi(G)$ , it is possible that the complex interactions between dynamics and graph connectivity are contained within the dynamic graph Laplacian. Indeed, Theorem 2.2.3 tell us that the subregions on a dynamic graph that are persistently highly interconnected are closely related to the second smallest magnitude eigenvalue of  $\mathcal{L}^{dyn}$ . In fact, in the proof of Theorem 2.2.3(1), it is shown that the eigenvector corresponding to  $\lambda_2$  indicates how the graph of interest should be partitioned. In particular, if the vertices of  $G$  are ordered according to the magnitude of each component of the degree normalised eigenvector of  $\lambda_2$ , then there exists a threshold in which the partition elements yielded would have a dynamic graph Cheeger constant that satisfies the inequality (2.21).

Although, vertex permutation dynamics on a graph is equivalent to a specific sequence of edge addition/deletion on a temporal network (for example, Figure 2.1 (a) and (b) can be represented in a manner similar to Figure 1 in [92]), there is a key difference between the temporal network approach of [42, 92] and the current dynamic spectral method for community structure detection on time-dependent graphs: The multislice extension of (6) in [82] by (1) in [92] does not follow the scheme of pushing forward, evolving, then pulling back of our normalised dynamic graph Laplacian in (2.18), because time-dependence of (1) in [92] is invoked by the coupling constants  $C_{jrs}$ .

One could also use the results Theory 2.2.3 to formulate a surrogate network generation algorithm. In particular, since the size of the second smallest magnitude eigenvalue of  $\mathcal{L}^{dyn}$  is an indication of the strength of evolving community structure of the dynamic graphs, a process based of the dynamic graph Laplacian can be used to generate random graphs of a fixed size with prescribed community structures (see [79, 111] for details of random graph generation).

## 2.3 Dynamics over $\tau$ time steps

If one has  $\tau - 1$  permutations  $\pi_1, \dots, \pi_{\tau-1}$ , which are applied in sequence to the graph, then one can naturally extend (2.1)-(2.10) to form dynamic graph Cheeger constants  $\mathbf{H}_G^\tau$  and  $\hat{\mathbf{H}}_G^\tau$  over  $\tau$  time steps. Denote  $\pi^{(0)} = \text{Id}$ ,  $\pi^{(t)} = \pi_t \circ \dots \circ \pi_2 \circ \pi_1$ ,

$t = 1, \dots, \tau - 1$ , and define

$$\mathbf{H}_G^\tau := \min_{V_1, V_2 \text{ partition } V} H_G^\tau(V_1, V_2), \quad (2.22)$$

where

$$H_G^\tau(V_1, V_2) = \frac{\sum_{t=0}^{\tau-1} |C_{\pi^{(t)}}(V_1, V_2)|}{\sum_{t=0}^{\tau-1} \min\{D_{\pi^{(t)}}(V_1), D_{\pi^{(t)}}(V_2)\}}. \quad (2.23)$$

Similarly, define

$$\hat{\mathbf{H}}_G^\tau := \min_{V_1, V_2 \text{ partition } V} \hat{H}_G^\tau(V_1, V_2), \quad (2.24)$$

where

$$\hat{H}_G^\tau(V_1, V_2) = \frac{\sum_{t=0}^{\tau-1} |C_{\pi^{(t)}}(V_1, V_2)|}{\min\{\sum_{t=0}^{\tau-1} D_{\pi^{(t)}}(V_1), \sum_{t=0}^{\tau-1} D_{\pi^{(t)}}(V_2)\}}. \quad (2.25)$$

To construct the  $\tau$ -time step dynamic Laplacian, denote by  $\mathbf{P}_t$  the permutation matrix for  $\pi_t$  (according to (2.13)) with  $\mathbf{P}_0 = \text{Id}$ , and define

$$\mathbf{L}^\tau = \frac{\sum_{t=0}^{\tau-1} (\mathbf{P}_t \cdots \mathbf{P}_2 \mathbf{P}_1)^{-1} \mathbf{L} (\mathbf{P}_t \cdots \mathbf{P}_2 \mathbf{P}_1)}{\tau}. \quad (2.26)$$

The  $\tau$ -time step normalised dynamic Laplacian can be found by

$$\mathcal{L}^\tau = \mathbf{D}^{-1/2} \mathbf{L}^\tau \mathbf{D}^{-1/2}, \quad (2.27)$$

where  $\mathbf{D}$  is the degree matrix for the initial graph  $G(V, E)$ . One has the following trivial extension of Theorem 2.2.3(1):

**Theorem 2.3.1.** *Let  $G = G(V, E)$  be a simple, connected graph. Define  $\hat{\mathbf{H}}_G^\tau$  and  $\mathcal{L}^\tau$  as in (2.24) and (2.27), respectively. If  $\lambda_2$  is the second largest (by magnitude) eigenvalue of  $\mathcal{L}^\tau$ , then*

$$\hat{\mathbf{H}}_G^\tau \leq \sqrt{-2\lambda_2}. \quad (2.28)$$

The results of Theorem 2.2.2 (see appendix) also hold in this  $\tau$ -step situation, however, we do not present a  $\tau$ -time step version of Theorem 2.2.3(2).

*Remark 2.3.2.* The expressions (2.24) and (2.25) calculate the quality of the cut on  $\pi^{(t)}(G)$  after each permutation  $\pi^{(t)}$ . If one is interested in a sequence of permutations  $\pi_1, \dots, \pi_{\tau-1}$  but only cares about the quality of the cut at time 0 and at time  $\tau - 1$ , then one would instead consider  $\pi^{(\tau-1)}$  as a single permutation and use (2.1)-(2.10) or (2.11)-(2.12) instead.

## 2.4 Numerical method and experiments

We can use the new Laplacian matrix  $\mathcal{L}^{dyn}$  to construct bisections of  $G$  that are robust to a single (resp. multiple) permutations. The algorithm to partition  $G$

**Algorithm 2.1:** Dynamic spectral partition for graphs

- 1 Let  $G$  be a connected graph, and  $\pi$  a graph isomorphism on  $G$ . Form the matrix  $\mathcal{L}^{dyn}$  according to (2.18).
- 2 Solve the matrix eigenvalue problem  $\mathcal{L}^{dyn} \mathbf{g}_2 = \lambda_2 \mathbf{g}_2$ , where  $\lambda_2$  is the first non-trivial eigenvalue of  $\mathcal{L}^{dyn}$ , with corresponding eigenvector  $\mathbf{g}_2$ .
- 3 Let  $\mathbf{f} = \mathbf{D}^{-1/2} \mathbf{g}_2$ , where  $\mathbf{D}$  is the degree matrix of  $G$ . Then for each  $\beta \in \{f_i\}_{i=1}^k$ , partition the vertex set  $V$  of  $G$  into  $V_1^\beta = \{v_i \in V : f_i \leq \beta\}$  and  $V_2^\beta = \{v_i \in V : f_i > \beta\}$ ; there are at most  $k - 1$  nontrivial partition of this form.
- 4 Output the partitions  $V = V_1^\beta \cup V_2^\beta$  that minimises  $H_G^{dyn}(V_1^\beta, V_2^\beta)$  or  $\hat{H}_G^{dyn}(V_1^\beta, V_2^\beta)$ .

under a single iteration of dynamics is a minor modification of the construction of bisections of a static graph in Section 1.1.3, which we outline in Algorithm (2.1)

For dynamic spectral partition over  $\tau$ -time steps, one applies the following modifications to Algorithm 2.1

1. In step 1 and 2 of Algorithm 2.1, replace  $\mathcal{L}^{dyn}$  with  $\mathcal{L}^\tau$  (2.27).
2. In step 4 of Algorithm 2.1, replace  $H_G^{dyn}(V_1^\beta, V_2^\beta)$  and  $\hat{H}_G^{dyn}(V_1^\beta, V_2^\beta)$  with  $H_G^\tau(V_1^\beta, V_2^\beta)$  (2.23) and  $\hat{H}_G^\tau(V_1^\beta, V_2^\beta)$  (2.25), respectively.

To illustrate our method, we apply vertex permutation dynamics to two graphs with very different connective structures. Firstly, a graph with obvious static community structures, where we apply dynamics to disrupt these community structures. Secondly, we consider dynamics to a large randomly generated graph, where there are no clear static community structures, nor dynamic community structures. In both cases, we search for community structures that are robust to the given vertex permutation dynamics. We note that Algorithm 2.1 can be applied to large graphs with thousands of vertices, because only the first nontrivial eigenvector of  $\mathcal{L}^{dyn}$  is required to generate the partitions  $\{V_1^\beta, V_2^\beta\}$  for some  $\beta \in \mathbb{R}$ , and at most  $k - 2$  comparisons are made between the partitions  $\{V_1^\beta, V_2^\beta\}$  to find a  $\beta$  that minimises  $H_G(V_1^\beta, V_2^\beta)$ .

### 2.4.1 Example 1: A structured graph

Let  $G$  be the 3-regular Ellingham-Horton 54-graph; see Figure 2.2.

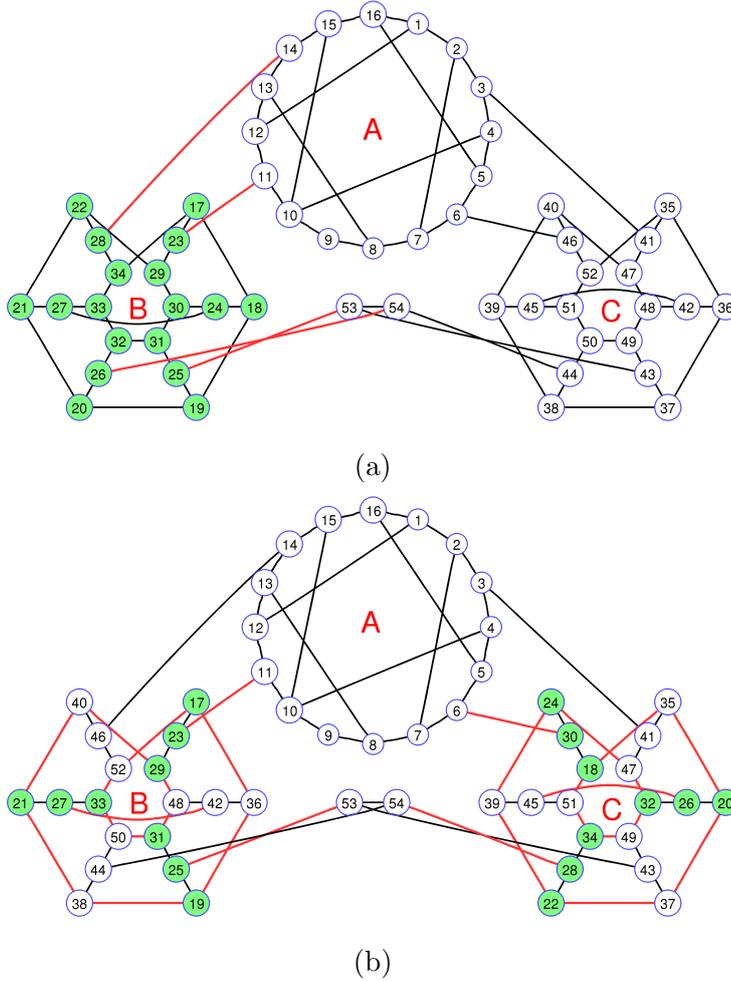


Figure 2.2: Ellingham-Horton 54-graph. Obvious static community structures are labelled “A”, “B”, “C”. Shown is the result of the spectral bisection method described in Section 1.1.3 using  $\mathcal{L}$ . The resulting partition is shown as  $V_1$  (colored vertices),  $V_2$  (non-colored vertices) and the partition boundary (red edges). (a)  $G$ :  $|C(V_1, V_2)| = 4$ ,  $D(V_1) = 54$ ,  $D(V_2) = 108$ . (b)  $\pi(G)$ :  $|C_\pi(V_1, V_2)| = 30$ ,  $D_\pi(V_1) = 54$ ,  $D_\pi(V_2) = 108$ .

### The standard (static) spectral bisection method

We first attempt to solve the static balanced bisection problem using the second eigenvector  $\mathbf{g}_2$  of the Laplacian matrix  $\mathcal{L}$  as described in Section 1.1.3. The vector  $\mathbf{f} = \mathbf{D}^{-1/2}\mathbf{g}_2$  (shown in Figure 2.3a) orders the vertices and produces at most  $k - 1$  distinct partitions of the form  $\{V_1^\beta, V_2^\beta\}$ ; we select the partition with the lowest value of  $H_G(V_1^\beta, V_2^\beta)$  given by (1.2). The results are shown in Figure 2.2a, with the vertices corresponding to  $V_1$  colored green and those in  $V_2$  uncolored. The edges that connect  $V_1$  and  $V_2$  are colored red. The corresponding numerical quantities are in the “ $\mathcal{L}$ ” column of Table 2.1. The degree counts of  $V_1$  and  $V_2$  are relatively unbalanced; this

is because the graph consists of three main clusters of approximately equal degree sum, and it is natural to statically partition the graph by grouping two clusters together. In practice, many graphs may not have a natural community structure. For example, consider a complete graph  $G$  of any degrees. There can be at most  $k - 1$  unique values of  $H_G(V_1, V_2)$  for any  $\{V_1, V_2\}$  that partitions  $G$ , with each unique value of  $H_G(V_1, V_2)$  corresponding to  $|V_1| = 1, 2, \dots, k - 1$ . Moreover, since  $|V_1^\beta| = 1, 2, \dots, k - 1$  for some  $\beta \in \mathbb{R}^k$ , the partition  $\{V_1^\beta, V_2^\beta\}$  is optimal for  $\mathbf{H}_G$  some  $\beta \in \mathbb{R}^k$ .

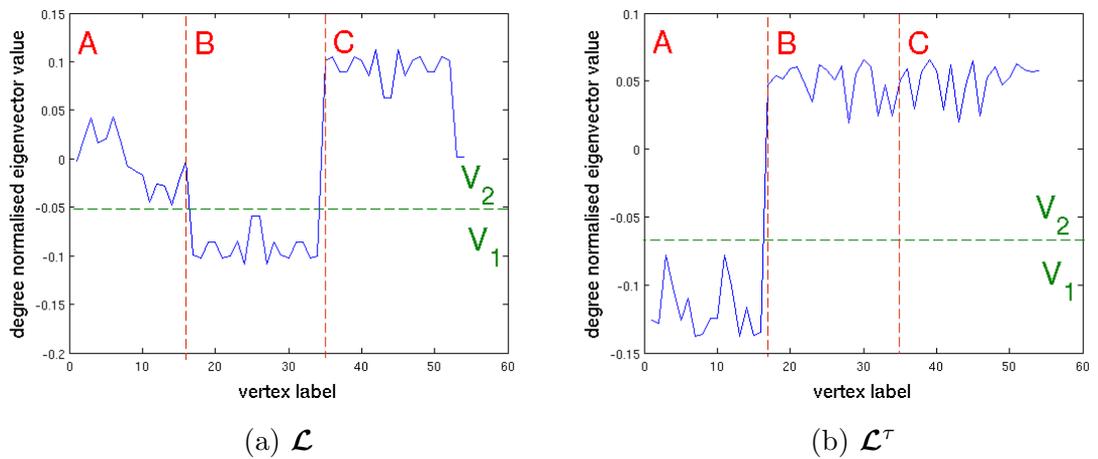


Figure 2.3: Plots of  $\mathbf{f} = \mathbf{D}^{-1/2}\mathbf{g}_2$ , where  $\mathbf{g}_2$  is the second eigenvector of either  $\mathcal{L}$  or  $\mathcal{L}^{dyn}$  on the Ellingham-Horton 54-graph. (a)  $\mathbf{f}$  from the static Laplacian  $\mathcal{L}$ . (b)  $\mathbf{f}$  from the dynamic Laplacian  $\mathcal{L}^{dyn}$ . The letters “A”, “B”, “C” and the vertical red lines refer to the static community structures labelled in Figure 2.2. The horizontal green line indicates the optimal value of  $\beta$  used for selecting  $V_1, V_2$ .

We now introduce a vertex permutation  $\pi_v : V \rightarrow V$ , which will disrupt the cluster structure. The particular permutation we apply to the vertex labels is the cyclic permutation  $(18, 36, 18 + 2, 36 + 2, 18 + 4, 36 + 4, 18 + 6, 36 + 6, \dots, 18 + 16, 36 + 16)$ . The vertex collections  $V_1$  and  $V_2$  in  $\pi(G)$  are shown in Figure 2.2b, colored green and white, respectively. The edges in  $\pi_e(E)$  that connect  $V_1$  and  $V_2$  are colored red, and one now sees a large increase in the number of these edges. Thus, the partition  $V_1, V_2$ , which nicely captured the cluster structure of the static graph, is not robust under the permutation  $\pi$ ; in other words,  $V_1, V_2$  do not capture community structures for both  $G$  and  $\pi(G)$ . The relevant numerical quantities are listed in the “ $\mathcal{L}$ ” column of Table 2.1. One sees a large increase in  $H_{\pi(G)}(V_1, V_2)$  compared to the value of  $H_G(V_1, V_2)$ .

## New dynamic spectral bisection method

We now seek to determine community structures that are robust under the permutation  $\pi$ . To do this, we form the matrix  $\mathcal{L}^\tau$  and compute the second eigenvector  $\mathbf{g}_2$ . The vector  $\mathbf{f} = \mathbf{D}^{-1/2}\mathbf{g}_2$  (shown in Figure 2.3b) orders the vertices and produces at most  $n - 1$  distinct partitions of the form  $\{V_1^\beta, V_2^\beta\}$  as described earlier in Section 2.2; we select the partition with the lowest value of  $H_G^{\text{dyn}}(V_1^\beta, V_2^\beta)$  given by (2.10). Because the degree of each vertex is 3, one has  $H_G^{\text{dyn}}(V_1^\beta, V_2^\beta) = \hat{H}_G^{\text{dyn}}(V_1^\beta, V_2^\beta)$  for all partitions  $V_1^\beta, V_2^\beta$ , thus we report only the former quantity.

The results are shown in Figure 2.4a and Figure 2.4b with the vertices corresponding to  $V_1$  colored green and those in  $V_2$  uncolored. The edges that connect  $V_1$  and  $V_2$  are colored red. In contrast to the partition in Figure 2.2, there are relatively few red edges in *both* Figure 2.4a *and* Figure 2.4b. The corresponding numerical quantities are in the “ $\mathcal{L}^{\text{dyn}}$ ” column of Table 2.1.

Quantity	$\mathcal{L}$	$\mathcal{L}^{\text{dyn}}$
$ C(V_1, V_2) $	4	4
$ C_\pi(V_1, V_2) $	30	4
$D(V_1), D(V_2)$	54, 108	114, 48
$D_\pi(V_1), D_\pi(V_2)$	54, 108	114, 48
$H_G(V_1, V_2)$	0.0741	0.0833
$H_{\pi(G)}(V_1, V_2)$	0.3148	0.0833

Table 2.1: Results of spectral bisection using the second eigenvectors of  $\mathcal{L}$  and  $\mathcal{L}^{\text{dyn}}$  for the Ellingham-Horton 54 graph. The column headed “ $\mathcal{L}$ ” contains evaluations using the partition  $\{V_1, V_2\}$  that minimises  $H_G(V_1^\beta, V_2^\beta)$ . The column headed “ $\mathcal{L}^{\text{dyn}}$ ” contains evaluations using the partition  $\{V_1, V_2\}$  that minimises  $H_G^{\text{dyn}}(V_1^\beta, V_2^\beta)$ . The partitions  $V_1, V_2$  are obtained using the methods described in Section 1.1.3 and Algorithm 2.1.

The value of  $H_G(V_1, V_2)$  produced via  $\mathcal{L}^{\text{dyn}}$  is slightly larger than that produced by  $\mathcal{L}$  (0.0833 vs. 0.0741), as the latter is tailored to minimising  $\mathbf{H}_G^{\text{dyn}}$ , however, the value of  $H_{\pi(G)}(V_1, V_2)$  produced by  $\mathcal{L}^{\text{dyn}}$  is much lower than that via  $\mathcal{L}$  (0.0833 vs. 0.3148). Note that the partition found by the degree normalised eigenvector  $\mathbf{f}$  in Figure 2.3a *cannot be found* as a partition from  $\mathbf{f}$  in Figure 2.3b because  $\mathbf{f}$  arising from the latter vector assigns extreme negative and positive values to the clusters B and C in Figure 2.2. Thus, the static Laplacian  $\mathcal{L}$  will not group together clusters B and C and prefers to adjoin cluster A to cluster B or C.

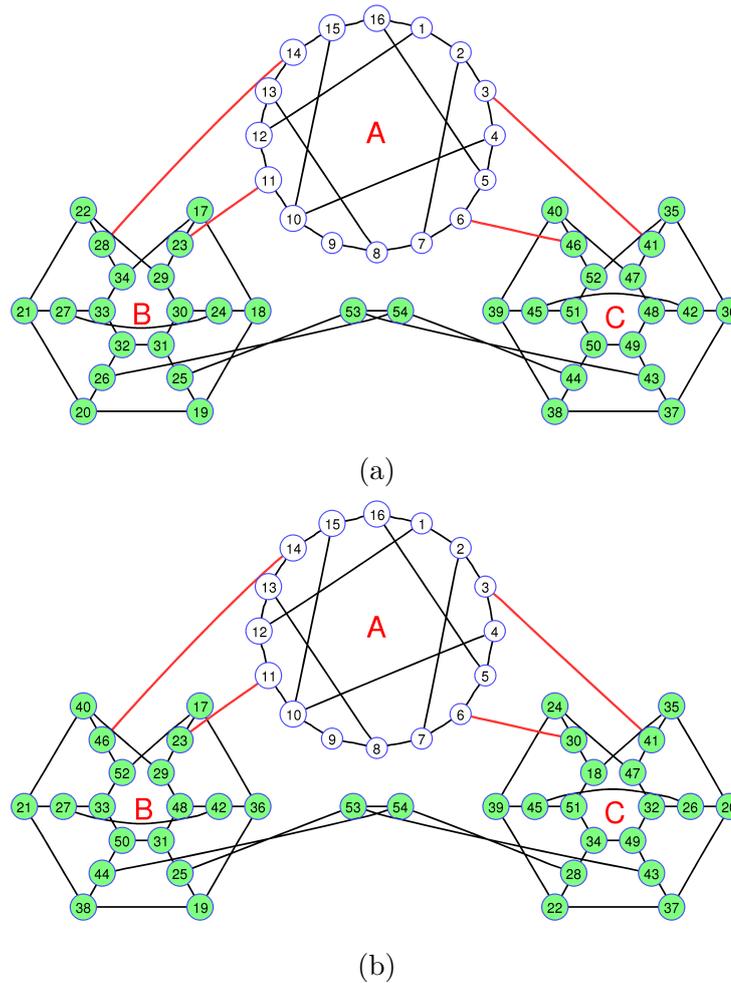


Figure 2.4: Ellingham-Horton 54-graph. Shown is the result of the spectral bisection method described in Section 2.3 using  $\mathcal{L}^\tau$ . The resulting partition is shown as  $V_1$  (colored vertices),  $V_2$  (non-colored vertices) and the partition boundary (red edges). (a)  $G$ :  $|C(V_1, V_2)| = 4$ ,  $D(V_1) = 114$ ,  $D(V_2) = 488$ . (b)  $\pi(G)$ :  $|C_\pi(V_1, V_2)| = 4$ ,  $D_\pi(V_1) = 114$ ,  $D_\pi(V_2) = 48$ .

### Dynamic spectral bisection with multiple permutations

We now demonstrate the dynamic spectral bisection of  $G$  over 10 time steps. We have a set of 9 graph isomorphisms  $\pi_t, t = 1, 2, \dots, 9$  applied in sequence to  $G$ . This set of transformations is designed to disrupt the cluster structure corresponding to the subregions B and C in Figure 2.4 via vertex interchange between the subregions  $B \cup C$  and A. In particular, when  $t$  is an even integer, the permutation on vertex labels corresponding to  $\pi_t$  is the concatenation of the eight disjoint transpositions:  $(18 + 2i, 36 + 2i)$ , for  $i = 0, 1, \dots, 8$ . When  $t$  is an odd integer the vertex label permutation  $\pi_{t,p}$  corresponding to  $\pi_t$  is given by  $\pi_{t,p}((t + 1)/2) = (t + 1)/2 + 17$ .

The sequence of graph isomorphisms  $\pi_t, t = 1, \dots, 9$  produces a total of 10

graphs, each corresponding to the state of the graph at time  $t$ ; i.e. the set of graphs  $\pi^{(t)}(G)$  where  $\pi^{(t)} = \pi_t \circ \dots \circ \pi_2 \circ \pi_1$ ,  $t = 1, \dots, 9$  and  $\pi^{(0)} = \text{Id}$ . We seek to determine a cluster structure that remains intact in *each* of these 10 graphs. To do this, we form the matrix  $\mathcal{L}^\tau$  using (2.27) and compute the second eigenvector  $\mathbf{g}_2$ . The vector  $\mathbf{f} = \mathbf{D}^{-1/2}\mathbf{g}_2$  orders the vertices and produces at most  $k - 1$  distinct partitions of the form  $\{V_1^\beta, V_2^\beta\}$  as described earlier in Section 2.2; we select the partition with the lowest value of  $H_G^\tau(V_1^\beta, V_2^\beta)$  given by (2.23).

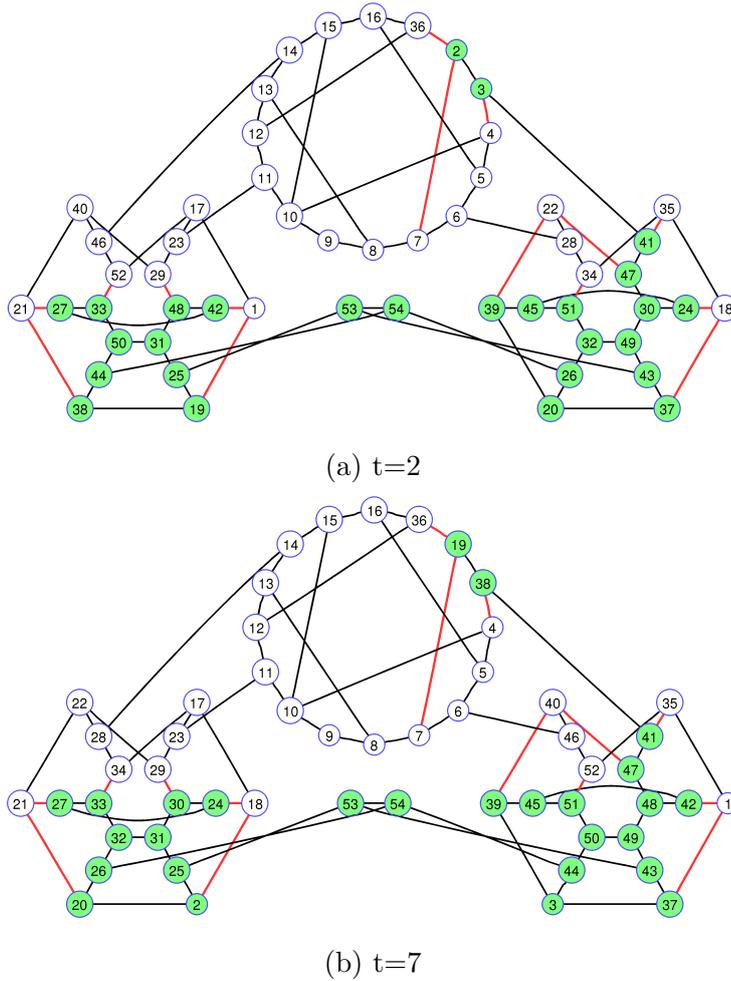


Figure 2.5: Ellingham-Horton 54-graph. Shown is the result of the spectral bisection method described in Section 2.3 using the 10 permutations  $\pi_0, \dots, \pi_9$  to create  $\mathcal{L}^\tau$  as defined in (2.27). The resulting partitions shown are  $V_1$  (colored vertices),  $V_2$  (non-colored vertices) and the partition boundary (red edges) at times  $t = 2$  and  $t = 7$ . (a)  $\pi^{(2)}(G)$ :  $|C_{\pi^{(2)}}(V_1, V_2)| = 15$ ,  $D_{\pi^{(2)}}(V_1) = 81$ ,  $D_{\pi^{(2)}}(V_2) = 81$ . (b)  $\pi^{(7)}(G)$ :  $|C_{\pi^{(7)}}(V_1, V_2)| = 15$ ,  $D_{\pi^{(7)}}(V_1) = 54$ ,  $D_{\pi^{(7)}}(V_2) = 81$ .

The results are shown in Figure 2.5a and Figure 2.5b for the time steps  $t = 2$  and  $t = 7$  respectively. The vertices corresponding to  $V_1$  are colored green and those in  $V_2$  uncolored. The edges that connects  $V_1$  and  $V_2$  are colored red. Note

that the community structures at the two different times  $t = 2$  and  $t = 7$  are exactly the same; in fact our computed partition displays the same community structure across each time step including the initial time with the corresponding graph  $G$ . The numerical quantities for  $\pi^{(t)}(G)$ ,  $t = 0, \dots, 9$  are  $|C_\pi^{(t)}(V_1, V_2)| = 15$  and  $D_\pi^{(t)}(V_1) = D_\pi^{(t)}(V_2) = 81$  for each  $t$ ; the partitions  $V_1$  and  $V_2$  are perfectly balanced in terms of degree. Finally one computes  $H_G^{10}(V_1, V_2) = 0.1852$  using (2.23). We note again that  $\hat{H}_G^{10}(V_1, V_2) = H_G^{10}(V_1, V_2)$  as all vertices have degree 3. Detailed information on each of the graphs  $\pi^{(t)}(G)$ ,  $t = 0, \dots, 9$  is provided in Table 2.2.

Quantity	$ C_\pi^{(t)}(V_1, V_2) $									$D_\pi^{(t)}(V_1)$	$D_\pi^{(t)}(V_2)$	$H_G(V_1, V_2)$	$H_G^\tau(V_1, V_2)$	
	$t$	0	1	1	3	4	5	6	7	8	9	1, ..., 9	1, ..., 9	-
$\mathcal{L}$	4	30	30	4	10	32	32	10	16	34	54	108	0.0741	0.3741
$\mathcal{L}^\tau$	15	15	15	15	15	15	15	15	15	15	81	81	0.1852	0.1852

Table 2.2: Results of spectral bisection of the Ellingham-Horton 54 graph, using the second eigenvectors of  $\mathcal{L}$  and  $\mathcal{L}^\tau$  to minimise  $\mathbf{H}_G$  and  $\mathbf{H}_G^{10}$ , respectively. The partitions  $V_1, V_2$  are obtained using the method described in Section 1.1.3 and Algorithm 2.1.

Recall that in addition to the large disruption of the clusters B and C by the vertex permutations, just over half of the 10 permutations induce vertex exchange between the cluster A and  $B \cup C$ . The dynamic community structure of the regions A and  $B \cup C$  is therefore weakened. Thus we see in Figure 2.5 a preference to group some vertices from A with those in  $B \cup C$ , and vice-versa.

### 2.4.2 Example 2: Randomly generated graph.

We now illustrate our method on a large random graph. We randomly generated a connected graph  $G$  on  $k = 1000$  vertices, with average degree approximately eight, as follows. Create a 4000-vector  $\mathbf{x}$  filled with uniformly randomly distributed integers sampled from  $\{1, 2, \dots, 1000\}$ . Create a second vector  $\mathbf{y}$  by sorting  $\mathbf{x}$  in ascending order. Produce 4000 edges of the form  $[x_i, y_i]$ ,  $i = 1, \dots, 4000$ , and remove all self-loops and duplicate edges. We arrive at a graph with 3985 edges (and a total degree sum of 7970).

The permutation  $\pi_v : V \rightarrow V$  is given by  $\pi_v(v_i) = v_{i+300} \pmod{1000}$ ,  $i = 1, \dots, 1000$ . We computed the second eigenvector of both  $\mathcal{L}$  (resp.  $\mathcal{L}^{dyn}$ ) for this graph and from these eigenvectors we created the corresponding partitions that minimise  $H_G(V_1, V_2)$  (resp.  $H_G^{dyn}(V_1, V_2)$ ,  $\hat{H}_G^{dyn}(V_1, V_2)$ ). The numerical results are summarised in Table 2.3.

Quantity	$\mathcal{L}$ minimising $H_G(V_1, V_2)$	$\mathcal{L}^{dyn}$ minimising $H_G^d(V_1, V_2)$	$\mathcal{L}^{dyn}$ minimising $\hat{H}_G^{dyn}(V_1, V_2)$
$ C(V_1, V_2) $	1122	1438	1445
$ C_\pi(V_1, V_2) $	2035	1550	1557
$D(V_1), D(V_2)$	3986, 3984	3776, 4194	3845, 4125
$D_\pi(V_1), D_\pi(V_2)$	3975, 3995	4006, 3964	4131, 3839
$H_G(V_1, V_2)$	0.2816	0.3808	0.3758
$H_G^{dyn}(V_1, V_2)$	0.3967	0.3860	0.3907
$\hat{H}_G^{dyn}(V_1, V_2)$	0.3966	0.3840	0.3769

Table 2.3: Results of spectral bisection using the second eigenvectors of  $\mathcal{L}$  and  $\mathcal{L}^{dyn}$  to find low values for  $H_G(V_1, V_2)$ ,  $H_G^{dyn}(V_1, V_2)$ , and  $\hat{H}_G^{dyn}(V_1, V_2)$ , for a randomly generated graph of 1000 vertices. The partitions  $V_1, V_2$  are obtained using the method described in Section 1.1.3 and Algorithm 2.1.

Referring first to the solution obtained from  $\mathcal{L}$ , Table 2.3 shows that the number of edges cut to disconnect the graph and minimise  $\mathbf{H}_G$  is just over one-quarter of all edges (1122 edges), indicating that there is no strong clustering in the random graph. Moreover, the bisections are almost perfectly balanced in terms of total degree counts. When subjected to the permutation dynamics  $\pi$ , the number of edges cut in  $\pi(G)$  almost doubles to 2035 edges. This is because there is no particular relation between the structure of  $G$  and the permutation  $\pi$ , so the bisection induced on  $\pi(G)$  is effectively random, and cuts about half of the total number of edges.

Considering the bisection obtained from  $\mathcal{L}^{dyn}$ , attempting to minimise  $H_G^{dyn}(V_1, V_2)$ , we see that this bisection cuts slightly more edges (1438 edges) than the bisection from  $\mathcal{L}$  (1122 edges) on  $G$ . However, when the dynamics of  $\pi$  is applied to the graph, the number of edges traversing  $V_1$  and  $V_2$  in  $\pi(G)$  increases only a little (to 1550 edges). Thus, one pays a little extra to bisect the initial graph, but this reaps large benefits when the dynamics are applied. The bisection obtained from  $\mathcal{L}^{dyn}$ , attempting to minimise  $\hat{H}_G^{dyn}(V_1, V_2)$ , cuts a slightly larger number of edges and is slightly less well degree-balanced in this example.

## 2.5 Conclusion to the chapter

Classical isoperimetry theory on graphs studies the connective structures of static graphs. In this chapter, we considered a dynamic extension of classical graph isoperimetry theory, whereby the graph connective structures are evolving due to vertex permutation dynamics. We search for community structures that are robust

to permutation dynamics, by introducing a dynamic balanced graph bisection problem designed for finding persistently highly interconnected subregions of dynamic graphs.

To find good solution to our balanced graph bisection problem, we constructed a dynamic graph Laplacian, and show that the spectrum of our dynamic graph Laplacian is characterised by a dynamically modified version of the Rayleigh quotient. Furthermore, we proved a dynamic graph Cheeger inequality, and proposed a natural extension of the Laplacian matrix-based spectral method of graph partitioning. Finally, we numerically demonstrated that eigenvectors of our dynamic graph Laplacian efficiently separate the graph into components that retain their community structure under dynamics.

The discrete dynamic isoperimetry problem on graphs naturally leads to the continuous dynamic isoperimetry problem on manifolds considered in [49]. The result of dynamic graph Cheeger inequality in this chapter, as well as the known connection between the normalised graph Laplacian and the Laplace-Beltrami operator (see e.g. [12]) shows promise for forming a theoretical connection between dynamic isoperimetry on graphs and dynamic isoperimetry on manifolds. Many prior works that relates graph partitioning problems and dynamical systems used Ulam-Galerkin method to approximate the Perron-Frobenius operator; e.g. [31, 107].



## Chapter 3

# A dynamic Laplacian for identifying Lagrangian coherent structures on weighted Riemannian manifolds.

Early attempts to characterise transport barriers in fluid dynamics include time-dependent invariant manifolds (such as lobe-dynamics [102]) and finite-time Lyapunov exponents [98, 99, 37, 68, 109]. More recently, in two-dimensional area-preserving flows, [69] proposed finding closed curves whose time-averaged length is stationary under small perturbations; this aim is closest in spirit<sup>1</sup> to the predecessor work of this chapter [49], though the latter theory applies in arbitrary finite dimensions and the curves need not be closed. In parallel to these efforts, the notion of almost-invariant sets [32] in autonomous systems spurred the development of probabilistic methods to transport based around the transfer operator. In relation to transport barriers, numerical observations [56] indicated connections between the boundaries of almost-invariant sets and invariant manifolds of low-period points. Transfer operator techniques were later extended to dynamical systems with general time dependence, with the introduction of coherent sets as the time-dependent analogues of almost-invariant sets [58, 48]. Topological approaches to phase space mixing have also been developed [63], including connections with almost-invariant sets [100]. A differential geometric perspective of shape coherence in transport was

---

<sup>1</sup>The extension [94] of [69] to three dimensions is less aligned with [49], as [94] asks for uniform expansion in all directions in the two-dimensional tangent space to potential Lagrangian coherent structures surfaces, whereas the approach of [49] in three-dimensions is simply concerned with surface growth without a uniform expansion restriction.

studied in [85, 86].

In [49], Froyland considered the identification of Lagrangian coherent structures by searching for subsets of a manifold whose boundary size to enclosed volume is minimised in a time-averaged sense under general time-dependent nonlinear dynamics. Solutions to this dynamic minimisation problem were constructed from eigenvectors of a dynamic Laplace operator, a time-average of pullbacks of Laplace operators under the dynamics. It was shown in [49] that the dynamic Laplace operator arises as a zero-diffusion limit of the transfer operator constructions for finite-time coherent sets in [48]. This result demonstrated that finite-time coherent sets (those sets that maximally resist mixing over a finite time interval), also had the persistently small boundary length to enclosed volume ratio property; intuitively this is reasonable because diffusive mixing between sets can only occur through their boundaries. Thus, finite-time coherent sets have dual minimising properties: slow mixing (probabilistic) and low boundary growth (geometric). The theory in [49] was restricted to the situation where the advective dynamics was volume-preserving, and to tracking the transport of a uniformly distributed tracer in Euclidean space.

In the current setting of weighted Riemannian manifolds, whereby the manifolds  $M$  and  $N$  are equipped with the full dimensional measures  $\mu_r$  and  $\nu_r$ , respectively. The generalisation of the above dynamic minimisation problem for a single iteration of transformation  $T : M \rightarrow N$  is

$$H_M^{dyn}(\Gamma) := \frac{\mu_{r-1}(\Gamma) + \nu_{r-1}(T\Gamma)}{2 \min\{\mu_r(M_1), \mu_r(M_2)\}}, \quad (3.1)$$

where  $M_1$  and  $M_2$  are connected components of  $M$  disconnected by  $\Gamma$ . Comparing (3.1) to the classical Cheeger ratio (1.8), the numerator of (3.1) not only measures the initial  $\mu_{r-1}$ -size of  $\Gamma$ , but also measures the  $\nu_{r-1}$ -size of image  $T\Gamma$ . We elaborate on the importance of the formulation (3.1) later in Section 3.2, for now we point out that (3.1) is the natural non-volume-preserving generalisation of Equation (1) in [49].

Beyond the generalised dynamic isoperimetric problem, our main contributions are firstly the formulation of a dynamic Sobolev constant (a dynamic version of the classical Sobolev constant given by Definition 2 on p. 96 in [22]) in our general setting and a corresponding proof of a dynamic version of the celebrated Federer-Fleming theorem (see Theorem 3.2 in [49] for the dynamic statement in the volume-preserving, uniform density, flat manifold setting), which equates the geometric Cheeger constant with the functional dynamic Sobolev constant. Secondly, we define a generalised version of the dynamic Laplace operator constructed by equation (4)

in [49]. In our general setting (see Section 3.3.1 for details), this operator is

$$\Delta^{dyn} := \frac{1}{2}(\Delta_\mu + \mathcal{H}^* \Delta_\nu \mathcal{H}), \quad (3.2)$$

where  $\Delta_\mu, \Delta_\nu$  are weighted Laplace-Beltrami operators, weighted by  $\mu_r, \nu_r$  respectively. The operator  $\mathcal{H} : L^2(M, m, \mu_r) \rightarrow L^2(N, n, \nu_r)$  is defined by normalising the transfer operator  $\mathcal{P} : L^1(M, V_m) \rightarrow L^1(N, V_n)$  via  $\mathcal{H}f = \mathcal{P}(f \cdot h_\mu)/h_\nu$  (previously defined in Section 4.1 of [48] in the flat, weighted manifold setting), where  $h_\mu, h_\nu$  are the densities of  $\mu_r, \nu_r$  with respect to the volume forms  $dV_m, dV_n$ , and  $\mathcal{H}^*$  is the adjoint of  $\mathcal{H}$ . We will see later that  $\mathcal{H}$  is simply composition with  $T^{-1}$  and  $\mathcal{H}^*$  is composition with  $T$ . See Section 3.3.2 for continuous time versions of  $\Delta^{dyn}$ . A related construction is considered in [9] from the point of view of heat flow, where they search for a single metric for a Laplace-Beltrami operator, rather than solving an isoperimetric-type problem, and follow ideas of [117] to consider flow in Lagrangian coordinates and make connections to almost-invariant sets subjected to time-dependent diffusion.

We prove a dynamic version of the well-known Cheeger inequality in our generalised setting (see [24] for the classic (static) Cheeger inequality and [49] for the dynamic Cheeger inequality in the volume-preserving, uniform density, flat manifold setting), which bounds the Cheeger constant above in terms of the dominant non-trivial eigenvalue of  $\Delta^{dyn} f = \lambda f$  (with natural Neumann-like boundary conditions). Finally, we prove that

$$\lim_{\epsilon \rightarrow 0} \frac{(\mathcal{H}_\epsilon^* \mathcal{H}_\epsilon - \text{Id})f}{\epsilon^2} = c \cdot \Delta^{dyn} f, \quad (3.3)$$

where  $\text{Id}$  is the identity,  $\mathcal{H}_\epsilon$  is an  $\epsilon$ -mollified version of  $\mathcal{H}$  (see (3.43)), used to compute finite-time coherent sets in [48] and  $c$  is an explicit constant. Because singular vectors of  $\mathcal{H}_\epsilon$  (eigenvectors of  $\mathcal{H}_\epsilon^* \mathcal{H}_\epsilon$ ) are used in [48], and eigenvectors of  $\Delta^{dyn}$  are used in the present work, this result shows that in the small perturbation limit, the purely probabilistic constructions of [48] coincide with the purely geometric constructions of 3.2.

This chapter is arranged as follows. In Section 3.1 we provide relevant background material from differential geometry. Section 3.2 describes the dynamic isoperimetric problem on weighted Riemannian manifolds and states the dynamic Federer-Fleming theorem. Section 3.3 details the dynamic Laplace operator on weighted manifolds and states the dynamic Cheeger inequality. In Section 3.4, we state the convergence result (3.3). Section 3.5 contains illustrative numerical experiments and most of the proofs are deferred to Appendix B. In comparison with [49], Theorem 3.2.4, 3.3.3, 3.3.4 and 3.4.1 in this chapter generalise, respectively Theorems 3.1, 4.1, 3.2 and 5.1 in [49].

### 3.1 Primer on differential geometry

Let  $M$  be a compact, connected  $r$ -dimensional  $C^\infty$  Riemannian manifold. We denote the boundary of  $M$  by  $\partial M$ . If  $\partial M$  is non-empty, then we assume that  $\partial M$  is  $C^\infty$ . We are interested in tracking the masses of the  $r$  and  $r - 1$  dimensional subsets of  $M$  as this manifold is transformed by a general smooth dynamical system. We now give a brief introduction of the key tools in differential geometry for performing the above task; additional details are provided in Appendix B.2.

Recall that to compute the  $r$ -dimensional volume of the objects in  $M$ , one considers a metric tensor on the tangent space  $\mathcal{T}_x M$  at the point  $x \in M$ . In particular, the Riemannian metric  $m$  on  $M$  associates each point  $x \in M$  with a symmetric bilinear form  $m(\cdot, \cdot)_x : \mathcal{T}_x M \times \mathcal{T}_x M \rightarrow \mathbb{R}$ , yielding a volume form  $\omega_m^r$  on  $M$  (see Appendix B.2.1 for more details). The differential  $r$ -form  $\omega_m^r$  defines an  $r$ -dimensional volume measure on any measurable subset  $U \subset M$  by  $V_m(U) := \int_U \omega_m^r$ . To describe the mass distribution of the objects in  $M$ , we consider a weighted Riemannian manifold  $(M, m, \mu_r)$ , where  $\mu_r$  is an absolutely continuous probability measure with respect to  $V_m$ ; that is, there exist  $h_\mu \in L^1(M, V_m)$  such that

$$\mu_r(U) = \int_U d\mu_r = \int_U h_\mu \cdot \omega_m^r,$$

for all measurable  $U \subset M$ , and  $\mu_r(M) = 1$ . Since any subset of  $M$  with  $\mu_r$  measure zero has no physical impact, without loss of generality we assume that the density  $h_\mu$  is uniformly bounded away from zero.

Let  $(N, n, \nu_r)$  be another weighted Riemannian manifold, where  $N$  is a compact, connected  $r$ -dimensional  $C^\infty$  Riemannian manifold,  $n$  the Riemannian metric on  $N$ , and  $\nu_r$  an absolutely continuous probability measure with respect to  $V_n$ . As before, we shall assume that the density  $h_\nu$  of  $\nu_r$  is uniformly bounded away from zero. Consider a general dynamical system  $T : M \rightarrow N$  that acts as a  $C^\infty$ -diffeomorphism from  $M$  onto  $N$ . For the purpose of modeling physical processes, we assume that no mass is lost under transport; that is, the measure  $\mu_r$  on  $M$  is transformed under the action of  $T$  to  $\nu_r := \mu_r \circ T^{-1}$ . Because the densities  $h_\mu, h_\nu$  are uniformly bounded away from zero,  $\nu_r = \mu_r \circ T^{-1}$ , and  $T$  is a diffeomorphism, the nondegeneracy of the metrics  $n, m$  implies that the Jacobian associated with  $T$  must be uniformly bounded above and away from zero (see Appendix B.2.3). We emphasise that  $n$  is not necessarily the pushforward of  $m$ , and that  $T$  is not an isometry from  $(M, m)$  to  $(N, n)$  in general.

Let  $\mathcal{T}M$  denote the tangent bundle of  $M$ ; that is,  $\mathcal{T}M := \cup_{x \in M} \{x\} \times \mathcal{T}_x M$ . A vector field  $\mathcal{V}$  on  $M$  is a section of the bundle  $\mathcal{T}M$ ; that is the image of  $x \in M$  under  $\mathcal{V}$  is the tangent vector  $\mathcal{V}_x \in \mathcal{T}_x M$ . For  $k \geq 1$ , we denote the space of  $k$ -times

continuously differentiable vector fields on  $M$  by  $\mathcal{F}^k(M)$ . For a pair  $\mathcal{V}, \mathcal{W} \in \mathcal{F}^k(M)$ , one can view  $m(\mathcal{V}, \mathcal{W}) : M \rightarrow \mathbb{R}$  as a  $C^k$  function on  $M$  given by  $m(\mathcal{V}, \mathcal{W})(x) = m(\mathcal{V}_x, \mathcal{W}_x)_x$  for all  $x \in M$ . Denote by  $\mathcal{T}^*M$  the dual bundle of  $\mathcal{T}M$ ; that is the cotangent bundle  $\mathcal{T}^*M : \cup_{x \in M} \{x\} \times \mathcal{T}_x^*M$ , where  $\mathcal{T}_x^*M$  is the vector dual of  $\mathcal{T}_xM$ . The covector fields on  $N$  are sections of the bundle  $\mathcal{T}^*M$ .

It is practical to associate the diffeomorphism  $T : M \rightarrow N$  with the linear tangent map  $T_*$  that takes vector fields on  $M$  to vector fields on  $N$ , which we now define. Let  $\sigma_x : (-\epsilon, \epsilon) \rightarrow M$  be a family of parameterised curves in  $M$ , with  $\sigma_x(0) = x \in M$ . Suppose for each  $x \in M$  that  $\mathcal{V}_x \in \mathcal{T}_xM$  is tangent to the curve  $\sigma_x$  at  $x$ . The action of  $\mathcal{V}_x$  on a differentiable function  $f$  at each point  $x \in M$  is defined to be the number

$$\mathcal{V}_x f|_x := \left. \frac{\partial(f \circ \sigma_x)}{\partial t} \right|_{t=0}; \quad (3.4)$$

that is  $\mathcal{V}_x f|_x$  measures the initial rate of change of  $f$  along a curve with tangent  $\mathcal{V}_x$  at the point  $x$ . The local pushforward map  $(T_*)_{x_0} : \mathcal{T}_{x_0}M \rightarrow \mathcal{T}_{T x_0}N$  is defined at a fixed point  $x_0 \in M$  as

$$[(T_*)_{x_0} \mathcal{V}_{x_0}]g|_{T x_0} := \mathcal{V}_{x_0}(g \circ T)|_{x_0},$$

for all  $g \in C^k(N, \mathbb{R})$ . The collection of local pushforward maps define a linear tangent map  $T_* : \mathcal{F}^k(M) \rightarrow \mathcal{F}^k(N)$  via

$$[(T_* \mathcal{V})g](T x) := [(T_*)_x \mathcal{V}_x]g|_{T x}, \quad (3.5)$$

for all  $x \in M$ , and  $g \in C^k(N, \mathbb{R})$ .

Next, we define the linear cotangent map  $T^*$  that takes covector fields on  $N$  to covector fields on  $M$  as follows. Given a vector field  $\mathcal{V}$  on  $M$ , the action of  $\mathcal{V}$  on a differentiable function  $f$  on  $M$  is a function  $\mathcal{V}f : M \rightarrow \mathbb{R}$  given by  $\mathcal{V}f(x) := \mathcal{V}_x f|_x$ . By the duality of the tangent and cotangent spaces, the cotangent vector fields are differential 1-forms  $df$  that map vector fields on  $M$  to functions on  $M$  via  $df(\mathcal{V}) := \mathcal{V}f$ . The cotangent mapping on differential 1-forms is defined by

$$[T^*(dg)]\mathcal{V} := dg(T_* \mathcal{V}) = \mathcal{V}(g \circ T) = (T_* \mathcal{V})g, \quad (3.6)$$

for all  $\mathcal{V} \in \mathcal{F}^k(M)$  and  $g \in C^k(N, \mathbb{R})$ . One can associate the cotangent mapping  $T^*$  with an exterior product of  $p$ -forms,  $1 \leq p \leq r$  (see Appendix B.2.1). In particular, since the metric tensor  $n$  is a symmetric 2-form on  $\mathcal{F}^k(N)$ , one defines the *pullback metric* of  $n$  by

$$T^*n(\mathcal{V}_1, \mathcal{V}_2)(x) := n(T_* \mathcal{V}_1, T_* \mathcal{V}_2)(T x), \quad (3.7)$$

for all vector fields  $\mathcal{V}_1, \mathcal{V}_2$  on  $M$ , and each point  $x \in M$ ; that is, the pullback metric  $T^*n$  is defined in such a way that  $T$  is an isometry from  $(M, T^*n)$  to  $(N, n)$ .

To compute the co-dimension 1 volume of  $r - 1$  dimensional subsets of  $(M, m, \mu_r)$  and  $(N, n, \nu_r)$ , one uses the induced Riemannian metric. Suppose  $\Gamma$  is a compact  $C^\infty$  co-dimension 1 subset of  $M$ . The embedding  $\Phi : \Gamma \hookrightarrow M$  induces a Riemannian metric on  $\Gamma$  via the pullback metric associated with  $\Phi$ ; that is  $\Phi^*m$  is the induced metric on  $\Gamma$ . Let  $\omega_m^{r-1}$  denote the  $(r - 1)$ -dimensional volume form corresponding to the induced metric  $\Phi^*m$  (i.e.  $\omega_m^{r-1} = \omega_{\Phi^*m}^r$ ). To describe the distribution of mass on  $\Gamma$ , we define the  $(r - 1)$ -dimensional measure  $\mu_{r-1}$  on  $M$  by

$$\mu_{r-1}(\Gamma) := \int_{\Gamma} h_{\mu} \cdot \omega_m^{r-1}, \quad (3.8)$$

where  $h_{\mu}$  is the density of  $\mu_r$ ; the measure  $\mu_{r-1}$  captures the mass distribution on  $\Gamma$  via  $h_{\mu}$ . Similarly, the co-dimension 1 mass distribution on a  $C^\infty$ , compact subset of  $N$  is captured by the  $r - 1$  dimensional measure  $\nu_{r-1}$  via the density  $h_{\nu}$  of  $\nu_r$ . We now provide an example to demonstrate that the  $\mu_{r-1}$  measure on certain hypersurfaces can be significantly increased under the action of a transformation  $T$ .

### 3.1.1 Shear on a two-dimensional cylinder

Let  $M = [0, 4)/\sim \times [0, 1]$  be a 2-dimensional cylinder in  $\mathbb{R}^2$ , where  $\sim$  is identification at interval endpoints; that is,  $M$  is periodic in the first coordinate with period 4. The Riemannian metric  $e$  on  $M$  is given by the Kronecker delta  $\delta_{ij}$ , so that the volume form  $\omega_e^2$  on  $M$  is  $\omega_m^2 = dx_1 dx_2$ . To form a weighted Riemannian manifold  $(M, e, \mu_2)$ , we set the density  $h_{\mu}$  of  $\mu_2$  to be a positive and periodic function  $h_{\mu}(x_1, x_2) = \frac{1}{8}(\sin(\pi x_1) + 2)$ .

Consider the hypersurface  $\Gamma = \{x \in M : x_1 = 1.5, 3.5\}$ ; we choose this surface because it is the solution of the classical “static” isoperimetric problem defined by minimising (3.9) without the second term in the numerator. The curve  $\Gamma$  is two vertical lines on  $M$  that pass over regions with minimal density  $h_{\mu}$  as shown in Figure 3.1a. One can compute  $\mu_1(\Gamma)$  analytically by noting that the induced Riemannian metric on  $\Gamma$  is given by  $dx_2$ ; thus

$$\mu_1(\Gamma) = \int_0^1 h_{\mu}(1.5, x_2) dx_2 + \int_0^1 h_{\mu}(3.5, x_2) dx_2 = 0.25.$$

Let us now apply the following transformation to  $M$ ,

$$T(x_1, x_2) = \left( x_1 + \frac{\cosh(2x_2) - 1}{2}, x_2 \right),$$

where the first coordinate is computed modulo 4. The map  $T$  is a nonlinear horizontal shear. The hypersurface  $\Gamma$  is transformed to  $T\Gamma$  under the action of  $T$  as shown

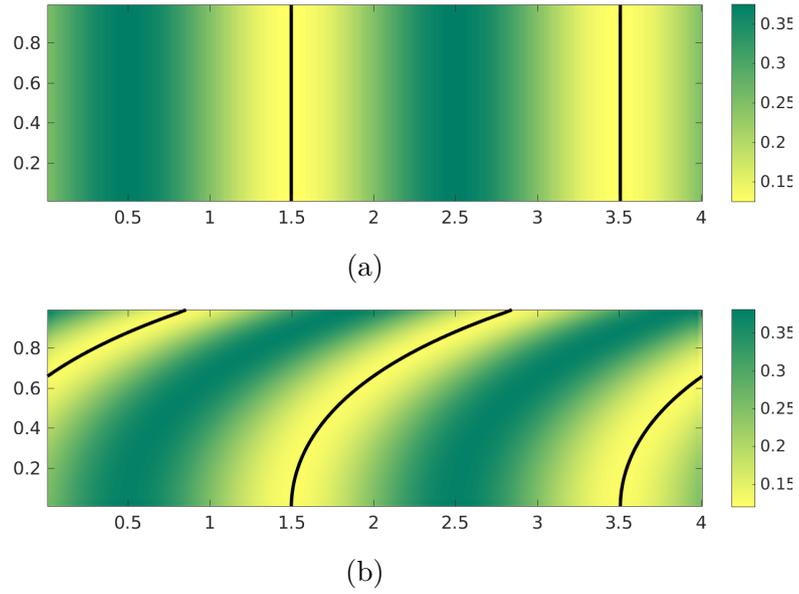


Figure 3.1: Deformation of 2-dimensional cylinder under nonlinear shear  $T$ . (a) Colours are values of  $h_\mu$ , and black lines are the hypersurface  $\Gamma$ . (b) Values of  $h_\mu \circ T^{-1}$ , and  $T\Gamma$ .

in Figure 3.1b. The shearing magnitude  $(\cosh(2x_2) - 1)/2$  is chosen to simplify the analytical computation of  $\nu_1(T\Gamma)$ . It is easy to verify that  $T$  is area-preserving. Since  $T$  is area-preserving and  $\nu_2 = \mu_2 \circ T^{-1}$ , one has

$$\int_{TM} h_\nu dx_1 dx_2 = \int_M h_\mu dx_1 dx_2 = \int_{TM} h_\mu \circ T dx_1 dx_2,$$

which implies  $h_\nu = h_\mu \circ T$  in this example.

To compute the  $\nu_1$  measure on  $T\Gamma$ , we parametrise the curve  $T\Gamma$  by  $T\Gamma = (\sigma_c(s), s)$  for  $s \in [0, 1]$ , where  $\sigma_c(s) = c + \frac{\cosh(2s)-1}{2}$ , for  $c = 1.5, 3.5$ . Furthermore, by using the fact that  $h_\nu = h_\mu \circ T$ , one has

$$h_\nu(\sigma_c(s), s) = h_\mu(c, s) = \frac{\sin(\pi c) + 2}{8} \Big|_{c=1.5, 3.5} = \frac{1}{8},$$

for all  $t \in [0, 1]$ . Therefore

$$\begin{aligned} & \nu_1(T\Gamma) \\ &= \int_0^1 \sqrt{1 + \left| \frac{\partial \sigma_{1.5}}{\partial s}(s) \right|^2} \cdot h_\nu(\sigma_{1.5}(s), s) ds + \int_0^1 \sqrt{1 + \left| \frac{\partial \sigma_{3.5}}{\partial s}(s) \right|^2} \cdot h_\nu(\sigma_{3.5}(s), s) ds \\ &= \frac{2}{8} \int_0^1 \sqrt{1 + \left| \frac{\partial \sigma_{1.5}}{\partial s} \right|^2} ds \\ &= \frac{2}{8} \int_0^1 \sqrt{1 + \sinh^2(2s)} ds = 0.4534. \end{aligned}$$

Thus the  $\nu_1$  measure of  $T\Gamma$  is almost double that of the  $\mu_1$  measure of  $\Gamma$ . Correspondingly, the numerator in (3.9) will be undesirably large. In Section 3.5 we show how to use our new machinery to find an improved choice for  $\Gamma$  that takes into account both the weight  $h_\mu$  and the dynamics of  $T$ .

## 3.2 A dynamic isoperimetric problem on weighted manifolds

Our goal is to detect Lagrangian coherent structures on the weighted Riemannian manifold  $(M, m, \mu_r)$ ; i.e. subsets of  $M$  that resist mixing with the surrounding phase space by having persistently small boundary size to internal size. Following [49], we introduce a version of the *dynamic isoperimetric problem*, generalised to the situation where the dynamics need not be volume preserving, and occurs on a possibly weighted, possibly curved manifold.

Let  $\Gamma$  be a compact  $C^\infty$ -hypersurface in  $M$  that disconnects  $M$  into two disjoint open subsets  $M_1$  and  $M_2$  with  $M_1 \cup \Gamma \cup M_2 = M$ . To begin with, we model the dynamics as a single iterate of  $T$ . The subsets  $M_1$  and  $M_2$  are transformed into  $N_1 := TM_1$  and  $N_2 := TM_2$ , with  $T\Gamma$  the disconnecting surface separating  $N_1$  and  $N_2$  in  $N$ . Consider the following optimisation problem:

**Definition 3.2.1.** Define the *dynamic Cheeger ratio*  $H_M^{dyn}$  by

$$H_M^{dyn}(\Gamma) = \frac{\mu_{r-1}(\Gamma) + \nu_{r-1}(T\Gamma)}{2 \min\{\mu_r(M_1), \mu_r(M_2)\}}. \quad (3.9)$$

The *dynamic isoperimetric problem* is defined by the optimisation problem

$$\mathbf{H}_M^{dyn} = \inf_{\Gamma} \{H_M^{dyn}(\Gamma)\}, \quad (3.10)$$

where  $\Gamma$  varies over all  $C^\infty$ -hypersurfaces in  $M$  that partition  $M$  into  $M = M_1 \cup \Gamma \cup M_2$ . The number  $\mathbf{H}_M^{dyn}$  is called the *dynamic Cheeger constant*.

Note that by the definition of  $\nu_r$ , one has  $\mu_r(M_1) = \nu_r(N_1)$  and  $\mu_r(M_2) = \nu_r(N_2)$ . Importantly, one does *not* have  $\mu_{r-1}(\Gamma) = \nu_{r-1}(T\Gamma)$  in general, because  $n$  is not necessarily the pushforward of  $m$  (see also the direct computation in Section 3.1.1). Thus, one could rewrite (3.9) as

$$H_M^{dyn}(\Gamma) = \frac{\mu_{r-1}(\Gamma)}{2 \min\{\mu_r(M_1), \mu_r(M_2)\}} + \frac{\nu_{r-1}(T\Gamma)}{2 \min\{\nu_r(TM_1), \nu_r(TM_2)\}}. \quad (3.11)$$

By searching over all  $C^\infty$ -hypersurfaces  $\Gamma$  in  $M$  to minimise  $H_M^{dyn}(\Gamma)$ , the first ratio term of (3.11) attempts to minimise mixing between the subsets  $M_1$  and  $M_2$  across

the boundary  $\Gamma$ , through the mechanism of small co-dimensional 1 mass  $\mu_{r-1}(\Gamma)$  at the initial time, and small co-dimensional 1 mass  $\nu_{r-1}(T\Gamma)$  at the final time. Having a persistently small boundary is consistent with slow mixing in the presence of small magnitude diffusion, and is also consistent with measures of mixing adapted to purely advective dynamics such as the mix-norm [87] and negative index Sobolev space norms [118]. The reason for the constraint  $\min\{\mu_r(M_1), \mu_r(M_2)\}$  is to ensure that  $M_1$  and  $M_2$  found, *both* have macroscopic  $r$ -dimensional mass to avoid trivial solutions. Thus, the optimal solution for (3.10) is a  $C^\infty$ -hypersurface that represents an excellent candidate for a Lagrangian coherent structure, in the sense that the corresponding subsets  $M_1$  and  $M_2$  are able to retain their resistance to mixing in the presence of the prescribed dynamics  $T$ .

To see why this problem is a truly dynamic problem, consider the 2-dimensional flat cylinder  $[0, 4]/\sim \times [0, 1]$  described in Section 3.1.1. The hypersurface  $\Gamma = \{x \in M : x_1 = 1.5, 3.5\}$  partitions  $M$  into two disjoint subsets  $M_1 = (1.5, 3.5) \times [0, 1]$  and  $M_2 = [0, 1.5) \times [0, 1] \cup (3.5, 4) \times [0, 1]$ , forming the partition  $M = M_1 \cup \Gamma \cup M_2$ . It is straightforward to calculate  $\mu_2(M_1) = \mu_2(M_2) = 0.5$ . We note that  $\Gamma$  is optimally minimising for the first ratio term of (3.11); thus mixing is minimised between  $M_1$  and  $M_2$ . However, under the action of  $T$ , the  $\nu_1$  measure on  $T\Gamma$  is almost doubled (from 0.25 to 0.4534). Thus, the sets  $M_1$  and  $M_2$  are not able to maintain their resistance to mixing, and therefore are poor candidates for Lagrangian coherent structures (LCSs).

### 3.2.1 Multiple discrete time steps and continuous time

The “single iterate” problem described above can easily be extended to multiple discrete time steps or continuous time. Let  $\{(M^t, m^t, \mu_r^t)\}_{t=1}^\tau$  be  $\tau$ ,  $r$ -dimensional weighted Riemannian manifolds, where each  $M^1, M^2, \dots, M^\tau$  is  $C^\infty$ , compact, and connected. For each  $1 \leq t \leq \tau$ , define co-dimension 1 measures  $\mu_{r-1}^t$  on  $M^t$  via the densities  $h_\mu^t$  of each  $\mu_r^t$  analogous to (3.8). Let us now consider a composition of several maps  $T_1, T_2, \dots, T_{\tau-1}$ , such that  $T_t(M^t) = M^{t+1}$  and  $\mu_r^t = \mu_r^{t+1} \circ T_t$  for  $t = 1, 2, \dots, \tau - 1$ . Denoting  $T^{(t)} = T_t \circ \dots \circ T_2 \circ T_1$ ,  $t = 1, \dots, \tau - 1$  and with  $T^{(0)}$  the identity map. These maps might arise, for example, as  $\tau$ -time maps of a time-dependent flow. If we wish to track the evolution of a coherent set under these maps, penalising the boundary of the evolved set  $T^{(t)}(\Gamma)$  after the application of each  $T_i$ , then we can define

$$H_M^\tau(\Gamma) := \frac{\frac{1}{\tau} \sum_{t=0}^{\tau-1} \mu_{r-1}^{t+1}(T^{(t)}\Gamma)}{\min\{\mu_r^1(M_1^1), \mu_r^1(M_2^1)\}}, \quad (3.12)$$

and consider the time-discrete dynamic optimisation problem

$$\mathbf{H}_M^\tau := \inf_{\Gamma} H_M^\tau(\Gamma), \quad (3.13)$$

as a natural generalisation of  $\mathbf{H}_M^{dyn}$ .

In continuous time, we consider an evolving Riemannian manifold  $M(t)$ ,  $t \in [0, \tau]$  under a (possibly time-dependent) ODE  $\dot{x} = F(x, t)$ , where  $F(x, t)$  is  $C^\infty$  at each  $x \in M(t)$ ; i.e. the initial manifold  $M(0)$  is transformed under the smooth flow maps  $T^{(t)} : M(0) \rightarrow M(t)$  arising from  $F$  for each  $t \in [0, \tau]$ . We denote the Riemannian metric on  $M(t)$  by  $m^t$ , and define absolutely continuous probability measures  $\mu_r^t$  on  $M(t)$  for each  $t \in [0, \tau]$ ; one has an evolving weighted Riemannian manifold  $(M(t), m^t, \mu_r^t)$ . Note that the metrics  $m^t$  need not be related for different  $t$ . For all  $t \in [0, \tau]$ , we assume  $\mu_r^0 = \mu_r^t \circ T^{(t)}$  on  $M(0)$ . Define

$$H_M^{[0, \tau]}(\Gamma) := \frac{\frac{1}{\tau} \int_0^\tau \mu_{r-1}^t(T^{(t)}\Gamma) dt}{\min\{\mu_r^0(M_1(0)), \mu_r^0(M_2(0))\}}, \quad (3.14)$$

and

$$\mathbf{H}_M^{[0, \tau]} := \inf_{\Gamma} H_M^{[0, \tau]}(\Gamma), \quad (3.15)$$

as a time-continuous generalisation of  $\mathbf{H}_M^{dyn}$ .

### 3.2.2 Dynamic Federer-Fleming theorem on weighted manifolds

Our first result on dynamic isoperimetry is the dynamic version of the Federer-Fleming theorem, which links  $\mathbf{H}_M^{dyn}$  with a function-based optimisation problem. The *gradient* of  $f \in C^1(M, \mathbb{R})$  denoted by  $\nabla_m f$  is a vector field satisfying

$$m(\nabla_m f, \mathcal{V}) = \mathcal{V}f, \quad (3.16)$$

for all  $\mathcal{V} \in \mathcal{F}^k(M)$ . To track the transformation of a function in  $L^1(M, V_m)$  under  $T$ , the standard tool in dynamical systems is the Perron-Frobenius operator  $\mathcal{P} : L^1(M, V_m) \rightarrow L^1(N, V_n)$  given by

$$\int_U \mathcal{P}h_\mu \cdot \omega_n^r = \int_{T^{-1}U} h_\mu \cdot \omega_m^r, \quad (3.17)$$

for all measurable  $U \subset N$ . For a point-wise definition of  $\mathcal{P}$ , see (B.22) in the appendix. Recalling  $h_\nu \in L^1(N, V_n)$  is the density of  $\nu_r$  with respect to  $\omega_n^r$ , and the fact that  $\nu_r = \mu_r \circ T^{-1}$ , one has

$$\int_U \mathcal{P}h_\mu \cdot \omega_n^r = \int_{T^{-1}U} h_\mu \cdot \omega_m^r = \mu_r(T^{-1}U) = \nu_r(U) = \int_U h_\nu \cdot \omega_n^r, \quad (3.18)$$

for all measurable  $U$  in  $N$ . Therefore,  $h_\nu = \mathcal{P}h_\mu$ . We define the pushforward operator  $\mathcal{H} : L^2(M, m, \mu_r) \rightarrow L^2(N, n, \nu_r)$  (from [48]) by

$$\mathcal{H}f := \frac{\mathcal{P}(f \cdot h_\mu)}{h_\nu}. \quad (3.19)$$

**Lemma 3.2.2.** *The operator  $\mathcal{H} : L^2(M, m, \mu_r) \rightarrow L^2(N, n, \nu_r)$  is well defined, may be expressed as  $\mathcal{H}f = f \circ T^{-1}$ , and has adjoint  $\mathcal{H}^*g = g \circ T$ .*

The proof of this result is given in the Appendix B (see Lemma B.2.4). Following [49], we define the *dynamic Sobolev constant*:

**Definition 3.2.3.** Define the dynamic Sobolev constant  $\mathbf{s}^{dyn}$  by

$$\mathbf{s}^{dyn} = \inf_f \frac{\int_M |\nabla_m f|_m d\mu_r + \int_N |\nabla_n \mathcal{H}f|_n d\nu_r}{2 \inf_{\alpha \in \mathbb{R}} \int_M |f - \alpha|_m d\mu_r}, \quad (3.20)$$

where  $f : M \rightarrow \mathbb{R}$  varies over all  $C^\infty$  functions on  $M$ , and  $\mathcal{H}$  is given by (3.19).

The dynamic Sobolev constant  $\mathbf{s}^{dyn}$  defined above admits the following geometric interpretation: consider the numerator of  $\mathbf{s}^{dyn}$ , one can show (by Lemma B.2.1 in the appendix) that

$$\int_M |\nabla_m f|_m d\mu_r = \int_{-\infty}^{\infty} \mu_{r-1}(\{f = \beta\}) dt,$$

and,

$$\begin{aligned} \int_N |\nabla_n \mathcal{H}f|_n d\nu_r &= \int_{-\infty}^{\infty} \nu_{r-1}(\{\mathcal{H}f = \beta\}) d\beta. \\ &= \int_{-\infty}^{\infty} \nu_{r-1}(T\{f = \beta\}) d\beta, \end{aligned}$$

where the final equality is due to Lemma 3.2.2. Furthermore, there is a deep connection between  $\mathbf{s}^{dyn}$  and the dynamic Cheeger constant  $\mathbf{H}_M^{dyn}$ . One has

**Theorem 3.2.4 (Dynamic Federer-Fleming theorem).** *Let  $(M, m, \mu_r)$  and  $(N, n, \nu_r)$  be weighted Riemannian manifolds, where  $M$  and  $N$  are  $C^\infty$ , compact and connected. Let  $T : M \rightarrow N$  be a  $C^\infty$  diffeomorphism, with  $\nu_r = \mu_r \circ T^{-1}$ . Assume the density of  $\mu_r$  is  $C^\infty$  and uniformly bounded away from zero. Define  $\mathbf{H}_M^{dyn}$  and  $\mathbf{s}^{dyn}$  by (3.9) and (3.20) respectively. Then*

$$\mathbf{s}^{dyn} = \mathbf{H}_M^{dyn}. \quad (3.21)$$

*Proof.* The inequality  $\mathbf{s}^{dyn} \geq \mathbf{H}_M^{dyn}$  is a straightforward modification of the corresponding result in [49]. The other direction is deferred to the appendix.  $\square$

Furthermore, in the notation of Section 3.2.1 one can define the continuous time-step dynamic Sobolev constant for continuous-time dynamics by

$$\mathbf{s}^{[0,\tau]} = \inf_f \frac{\frac{1}{\tau} \int_0^\tau \left( \int_{M(t)} |\nabla_{m^t} f|_{m^t} d\mu_r^t \right) dt}{\inf_\alpha \int_{M(0)} |f - \alpha|_{m^0} d\mu_r^0}. \quad (3.22)$$

Again by the linearity of our construction, it is straightforward to obtain a dynamic Federer-Fleming theorem for continuous-time dynamics; that is

$$\mathbf{s}^{[0,\tau]} = \mathbf{H}_M^{[0,\tau]}. \quad (3.23)$$

The proof is obtained by a straightforward modification of the proof of Theorem 3.2.4 analogous to the continuous-time modification in the proof of Corollary 3.3 in [49].

### 3.3 The dynamic Laplace operator on weighted manifolds

In this section, we further develop the theory of dynamic isoperimetry established for  $\mathbb{R}^r$  in [49], to obtain results that hold on weighted, non-flat Riemannian manifolds  $(M, m, \mu_r)$  for non-volume-preserving dynamics. More precisely, we define the dynamic Laplace operator and state and prove dynamic versions of Cheeger's inequality. The dynamic Laplace operator will be the key object in the computation of solutions of the dynamic isoperimetric problem.

#### 3.3.1 The dynamic Laplace-Beltrami operator

Classical isoperimetric theory has deep connections with the Laplace-Beltrami operator (see [17, 22, 80, 90]). It is well known that one can recover certain geometrical information about a manifold  $M$  from the spectrum of Laplace-Beltrami operator [101, 105]. In this work, our domain of interest is a weighted Riemannian manifold  $(M, m, \mu_r)$ . The dynamics  $T$  maps  $M$  onto  $N = T(M)$ . The geometric properties of  $N$  can be drastically different to  $M$ , and we are motivated to construct an operator on  $(M, m, \mu_r)$  whose spectrum reveals important geometric structures on both  $M$  and  $N$ .

For an unweighted Riemannian manifold  $M$ , the standard Laplace-Beltrami operator is defined as the composition of the divergence with the gradient [22]. Let  $U \subseteq M$  be open, with  $C^\infty$  boundary  $\partial U$  and unit normal bundle  $\mathbf{n}$  along  $\partial U$ ; i.e.

for  $\mathcal{W} \in \mathcal{F}^k(\partial U)$ ,  $m(\mathcal{W}, \mathbf{n})(x) = 0$  for all  $x \in \partial U$ . The *divergence* of  $\mathcal{V} \in \mathcal{F}^1(M)$ , denoted by  $\operatorname{div}_m \mathcal{V}$  is a function satisfying

$$\int_U \operatorname{div}_m \mathcal{V} \cdot \omega_m^r := \int_{\partial U} m(\mathcal{V}, \mathbf{n}) \cdot \omega_m^{r-1}, \quad (3.24)$$

for all open  $U \subseteq M$ . The *Laplace-Beltrami* operator acting on a function  $f \in C^2(M, \mathbb{R})$  is defined by  $\Delta_m f := \operatorname{div}_m(\nabla_m f)$ , where  $\nabla_m$  is as in (3.16).

Recall that in the setting of a weighted Riemannian manifold  $(M, m, \mu_r)$ , if  $h_\mu$  is the density of  $\mu_r$ , then when computing weighted volumes, the volume form  $\omega_m^r$  is scaled by  $h_\mu$  at each point in  $M$ . According to the definition (3.16), the gradient does not depend on the weight  $h_\mu$ , because the metric  $m$  is independent of  $h_\mu$ . However, the divergence given by (3.24) does depend on  $h_\mu$  because it is defined in terms of  $\omega_m^r$ . We define the *weighted divergence*  $\operatorname{div}_\mu$  of a  $\mathcal{V} \in \mathcal{F}^1(M)$  for  $(M, m, \mu_r)$  by

$$\operatorname{div}_\mu \mathcal{V} := \frac{1}{h_\mu} \operatorname{div}_m(h_\mu \mathcal{V}), \quad (3.25)$$

where the density  $h_\mu$  of  $\mu_r$  is assumed to be  $C^1(M, \mathbb{R})$ . Note that by (3.24)

$$\int_U (\operatorname{div}_\mu \mathcal{V}) \cdot h_\mu \omega_m^r = \int_U \operatorname{div}_m(h_\mu \mathcal{V}) \cdot \omega_m^r = \int_{\partial U} m(h_\mu \mathcal{V}, \mathbf{n}) \cdot \omega_m^{r-1} = \int_{\partial U} m(\mathcal{V}, \mathbf{n}) \cdot h_\mu \omega_m^{r-1}. \quad (3.26)$$

Hence, the definition (3.25) for weighted divergence is analogous to the unweighted version (3.24).

Now as a consequence of (3.25) and the well-known fact that  $\operatorname{div}_m(h_\mu \mathcal{V}) = \operatorname{div}_m(\mathcal{V}) + m(\nabla_m h_\mu, \mathcal{V})$  (see e.g equation (13) p.3 in [22]), one has the following definition for the *weighted Laplacian* on a weighted Riemannian manifold  $(M, m, \mu_r)$ :

$$\Delta_\mu f := \operatorname{div}_\mu(\nabla_m f) = \frac{1}{h_\mu} \operatorname{div}_m(h_\mu \nabla_m f) = \Delta_m f + \frac{m(\nabla_m h_\mu, \nabla_m f)}{h_\mu}, \quad (3.27)$$

for all  $f \in C^2(M, \mathbb{R})$ . Analogous to (3.27), one forms the weighted Laplacian  $\Delta_\nu$  on  $N$  with respect to the metric  $n$  and density  $h_\nu$  for the weighted Riemannian manifold  $(N, n, \nu_r)$ .

We now describe the construction of the *dynamic* version of  $\Delta_\mu$ , where we push forward and pull back functions between  $L^2(M, m, \mu_r)$  and  $L^2(N, n, \nu_r)$  by using  $\mathcal{H}$  (3.19) and its adjoint  $\mathcal{H}^*$ .

**Definition 3.3.1.** Assume the density of  $\mu_r$  is  $C^1(M, \mathbb{R})$ . Define the *dynamic Laplacian*  $\Delta^{dyn} : C^2(M, \mathbb{R}) \rightarrow C^0(M, \mathbb{R})$  by

$$\Delta^{dyn} := \frac{1}{2}(\Delta_\mu + \mathcal{H}^* \Delta_\nu \mathcal{H}), \quad (3.28)$$

where the weighted Laplacians  $\Delta_\mu, \Delta_\nu$  are given by (3.27), and  $\mathcal{H}$  is as in (3.19).

The first term in the RHS of (3.28) is the weighted Laplacian  $\Delta_\mu$  on  $f \in C^2(M, \mathbb{R})$ . The second term pushes  $f$  forward by  $\mathcal{H}$  to the function  $\mathcal{H}f$ . This is then followed by the application of the weighted Laplacian  $\Delta_\nu$  to the function  $\mathcal{H}f$ . The weighted Laplacian  $\Delta_\nu$  provides geometric information on the weighted manifold  $(N, n, \nu_r)$ . The result is finally pulled back to a continuous function on  $M$  via  $\mathcal{H}^*$ . For example, consider the familiar setting of  $(M, e, \ell)$ , where  $M$  is an open subset of  $\mathbb{R}^r$ , with  $\ell$  the Lebesgue measure on  $M$  and  $e$  the standard Euclidean metric (i.e. on each point in  $M$ ,  $e_{ij} = \delta_{ij}$  for all  $1 \leq i, j \leq r$ ). If  $T : M \rightarrow N$  is volume preserving, then in the standard Euclidean coordinates  $\{x_i\}_{i=1}^r$  for  $M$ , and  $\{y_i\}_{i=1}^r$  for  $N$ , one has

$$\Delta^{dyn} f = \frac{1}{2} \sum_{i=1}^r \left( \frac{\partial^2 f}{\partial x_i^2} + \frac{\partial^2 (f \circ T^{-1})}{\partial y_i^2} \circ T \right),$$

for all  $f \in C^2(M, \mathbb{R})$ ; i.e.  $\Delta^{dyn}$  is precisely the definition of the dynamic Laplacian in [49] (where it is denoted by  $\hat{\Delta}$ ).

Corollary B.2.7 in the appendix provides an alternate representation of  $\Delta^{dyn}$ :

$$\Delta^{dyn} f = \frac{1}{2} (\Delta_m + \mathcal{H}^* \Delta_n \mathcal{H}) f + \frac{1}{2} \left( \frac{m(\nabla_m h_\mu, \nabla_m f)}{h_\mu} + \frac{n(\nabla_n h_\nu, \nabla_n \mathcal{H}f) \circ T}{h_\nu \circ T} \right). \quad (3.29)$$

The effect of the densities  $h_\mu, h_\nu$  is completely captured by the terms in the second parentheses of (3.29). Finally and importantly, we have

**Proposition 3.3.2.** *The operator  $\Delta^{dyn}$  may be represented as*

$$\Delta^{dyn} = \frac{1}{2} (\Delta_\mu + \Delta_{\tilde{\mu}}) f, \quad (3.30)$$

where  $\Delta_{\tilde{\mu}}$  is the weighted Laplace-Beltrami operator on  $M$  defined by (3.27) with respect to the metric  $T^*n$  and density  $\mathcal{H}^*h_\nu = h_\nu \circ T$ .

For the proof, see Corollary B.2.8 in the appendix. We briefly discuss some special cases of Proposition 3.3.2. If  $(M, m) = (N, n)$ , then  $\Delta_{\tilde{\mu}}$  in (3.30) is the weighted Laplace-Beltrami operator on  $M$  with respect to the metric  $T^*m$  and density  $\mathcal{H}^*h_\nu = \frac{h_\mu}{|\det J_T|}$ , where  $J_T$  is the Jacobian matrix associated with  $T$  (see (B.7)). If  $N = T(M) \subset \mathbb{R}^d$ ,  $m = n = e$ , and  $T$  is volume preserving, then  $\Delta_{\tilde{\mu}}$  is the Laplace-Beltrami operator on  $M$  with respect to the metric  $T^*e$  and density  $\mathcal{H}^*h_\nu = h_\mu$ . Finally, if  $h_\mu \equiv 1$  (uniform density) and  $T$  is volume preserving, one is in the setting of [49], and  $\Delta_{\tilde{\mu}}$  in (3.28) is the unweighted Laplace-Beltrami operator with respect to the metric  $T^*e$ .

### 3.3.2 Continuous time

We now describe a time-continuous<sup>2</sup> version of (3.28). Let  $(M(t), m^t, \mu_r^t)$  be an evolving weighted Riemannian manifold as in Section 3.2.1, with flow maps  $T^{(t)} : M(0) \rightarrow M(t)$  arising from a (possible time-dependent) ODE  $\dot{x} = F(x, t)$ , where  $F(x, t)$  is  $C^\infty$  at each  $x \in M(t)$ . We define a time-continuous Perron-Frobenius operator  $\mathcal{P}^{(t)} : L^1(M(0), \mu_r^0) \rightarrow L^1(M(t), \mu_r^t)$  by  $\int_{M(t)} \mathcal{P}^{(t)} f \cdot \omega_{m^t}^r = \int_{M(0)} f \cdot \omega_{m^0}^r$  for all  $t \in [0, \tau]$ . One now has the time-continuous pushforward operator  $\mathcal{H}^{(t)} : L^2(M(0), m^0, \mu_r^0) \rightarrow L^2(M(t), m^t, \mu_r^t)$  given by

$$\mathcal{H}^{(t)} f := \frac{\mathcal{P}^{(t)}(f \cdot h)}{\mathcal{P}^{(t)} h}, \quad (3.31)$$

for all  $t \in [0, \tau]$ , where  $h$  is the density of the initial measure  $\mu_r^0$ .

Define the time-continuous generalisation of  $\Delta^{dyn}$  as

$$\Delta^{[0, \tau]} f := \frac{1}{\tau} \int_0^\tau (\mathcal{H}^{(t)})^* \Delta_{\mu, t} \mathcal{H}^{(t)} f dt, \quad (3.32)$$

where  $\Delta_{\mu, t}$  is the weighted Laplacian given by (3.27), with respect to the metric  $m^t$  and measure  $\mu_r^t$  for each  $t \in [0, \tau]$ . Furthermore, by a straightforward modification of Corollary B.2.8 in Appendix B, one has

$$(\mathcal{H}^{(t)})^* \Delta_{\mu, t} \mathcal{H}^{(t)} = \Delta_{\bar{\mu}, t},$$

for each  $t \in [0, \tau]$ , where  $\Delta_{\bar{\mu}, t}$  is a weighted Laplacian on  $M$  defined by (3.27) with respect to the metric  $(T^{(t)})^*(m^t)$  and density  $\mathcal{P}^{(t)} h \circ T^{(t)}$ . Hence, one may express (3.32) as

$$\Delta^{[0, \tau]} f = \frac{1}{\tau} \int_0^\tau \Delta_{\bar{\mu}, t} f dt. \quad (3.33)$$

### 3.3.3 Spectral theory and a dynamic Cheeger inequality on weighted manifolds

In standard isoperimetric theory for a compact, connected Riemannian manifold  $M$ , one may use the spectrum of the Laplace-Beltrami operator  $\Delta_m$  to reveal geometric information about  $M$ . Variational properties characterise the spectrum of  $\Delta_m$  (see e.g. p.13 in [22] or p.210 in [88]). Extensions of these variational properties, which carry dynamic information, can be developed for dynamic Laplacian on a compact subset of  $\mathbb{R}^r$ , under volume-preserving dynamics as in Theorem 3.2 in [49]. Here, we generalise Theorem 3.2 in [49] to weighted, non-flat Riemannian manifolds, subjected to non-volume-preserving dynamics.

---

<sup>2</sup>The time-discrete version of (3.28) is constructed similarly compared to the time-continuous version.

**Theorem 3.3.3.** *Let  $(M, m, \mu_r)$  and  $(N, n, \nu_r)$  be weighted Riemannian manifolds, where  $M$  and  $N$  are  $C^\infty$ , compact and connected. Let  $T : M \rightarrow N$  be a  $C^\infty$ -diffeomorphism such that  $\nu_r = \mu_r \circ T^{-1}$ . Define  $\Delta^{\text{dyn}}$  and  $T^*n$  by (3.28) and (3.7) respectively. Consider the eigenvalue problem*

$$\Delta^{\text{dyn}}\phi = \lambda\phi, \quad (3.34)$$

with initial Neumann-type boundary condition

$$m([\nabla_m + \nabla_{T^*n}]\phi, \mathbf{n})_x = 0, \quad \forall x \in \partial M, \quad (3.35)$$

where  $\mathbf{n}$  is the normal bundle along  $\partial M$ . Assume the density of  $\mu_r$  is  $C^\infty$  and uniformly bounded away from zero.

1. The eigenvalues of  $\Delta^{\text{dyn}}$  are nonpositive, real, and form a decreasing sequence  $0 = \lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \dots$  with  $\lambda_k \rightarrow -\infty$ , as  $k \rightarrow \infty$ .
2. The corresponding eigenfunctions  $\phi_1, \phi_2, \dots$  are in  $C^\infty(M, \mathbb{R})$ , the eigenfunction  $\phi_1$  is constant, and eigenfunctions corresponding to distinct eigenvalues are pairwise orthogonal in  $L^2(M, m, \mu_r)$ .
3. Let  $\langle \cdot, \cdot \rangle_\mu$  denote the inner-product on  $L^2(M, m, \mu_r)$ , and  $|\cdot|_m = \sqrt{m(\cdot, \cdot)}$  the norm on tangent spaces induced by the metric tensor  $m$ . Define  $F^0 = L^2(M, m, \mu_r)$  and  $F^k := \{f \in L^2(M, m, \mu_r) : \langle f, \phi_i \rangle_\mu = 0 \text{ for } i = 1, \dots, k\}$ , for  $k = 1, 2, \dots$ , then

$$\lambda_k = - \inf_{f \in F^{k-1}} \frac{\int_M |\nabla_m f|_m^2 d\mu_r + \int_N |\nabla_n \mathcal{H}f|_n^2 d\nu_r}{2 \int_M f^2 d\mu_r} \quad (3.36)$$

$$= - \inf_{f \in F^{k-1}} \frac{\int_M (|\nabla_m f|_m^2 + |\nabla_{T^*n} f|_{T^*n}^2) d\mu_r}{2 \int_M f^2 d\mu_r}, \quad (3.37)$$

where  $\mathcal{H}$  is given by (3.19). Moreover, the infimum of (3.36) is attained by  $f = \phi_k$ .

*Proof.* See Appendix B. □

Equation (3.34) shows that the eigenvalues of  $\Delta^{\text{dyn}}$  take on larger negative values when  $|\nabla_m f|_m$  is large with respect to  $\mu_r$  and  $|\nabla_n \mathcal{H}f|_n$  is large with respect to  $\nu_r$ . To obtain  $\lambda_k$  close to zero, one needs  $f$  and  $\mathcal{H}f$  to have low gradient, and particularly in regions of high  $\mu_r$  and  $\nu_r$  mass, respectively. Compare this to (3.20), (3.22) and (3.23), which make connections with level sets of  $f$  and the pushforward  $\mathcal{H}f$ . Another way to state that  $\lambda_k$  is close to zero is to say that one needs the level sets of  $f$  and  $\mathcal{H}f$  to be not large with respect to  $\mu_{r-1}$  and  $\nu_{r-1}$ , respectively. This

probably means a combination of not being large according to  $\omega_m^{r-1}$  and  $\omega_n^{r-1}$  (e.g. if  $M$  is two-dimensional, the level sets are generally a *small number of short curves* produced by an  $f$  which is not very oscillatory), and avoiding high density areas of  $(M, m, \mu_r)$  and  $(N, n, \nu_r)$ .

The following theorem provides an upper bound on how bad (how large) the average size of an evolving boundary  $\Gamma$  can be; it bounds above the geometric quantity  $\mathbf{H}_M^{dyn}$  in terms of  $\lambda_2$ , the first nontrivial eigenvalue of  $\Delta^{dyn}$ . The classical “static” version of this result, due to Cheeger [24], can be intuitively described in terms of heat flow. Consider heat flow (generated by the Laplace operator) on a solid dumbbell in two dimensions with a narrow neck. By initialising “positive heat” on one side of the dumbbell and “negative heat” on the other side, the rate at which the heat flow equilibrates will be slow because of the narrow neck. The eigenvalue  $\lambda_2$  will be close to zero because of this slow equilibration. Of course, the narrow neck means that it is possible to very cheaply partition the dumbbell  $M$  into two pieces  $M_1, M_2$ , with  $\Gamma$  cutting across the neck. Cheeger showed that a small  $\lambda_2$  implied a small  $\mathbf{H}_M^{dyn}$  (a cheap way of disconnecting  $M$ ). Theorem 3.3.4 injects general nonlinear dynamics into these ideas, and extends Theorem 3.2 of [49] to weighted manifolds and non-volume-preserving dynamics. In terms of heat flow, we are effectively averaging the heat flow geometry across the time duration over which our dynamics acts.

**Theorem 3.3.4 (Dynamic Cheeger inequality).** *Let  $(M, m, \mu_r)$  and  $(N, n, \nu_r)$  be weighted Riemannian manifolds, where  $M$  and  $N$  are  $C^\infty$ , compact and connected. Let  $\Delta^{dyn}$  and  $H_M^{dyn}$  be defined by (3.28) and (3.9) respectively. Assume the density of  $\mu_r$  is  $C^\infty$  and uniformly bounded away from zero. If  $\lambda_2$  is the smallest magnitude nonzero eigenvalue of the eigenproblem (3.34)-(3.35) with eigenfunction  $\phi_2$ , then*

$$\mathbf{H}_M^{dyn} \leq \inf_{\beta \in (-\infty, \infty)} H_M^{dyn}(\{\phi_2 = \beta\}) \leq 2\sqrt{-\lambda_2}. \quad (3.38)$$

*Proof.* See Appendix B. □

By the linearity of our construction with respect to time, it is straightforward to use variational properties to characterise the spectrum of  $\Delta^{[0, \tau]}$  (see (3.33)) as in Theorem 3.3.3. Moreover, by a modification (see Appendix B.6.1 for details), one can obtain a continuous-time dynamic Cheeger inequality

$$\mathbf{H}_M^{[0, \tau]} \leq 2\sqrt{-\lambda_2^{[0, \tau]}},$$

where  $\lambda_2^{[0, \tau]}$  is the second eigenvalue of  $\Delta^{[0, \tau]}$  defined in (3.33). Consequently, one can find good solutions to the continuous-time optimisation problem (3.15) by a process identical to that outlined in Algorithm 3.1 below.

Theorem 3.3.3 and Theorem 3.3.4 suggest the following dynamic spectral partitioning strategy: Let  $(M, m, \mu_r)$  be a weighted Riemannian manifold. Suppose one wishes to find a  $C^\infty$  co-dimension 1 surface  $\Gamma$  in  $M$ , such that  $\Gamma$  partitions  $M = M_1 \cup \Gamma \cup M_2$  and the dynamic Cheeger ratio  $H_M^{dyn}(\Gamma)$  given by (3.9), (3.12), or (3.14) is small. Instead of searching over *all possible* smooth co-dimension 1 surfaces in  $M$  to find the optimal solution for (3.10), one can consider a significantly smaller collection of  $C^\infty$  co-dimension 1 surfaces generated by the eigenfunction  $\phi_2$ , and still find a good solution to (3.10). This is the content of Algorithm 3.1, which is standard in manifold learning and graph partitioning, and is also used in [49, 51].

**Algorithm 3.1:** Dynamic spectral partitioning

- 1 Given a weighted Riemannian manifold  $(M, m, \mu_r)$  and a single iteration of transformation  $T$  on  $M$ , form the weighted Laplacian  $\Delta_\mu$  as in (3.32)
- 2 Solve the eigenvalue problem  $\Delta^{dyn}\phi_2 = \lambda_2\phi_2$ , where  $\lambda_2$  is the first non-trivial eigenvalue of  $\Delta^{dyn}$ , with corresponding  $C^\infty(M)$  eigenfunction  $\phi_2$ .
- 3 For each  $\beta \in [\min \phi_2, \max \phi_2]$ , partition  $M$  into  $M = M_1^\beta \cup \Gamma^\beta \cup M_2^\beta$  via  $M_1^\beta = \{x \in M : \phi_2(x) < \beta\}$ ,  $M_2^\beta = \{x \in M : \phi_2(x) > \beta\}$ , and the  $C^\infty$  hypersurface  $\Gamma^\beta = \{x \in M : \phi_2(x) = \beta\}$ .
- 4 Compute  $H_M^{dyn}(\Gamma^\beta)$  for each  $\beta \in [\min \phi_2, \max \phi_2]$  and extract the optimal  $\beta_0$ ; the hypersurface  $\Gamma^{\beta_0}$  is an approximate solution to the dynamic isoperimetric problem (3.9).

*Remark 3.3.5.* Algorithm 3.1 can be extended to multi-element partitions if one is searching for multiple coherent objects. Early transfer operator based methods (e.g. [32, 34]) proposed the use of a numerical spectral gap as a heuristic for determining the number of almost-invariant sets; that is, a gap between  $\lambda_k$  and  $\lambda_{k+1}$  indicates that  $k$  is a natural number<sup>3</sup> of almost-invariant sets to search for. This idea is commonly used in the transfer operator community and is equally applicable to finite-time coherent sets [48] (where one would look for a gap in the singular value spectrum) and to the dynamic Laplace operator [51]. Such a heuristic has also been used for eigenvalues of (static) Laplace-Beltrami operators and their discrete graph-based counterparts in manifold learning (see e.g. the review [84]), where it is called the eigengap heuristic. Once an estimate of a natural number  $k \geq 1$  of coherent objects has been determined in this way, one embeds the eigenfunctions  $\phi_2, \dots, \phi_{k+1}$

<sup>3</sup>In settings where there is a good functional analytic setup for the transfer operator  $\mathcal{P}$ , one defines the number of almost-invariant (resp. coherent) sets as the number of eigenvalues (resp. Lyapunov exponents) outside the essential spectrum [31] (resp. [55]).

in  $k$ -dimensional Euclidean space, as per e.g. [110]). One can then employ standard clustering methods to identify  $k + 1$  distinct coherent objects  $M_1, \dots, M_{k+1}$ . In the case of weighted manifolds, the balancing of the  $\mu_r$  measures of the sets  $M_1, \dots, M_k$  is important. This could be achieved by, for example, weighted fuzzy clustering, analogous to the algorithm in [47].

### 3.4 Geometry and probability: linking finite-time coherent sets with dynamic isoperimetry

We demonstrate that the probabilistic approach for identifying coherent structures in [48] is tightly connected to the dynamic Laplacian given by (3.28), extending Theorem 5.1 of [49] to the non-volume-preserving, weighted manifold setting. Let  $(M, e, \mu_r)$  and  $(N, e, \nu_r)$  be weighted Riemannian manifolds, where  $\mu_r, \nu_r$  are absolutely continuous probability measures with respect to the Lebesgue measure  $\ell_r$ ,  $e$  the Euclidean metric, and  $M$  a compact,  $r$ -dimensional subset of  $\mathbb{R}^r$ . In [48] (see (3.40)-(3.43) below for the corresponding treatment in  $(M, m, \mu_r)$ ), one applies local diffusion on  $M$ , by locally averaging the functions in  $L^1(M, \ell)$  via the operator  $\mathcal{D}_{X,\epsilon} : L^1(X, \ell_r) \rightarrow L^1(X_\epsilon, \ell_r)$ , where  $X \subset X_\epsilon \subseteq \mathbb{R}^r$ . By composing  $\mathcal{D}_{X,\epsilon}$  with the Perron-Frobenius operator  $\mathcal{P}$ , one obtains an operator  $\mathcal{P}_\epsilon$  that applies local diffusion to  $X$  before and after the application of dynamics. Furthermore, it was shown in [48], that by normalising  $\mathcal{P}_\epsilon$ , one obtains an operator  $\mathcal{H}_\epsilon : L^2(X, \mu_r) \rightarrow L^2(X'_\epsilon, \nu_r)$ , where  $X$  and  $X'_\epsilon$  are subsets of  $\mathbb{R}^r$ ; the operator  $\mathcal{H}_\epsilon$  applies diffusion and advection to the weighted space  $(M, e, \mu_r)$ , where  $e$  is the Euclidean metric on  $\mathbb{R}^r$ . Finally, it was shown in [48] that the leading sub-dominant singular vectors of the operator  $\mathcal{H}_\epsilon$  correspond to finite-time coherent sets.

Theorem 5.1 in [49] states that if  $T$  is volume preserving and  $\ell_r = \mu_r = \nu_r$ , then

$$\lim_{\epsilon \rightarrow 0} \frac{(\mathcal{H}_\epsilon^* \mathcal{H}_\epsilon - \text{Id})f}{\epsilon^2}(x) = \frac{1}{2}(\Delta_e + \mathcal{P}^* \Delta_e \mathcal{P})f(x), \quad (3.39)$$

for all  $x \in M \subset \mathbb{R}^r$ , where  $\mathcal{P}^*$  is the adjoint of the Perron-Frobenius operator  $\mathcal{P}$  with respect to the standard inner-product  $\langle \cdot, \cdot \rangle_e$ ; namely composition with  $T$  (the Koopman operator), and  $\mathcal{H}_\epsilon^*$  is the adjoint of  $\mathcal{H}_\epsilon$  with respect to a weighted inner-product (see (3.44) below).

In the following, we first generalise the above constructions to a weighted Riemannian manifold setting. We then improve the point-wise convergence (3.39) to a uniform convergence over all  $f \in C^3(M, \mathbb{R})$ . Let  $E_\rho(x)$  denote the Euclidean ball, centered at  $x$  with radius  $\rho$ . Define  $q : \mathbb{R}^+ \rightarrow \mathbb{R}$  with support in the open interval

$(0, 1)$ , such that for any vector  $\mathbf{v} = v - \mathbf{0} = (v_1, v_2, \dots, v_r) \in \mathbb{R}^r$

$$\int_{E_1(0)} v_i v_j q(|\mathbf{v}|^2) d\ell(v) = \begin{cases} 0 & \text{if } i \neq j \\ c & \text{if } i = j \end{cases}, \quad (3.40)$$

for some fixed constant  $c$ , and the integral is over all unit vectors in  $\mathbb{R}^r$  centered at the origin. For  $\epsilon > 0$ , let  $Q_{m,\epsilon}(x, z) := \epsilon^{-r/2} q(\text{dist}_m(x, z)/\sqrt{\epsilon})$  be a family of functions, where  $\text{dist}_m$  is the Riemannian distance function on  $M$  with respect to the metric  $m$ . For open subsets  $X \subset X_\epsilon \subseteq M$  and each  $\epsilon > 0$ , define the *diffusion operator*  $\mathcal{D}_{X,\epsilon} : L^1(X, V_m) \rightarrow L^1(X_\epsilon, V_m)$  by

$$\mathcal{D}_{X,\epsilon} f(x) = \int_X Q_{m,\epsilon}(x, y) f(y) \cdot \omega_m^r(y), \quad (3.41)$$

for all  $x \in M$ . If necessary we rescale  $q$  so that for  $\epsilon$  sufficiently small  $\mathcal{D}_{X,\epsilon} \mathcal{I}_X = \mathcal{I}_{X_\epsilon}$ , where  $\mathcal{I}_X : M \rightarrow \{0, 1\}$  is the characteristic function. In particular, we assume  $\lim_{\epsilon \rightarrow 0} \int_0^1 q(v) dv = 1$ . One can interpret  $\mathcal{D}_{X,\epsilon}$  as a mollifier on  $f$ , that averages  $f$  at the point  $x \in X$  over the  $\epsilon$ -neighbourhood of  $x$  according to the distribution  $q$ . Similarly for  $Y'_\epsilon \subset Y_\epsilon \subseteq N$  we define a local diffusion operator  $\mathcal{D}_{Y'_\epsilon,\epsilon} : L^1(Y'_\epsilon, V_n) \rightarrow L^1(Y_\epsilon, V_n)$  by  $\mathcal{D}_{Y'_\epsilon,\epsilon} f(x) := \int_{Y'_\epsilon} Q_{n,\epsilon}(x, y) f(y) \cdot \omega_n^r(y)$ .

Recall the definition of the Perron-Frobenius operator  $\mathcal{P}$  given by (3.17). Set  $Y'_\epsilon = TX_\epsilon$ , one has an advection-diffusion process between  $L^1(X, V_m)$  and  $L^1(Y_\epsilon, V_n)$ , given by the following diagram:

$$L^1(X, V_m) \xrightarrow{\mathcal{D}_{X,\epsilon}} L^1(X_\epsilon, V_m) \xrightarrow{\mathcal{P}} L^1(Y'_\epsilon, V_n) \xrightarrow{\mathcal{D}_{Y'_\epsilon,\epsilon}} L^1(Y_\epsilon, V_n). \quad (3.42)$$

We form  $\mathcal{P}_\epsilon : L^1(X, V_m) \rightarrow L^1(Y_\epsilon, V_n)$  according to (3.42) via the composition  $\mathcal{P}_\epsilon f := \mathcal{D}_{Y'_\epsilon,\epsilon} \circ \mathcal{P} \circ \mathcal{D}_{X,\epsilon} f$ . Normalising  $\mathcal{P}_\epsilon$  yields the operator

$$\mathcal{H}_\epsilon f(y) := \frac{\mathcal{P}_\epsilon(f \cdot h_\mu)}{\mathcal{P}_\epsilon h_\mu} \Big|_y = \int_X \kappa_\epsilon(x, y) f(x) d\mu_r(x), \quad (3.43)$$

where

$$\kappa_\epsilon(x, y) := \frac{\int_{X_\epsilon} Q_{n,\epsilon}(y, Tz) Q_{m,\epsilon}(z, x) \cdot \omega_m^r(z)}{\int_{X_\epsilon} \left( \int_{X_\epsilon} Q_{n,\epsilon}(y, Tz) Q_{m,\epsilon}(z, x) \cdot \omega_m^r(z) \right) d\mu_r(x)}.$$

Let  $h_{\nu_\epsilon} = \mathcal{P}_\epsilon h_\mu$ , and define  $\nu_{r,\epsilon} := dh_{\nu_\epsilon}/dV_n$ . If  $\kappa_\epsilon(x, y) \in L^2(X \times Y_\epsilon, \mu_r \times \nu_{r,\epsilon})$  then  $\mathcal{H}_\epsilon : L^2(X, \mu_r) \rightarrow L^2(Y_\epsilon, \nu_{r,\epsilon})$  is compact (by Lemma 1 in [48]).

By obvious modification of the arguments in [48], one can verify that the adjoint operator  $\mathcal{H}_\epsilon^* : L^2(Y_\epsilon, \nu_{r,\epsilon}) \rightarrow L^2(X, \mu_r)$  is given by the composition

$$\mathcal{H}_\epsilon^* g = \mathcal{D}_{X,\epsilon}^* \circ \mathcal{H}^* \circ \mathcal{D}_{Y'_\epsilon,\epsilon}^* g. \quad (3.44)$$

Note that for small  $\epsilon$ ,  $\mathcal{H}_\epsilon \mathcal{I}_X = \mathcal{I}_{Y_\epsilon}$  and  $\mathcal{H}_\epsilon^* \mathcal{I}_{Y_\epsilon} = \mathcal{I}_X$ , the leading singular values  $\mathcal{H}_\epsilon$  approaches 1 as  $\epsilon \rightarrow 0$ , with corresponding left and right singular vectors  $\mathcal{I}_X$  and  $\mathcal{I}_{Y_\epsilon}$  (by Proposition 2 in [48]).

By construction, with a suitable choice for  $q$  the leading singular value of  $\mathcal{H}_\epsilon$  is always 1, and the second leading singular vector of  $\mathcal{H}_\epsilon$  is used to partition  $X \subset M$  into finite-time coherent sets in [48]. The operator  $\mathcal{H}_\epsilon$  applies local diffusion on  $X \subset M$ , before and after  $X$  is transformed into  $Y_\epsilon$  under the action  $T$ . Similarly, the operator  $\mathcal{H}_\epsilon^*$  applies local diffusion on  $Y_\epsilon$ , before and after  $Y_\epsilon \subset N$  is pulled-back to  $X \subset M$  under  $T$ . Therefore, if  $X$  contains finite-time coherent sets (and  $Y_\epsilon$  contains their images), then there will be a tendency for the boundaries of these coherent sets to be small both before and after advection in order to minimise diffusive mixing through their boundaries. The reason for adding diffusion is to give compactness of  $\mathcal{H}_\epsilon$  acting on  $L^2$ , ensuring the singular values of  $\mathcal{H}_\epsilon$  close to 1 are isolated, and to detect subsets of  $X$  and  $Y_\epsilon$  that have small boundary both before and after the application of  $T$ ; see Section 4 in [48] for details.

An interesting question is “what happens in the limit  $\epsilon \rightarrow 0$ ?” The composition  $\mathcal{H}_\epsilon^* \mathcal{H}_\epsilon$  is approximately the identity for small  $\epsilon$ , which appears to provide no dynamical information. However, by subtracting the identity and rescaling by  $\epsilon$ , one can extract the next term in an  $\epsilon$ -expansion of  $\mathcal{H}_\epsilon^* \mathcal{H}_\epsilon$ . The following result generalises Theorem 5.1 in [49] for  $\mathbb{R}^r$  to the case of non-flat weighted Riemannian manifolds; subjected to non-volume-preserving dynamics.

**Theorem 3.4.1.** *Let  $(M, m, \mu_r)$  and  $(N, n, \nu_r)$  be weighted Riemannian manifolds, where  $M$  and  $N$  are  $C^\infty$ , compact and connected. Let  $T : M \rightarrow N$  be a  $C^\infty$  diffeomorphism. Assume  $\nu_r = \mu_r \circ T^{-1}$ , and the density of  $\mu_r$  is  $C^3$ . Define  $\Delta^{dyn}$  by (3.28), and  $\mathcal{H}_\epsilon$  and its adjoint  $\mathcal{H}_\epsilon^*$  by (3.43) and (3.44) respectively. There exists a constant  $c$  such that*

$$\lim_{\epsilon \rightarrow 0} \left( \sup_{\|f\|_{C^3(M, \mathbb{R})} \leq 1} \left\| \frac{(\mathcal{H}_\epsilon^* \mathcal{H}_\epsilon - \text{Id})f}{\epsilon} - c \cdot \Delta^{dyn} f \right\|_{C^0(M, \mathbb{R})} \right) = 0, \quad (3.45)$$

where the constant  $c$  is as in (3.40).

*Proof.* See Appendix B. □

As in the analogous result for small magnitude diffusion presented in Theorem 5.1 [49], one now has a geometric interpretation of finite-time coherent sets considered in [48]. Due to Theorem 3.4.1, given  $\epsilon$  sufficiently small, the action of the operator  $\mathcal{H}_\epsilon^* \mathcal{H}_\epsilon - I$  is approximated by the action of the dynamic Laplacian  $\Delta^{dyn}$ . Thus, one has a dual interpretation of finite-time coherent sets as defined probabilistically in [48] to minimise global mixing (including now in the weighted, non-volume-preserving situation), and as defined geometrically in [49] and the present chapter using the notion of dynamical isoperimetry to force small boundary size

under nonlinear dynamics. A similar result has also been proven in [9] for a slightly differently defined  $\Delta^{dyn}$ .

## 3.5 Numerical experiments

In this section we use Theorems 3.3.3 and 3.3.4 to compute solutions to the dynamic isoperimetric problem (3.9). Our examples will showcase Lagrangian coherent structures on weighted domains with non-volume-preserving dynamics. To keep the numeric simple, we do not explicitly model curvature and limit the dimension in the examples to 2. However, numerical examples in 3 dimensions are complete feasible. We consider 2-dimensional weighted, flat Riemannian manifolds  $(M, e, \mu_2)$  and  $(N, e, \nu_2)$ , where  $M$  and  $N$  are 2-dimensional compact subsets of  $\mathbb{R}^2$ , and  $e$  is the Euclidean metric. We consider measures  $\mu_2$  with smooth densities  $h_\mu$  that are uniformly bounded away from zero, and nonlinear dynamics  $T : M \rightarrow N$  such that  $\nu_2 = \mu_2 \circ T^{-1}$ . Before we give the specific details on the 2-dimensional weighted Riemannian manifolds  $(M, e, \mu_2)$ ,  $(N, e, \nu_2)$  and the transformations  $T$ , we outline the numerical discretisation of the weighted Laplacian  $\Delta^{dyn}$  defined by (3.28) and the operator  $\mathcal{H}$ . We have employed a very simple low-order method, but in principle any standard operator approximation method can be used instead.

### 3.5.1 Numerical approximation for $\mathcal{H}$ and $\mathcal{H}^*$

To obtain a numerical approximation for  $\mathcal{H}$ , we start with tracking the time evolution of the density  $h_\mu$  under  $T$ . To achieve this, we numerically estimate the Perron-Frobenius operator  $\mathcal{P}$  using Ulam's method [120]. We follow the construction of [58]: partition  $M$  and  $N$  into the collections of small boxes  $\{B_1, \dots, B_I\}$  and  $\{C_1, \dots, C_J\}$  respectively, and let  $\mathbf{T}$  be the transition matrix of volume transport between the boxes in  $M$  and boxes in  $N$  under the action of  $T$ . Let  $z_{i,k}$ ,  $k = 1, \dots, K_i$  be  $K_i$  uniformly distributed test points in the box  $B_i$ . We numerically estimate the entries of  $\mathbf{T}$  by computing

$$\mathbf{T}_{ij} = \frac{\#\{z_{i,q} \in B_i : T(z_{i,q}) \in C_j\}}{K_i}. \quad (3.46)$$

The matrix  $\mathbf{T}$  is a row-stochastic matrix, where the  $(i, j)^{th}$  entry estimates the conditional probability of a randomly chosen point in  $B_i$  entering  $C_j$  under the application of  $T$ . The connection between the matrix  $\mathbf{T}$  and the operator  $\mathcal{P}$  is as follows. Denote by  $\pi_I : L^1(M, e, \nu_m) \rightarrow \text{span}\{\mathcal{I}_{B_1}, \dots, \mathcal{I}_{B_I}\}$  and  $\theta_J : L^1(N, e, \nu_r) \rightarrow \text{span}\{\mathcal{I}_{C_1}, \dots, \mathcal{I}_{C_J}\}$  the orthogonal Ulam projections formed by taking expectations on partition elements. Define  $\mathcal{P}_{I,J} := \theta_J \circ \mathcal{P}$ . One has  $\mathcal{P}_{I,J} : \text{span}\{\mathcal{I}_{B_1}, \dots, \mathcal{I}_{B_I}\} \rightarrow$

$\text{span}\{\mathcal{I}_{C_1}, \dots, \mathcal{I}_{C_J}\}$ , so that  $\mathbf{T}$  is the matrix representation of  $\mathcal{P}_{I,J}$  under left multiplication.

We discretise the density  $h_\mu$  of  $\mu_r$  to a column vector  $\mathbf{u}$  of length  $I$ , by setting  $u_i = \mu_r(B_i)$ . If some sets  $B_i$  have zero  $\mu_r$ -measure, then we remove them from our collection as there is no mass to be transported. We therefore assume that  $u_i > 0$  for all  $i = 1, \dots, I$ . To approximate the density  $h_\nu$  of  $\nu_r$ , we use the fact that  $h_\nu = \mathcal{P}h_\mu$  (by (3.18)). Thus  $\mathbf{v} = \mathbf{T}^\top \mathbf{u}$  is the numerical approximation of  $h_\nu$ . We assume  $v_j > 0$  (if  $v_j = 0$ , then we remove the corresponding sets  $C_j$  because they represent  $\nu_r(C_j) = 0$ ).

To numerically estimate  $\mathcal{H}$  given by (3.19), we use the matrix  $\mathbf{T}$  and the vectors  $\mathbf{u}$  and  $\mathbf{v}$ . In particular, the components of  $\theta_J(\mathcal{H}f)$  are approximated by

$$[\mathcal{H}f]_j \approx \sum_{i=1}^I \frac{\mathbf{T}_{ji}(f_i u_i)}{v_j}, \quad (3.47)$$

where  $f_i$  are the components of the vector  $\mathbf{f} := \pi_I f$ . Define the  $I \times J$  matrix  $\tilde{\mathbf{T}}$  by

$$\tilde{\mathbf{T}}_{ij} := \mathbf{T}_{ij} u_i / v_j. \quad (3.48)$$

Then (3.47) is equivalent to  $\theta_J(\mathcal{H}f) \approx \tilde{\mathbf{T}}^\top \mathbf{f}$ ; that is the matrix  $\tilde{\mathbf{T}}$  under left multiplication is the numerical approximation of  $\mathcal{H}$ . To numerically estimate  $\mathcal{H}^*$  from  $\mathcal{H}$ , we note by definition  $\langle \mathcal{H}f, g \rangle_\mu = \langle f, \mathcal{H}^*g \rangle_\nu$ , for all  $f \in L^2(M, m, \mu_r)$  and  $g \in L^2(N, n, \nu_r)$ . Hence,

$$\sum_{i=1}^I f_i \cdot [\mathcal{H}^*g]_i \cdot u_i \approx \langle f, \mathcal{H}^*g \rangle_\mu = \langle \mathcal{H}f, g \rangle_\nu \approx \sum_{i=1}^I \sum_{j=1}^J \tilde{\mathbf{T}}_{ij} f_i \cdot g_j \cdot v_j = \sum_{i=1}^I f_i \cdot \sum_{j=1}^J \mathbf{T}_{ij} g_j \cdot u_i,$$

where  $[\mathcal{H}^*g]_i$  and  $g_j$  are the components of the vectors  $\pi_I(\mathcal{H}f)$  and  $\theta_J g$  respectively. Therefore, we have

$$[\mathcal{H}^*g]_i \approx \sum_{j=1}^J \mathbf{T}_{ij} g_j. \quad (3.49)$$

The operator  $\mathcal{H}^*$  is numerically estimated by the matrix  $\mathbf{T}$  under right multiplication.

### 3.5.2 Finite-difference estimate for $\Delta^{dyn}$

To numerically solve the eigenvalue problem  $\Delta_\mu f = \lambda f$  on  $(M, e, \mu_r)$ , we discretise  $\Delta_\mu$  using the second equality of (3.27); that is

$$\Delta_\mu f = \frac{1}{h_\mu} \text{div}_e(h_\mu \nabla_e f). \quad (3.50)$$

In preparation for the numerical approximations for our 2-dimensional examples, which will be a rectangle, cylinder or torus, we construct a  $K$  by  $L$  grid system for  $M$ . Let  $(x_1, x_2)$  be Euclidean coordinates on  $M$ . We cover  $M$  with  $I$  grid boxes  $\{B_i\}_{i=1}^I$  of uniform size  $b_{x_1} \times b_{x_2}$  (one can easily consider the more general case of nonuniform box sizes), and re-index the boxes  $\{B_i\}_{i=1}^I$  to  $\{B_{k,l}\}_{1 \leq k \leq K, 1 \leq l \leq L}$ , indexing the  $x_1$ -direction with  $k$ , and the  $x_2$ -direction with  $l$ ; clearly  $K \times L = I$ . The re-indexing of the grid boxes  $B_{k,l}$  induces a re-index of  $\mathbf{f}$  via

$$\mathbf{f} := (f_{1,1}, f_{2,1}, \dots, f_{K,1}, f_{1,2}, f_{2,2}, \dots, f_{1,L}, f_{2,L}, \dots, f_{K,L}), \quad (3.51)$$

and similar re-index of  $\mathbf{u}$  and  $\mathbf{v}$ .

We employ standard finite-difference schemes to obtain numerical approximations for the RHS of (3.50). Starting with the approximation of  $h_\mu \nabla_e f$ , one has in Euclidean coordinates  $(x_1, x_2)$ , the vector  $h_\mu \nabla_e f = h_\mu (\partial f / \partial x_1, \partial f / \partial x_2)$ . To compute the derivatives  $\partial f / \partial x_1$  and  $\partial f / \partial x_2$  numerically, we apply the standard central-difference technique to obtain on the grid box  $B_{k,l}$ ,

$$\frac{\partial f}{\partial x_1} \approx \frac{f_{k+1,l} - f_{k-1,l}}{2b_{x_1}} \quad \text{and} \quad \frac{\partial f}{\partial x_2} \approx \frac{f_{k,l+1} - f_{k,l-1}}{2b_{x_2}},$$

thus on the grid box  $B_{k,l}$

$$h_\mu \nabla_e f \approx \left( u_{k,l} \frac{f_{k+1,l} - f_{k-1,l}}{2b_{x_1}}, u_{k,l} \frac{f_{k,l+1} - f_{k,l-1}}{2b_{x_2}} \right). \quad (3.52)$$

Next, we numerically solve the divergence  $\text{div}_e$  applied to the RHS of (3.52). By central-difference approximations, one has on the grid box  $B_{k,l}$

$$\begin{aligned} \Delta_\mu f = \frac{1}{h_\mu} (\text{div}_e(h_\mu \nabla_e f)) &\approx \frac{1}{u_{k,l}} \left[ u_{k+1,l} \frac{f_{k+2,l} - f_{k,l}}{4b_{x_1}^2} - u_{k-1,l} \frac{f_{k-2,l} - f_{k,l}}{4b_{x_1}^2} \right. \\ &\quad \left. + u_{k,l+1} \frac{f_{k,l+2} - f_{k,l}}{4b_{x_2}^2} - u_{k,l-1} \frac{f_{k,l} - f_{k,l-2}}{4b_{x_2}^2} \right]. \end{aligned} \quad (3.53)$$

Denote the resulting finite-difference approximation of  $\Delta_\mu$  by the  $I \times I$  matrix  $\mathbf{\Delta}_\mu$ . Rearranging (3.53), then  $\mathbf{\Delta}_\mu$  applied to  $\mathbf{f}$  is a vector of length  $I$  with components

$$\begin{aligned} [\mathbf{\Delta}_\mu \mathbf{f}]_{k+K(l-1)} &= \frac{1}{4b_{x_1}^2} \frac{u_{k+1,l}}{u_{k,l}} f_{k+2,l} + \frac{1}{4b_{x_1}^2} \frac{u_{k-1,l}}{u_{k,l}} f_{k-2,l} + \frac{1}{4b_{x_2}^2} \frac{u_{k,l+1}}{u_{k,l}} f_{k,l+2} \\ &\quad + \frac{1}{4b_{x_2}^2} \frac{u_{k,l-1}}{u_{k,l}} f_{k,l-2} - \left( \frac{1}{4b_{x_1}^2} \frac{u_{k+1,l} + u_{k-1,l}}{u_{k,l}} + \frac{1}{4b_{x_2}^2} \frac{u_{k,l+1} + u_{k,l-1}}{u_{k,l}} \right) f_{k,l}, \end{aligned} \quad (3.54)$$

for  $1 \leq k \leq K$ ,  $1 \leq l \leq L$ . Note that if  $u_{k,l}$  is constant for all  $1 \leq k \leq K$  and  $1 \leq l \leq L$ , then the expression (3.54) becomes the standard 5-point stencil Laplace matrix. Moreover, as in the standard 5-point stencil approximations, the error of the approximate (3.53) is  $\max\{\mathcal{O}(b_{x_1}^4), \mathcal{O}(b_{x_2}^4)\}$ .

To treat the numerical approximation of  $\Delta_\mu$  at the boundary of  $M$ , we apply the usual Neumann boundary condition  $e(\nabla_e \varphi, \mathbf{n})_x = 0$  for all  $x \in \partial M$  (where  $\mathbf{n}$  is unit normal to  $\partial M$ ). This Neumann boundary condition is imposed by symmetric reflection [116] in the above modified finite-difference scheme as follows: Consider the grid boxes  $B_{1,l}$  for  $1 \leq l \leq L$ ; one has a boundary on the left side edge of each of these grid boxes. By construction, the unit normal  $\mathbf{n}$  along the left side edge of the grid boxes  $\{B_{1,l}\}_{l=1}^L$  is given by  $(-1, 0)$ . Therefore, the boundary condition  $e(\nabla_e \varphi, \mathbf{n})_x = 0$  is satisfied by reflecting the artificial  $f_{0,l} = f_{2,l}$ ,  $f_{-1,l} = f_{1,l}$  and  $u_{0,l} = u_{2,l}$  for all  $1 \leq l \leq L$ . One applies similar symmetric reflections to all  $B_{k,l}$  at the boundary of  $M$ .

By definition (3.28), and the numerical approximations we obtained for  $\Delta_\mu$ ,  $\Delta_\nu$ ,  $\mathcal{H}$  and  $\mathcal{H}^*$ , one has the finite-difference approximation for the weighted dynamic Laplacian  $\Delta^{dyn}$  given by

$$\Delta^{dyn} = \Delta_\mu + \mathbf{T} \Delta_\nu \tilde{\mathbf{T}}^\top, \quad (3.55)$$

where the matrices  $\mathbf{T}$  and  $\tilde{\mathbf{T}}$  are given by (3.46) and (3.48) respectively, and  $\Delta_\mu$ ,  $\Delta_\nu$  by (3.54). We note that the matrices  $\Delta_\mu$ ,  $\Delta_\nu$ ,  $\mathbf{T}$  and  $\tilde{\mathbf{T}}$  are sparse and consequently  $\Delta^{dyn}$  is sparse. One can numerically solve the finite dimensional eigenvalue problem  $\Delta^{dyn} \mathbf{f} = \lambda \mathbf{f}$  for small eigenvalues  $\lambda$ , and in particular  $\lambda_2$  and corresponding eigenfunction  $\phi_2$ . To find a good solution  $\Gamma$  to the dynamic isoperimetric problem (3.9), one can use the level sets of  $\phi_2$  as candidates for  $\Gamma$  as in Algorithm 3.1.

### 3.5.3 Case study 1: dynamics on a cylinder

We now demonstrate our technique on a weighted 2-dimensional cylinder  $(M, e, \mu_2)$ , where  $M = [0, 4] / \sim \times [0, 1]$  and  $h_\mu(x_1, x_2) = \frac{1}{8}(\sin(\pi x_1) + 2)$  as in Section 3.1.1. We set our computational resolution for  $M$  to be  $K \times L = 256 \times 64$  square grid boxes  $B_{k,l}$  of side length  $b = 1/64$ , and select the number of test points in each grid box to be  $Q = 400^4$ . We consider two different types of nonlinear transformations  $T_1$  and  $T_2$  acting on  $M$ :

$$T_1(x_1, x_2) = \left( x_1 + \frac{\cosh(2x_2) - 1}{2}, x_2 \right), \quad (3.56)$$

$$T_2(x_1, x_2) = (x_1 + x_2, x_2 + 0.1x_2 \sin(2\pi x_2)), \quad (3.57)$$

where the first coordinate is computed modulo 4 in both cases. The map  $T_1$  is the area-preserving, nonlinear horizontal shear from the example considered in Section 3.1.1. The map  $T_2$  is a linear horizontal shear, composed with vertical area-

<sup>4</sup>One can also use as few as 25 points per box and still obtain good results.

distortion; i.e.

$$T_2(x_1, x_2) = \hat{T}_2(x_1 + x_2, x_2),$$

where

$$\hat{T}_2(x_1, x_2) = (x_1, x_2 + 0.1x_2 \sin(2\pi x_2)),$$

compresses the mass distribution of  $M$  in towards the horizontal line  $x_2 = 0.5$ .

### The transformation $T_1$ on $M$

We optimally partition  $M$  using Algorithm 3.1 to find a good solution  $\Gamma$  to the dynamic isoperimetric problem (3.9). First, we consider the dynamic Laplacian for  $T_1$  acting on  $M$ . In step 1 of Algorithm 3.1, we construct the matrix  $\mathbf{\Delta}^{dyn}$  given by (3.55) as the numerical approximation of  $\Delta^{dyn}$  via the finite-difference scheme outlined in Section 3.5.2, and numerically solve the finite-dimensional eigenproblem  $\mathbf{\Delta}^{dyn}\phi = \lambda\phi$ . The leading numerical eigenvalues  $\lambda_1, \lambda_2, \dots, \lambda_7$  of  $\mathbf{\Delta}^{dyn}$  are  $0, -0.6046, -1.3739, -2.3221, -3.2886, -3.4091, -3.7056 \dots$ . The components of the numerical eigenvector  $\phi_2$  corresponding to  $\lambda_2$  takes on at most  $256 \times 64$  unique values; at most one value on each of the grid box. Step 2 of Algorithm 3.1, generates partitions of  $M = M_1^\beta \cup \Gamma^\beta \cup M_2^\beta$  from level sets of  $\phi_2$ . Finally, one computes  $H_M^{dyn}(\Gamma^\beta)$  for each  $t$ , and finds the optimal  $\Gamma^{\beta_0}$  as a solution to the dynamic optimisation problem (3.9); the results are shown in Figure 3.2.

It was found that the hypersurface  $\Gamma^{\beta_0}$  is  $H_M^{dyn}$  *minimising* for  $\beta_0 = -1.211 \times 10^{-5}$ ; see figures 3.2a and 3.2c. Note that the densities  $h_\mu$  form a region of low  $\mu_1$ -mass in  $M$  about the lines  $\{x \in M : x_1 = 1.5, 3.5\}$ . Thus, to minimise the  $\mu_1$ -mass of the hypersurface  $\Gamma^{\beta_0}$  in  $M$ , it is advantageous to have  $\Gamma^{\beta_0}$  as short curves in close proximity to the vertical lines  $\{x \in M : x_1 = 1.5, 3.5\}$ . Moreover, to effectively counter the shearing imposed by  $T_1$  so that the size of  $\nu_1$ -mass of  $T_1\Gamma^{\beta_0}$  stays persistently small in  $TM$ , the curve  $\Gamma^{\beta_0}$  bends horizontally towards the left progressively more as  $x_2$  approaches 1 from 0. The  $\mu_1$ -mass of  $\Gamma^{\beta_0}$  and its image under  $T_1$  are  $\mu_1(\Gamma^{\beta_0}) = 0.3088$  and  $\nu_1(T_1\Gamma^{\beta_0}) = 0.3815$ ; the level surface  $\Gamma^{\beta_0}$  experiences significantly reduced deformation under the action of  $T_1$ , compared to the results of Section 3.1.1 shown by Figure 3.1b. Moreover, the partition  $M = M_1^{\beta_0} \cup \Gamma^{\beta_0} \cup M_2^{\beta_0}$  has a perfectly balanced  $\mu_2$ -mass distribution between  $M_1^{\beta_0}$  and  $M_2^{\beta_0}$ . One has  $H_M^{dyn}(\Gamma^{\beta_0}) = 0.6903$ , and this solution is a suitable candidate for LCSs on  $(M, e, \mu_2)$ .

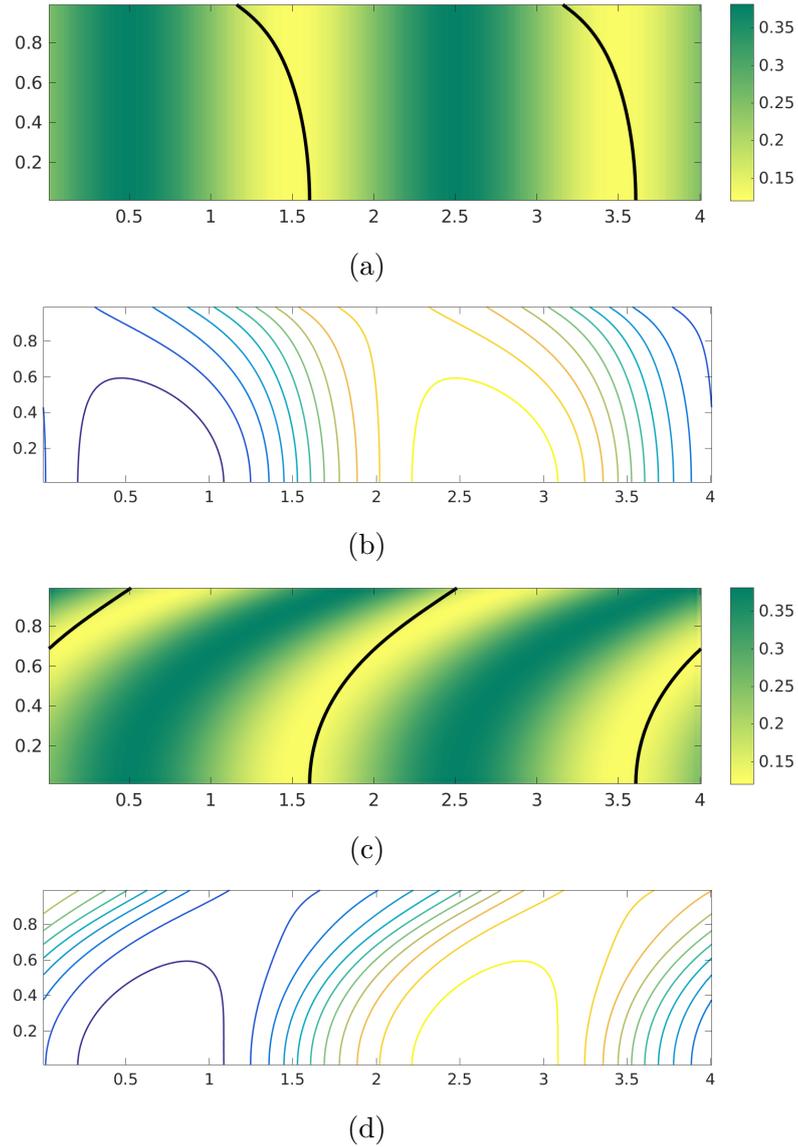


Figure 3.2: Partition of  $M$  using the eigenvector  $\phi_2$  of  $\Delta^{dyn}$  under nonlinear shear  $T_1$  given by (3.56). (a) Colours are values of  $h_\mu$ , and black lines are the level surface  $\Gamma^{\beta_0} = \{\phi_2 = -1.211 \times 10^{-5}\}$ . (b) The level surfaces of  $\phi_2$ . (c) Colours are the values of  $h_\nu$ , and black lines are the level surface  $T_1\Gamma^{\beta_0}$ . (d) The level surfaces of  $\mathcal{H}\phi_2$ .

### The transformation $T_2$ on $M$

We repeat the above numerical experiment on the 2-cylinder, replacing the transformation  $T_1$  with  $T_2$ , and setting the initial mass density  $h_\mu$  to be uniformly distributed on  $M$ . The leading numerical eigenvalues  $\lambda_1, \lambda_2, \dots, \lambda_6$  of  $\Delta^{dyn}$  are  $0, -0.7747 \pm 0.0092i, -3.0900 \pm 0.0199i, 3.8702, -4.5674 \pm 0.0250i$ . In this example, although  $\Delta_\mu$  and  $\Delta_\nu$  have real eigenvalues to numerical precision, when combined to form  $\Delta^{dyn}$ , one obtains small imaginary parts. The eigenvalues  $\lambda_2, \lambda_3$  should be real and equal (i.e.  $\lambda_2$  has multiplicity 2), because of the symmetry obtained by trans-

lating all objects in the  $x$ -coordinate direction. As before, we apply Algorithm 3.1 to partition  $M$  using the level surfaces of the second eigenfunction  $\phi_2$  corresponding to  $\lambda_2 = -0.7747$ ; the results are shown in Figure 3.3.

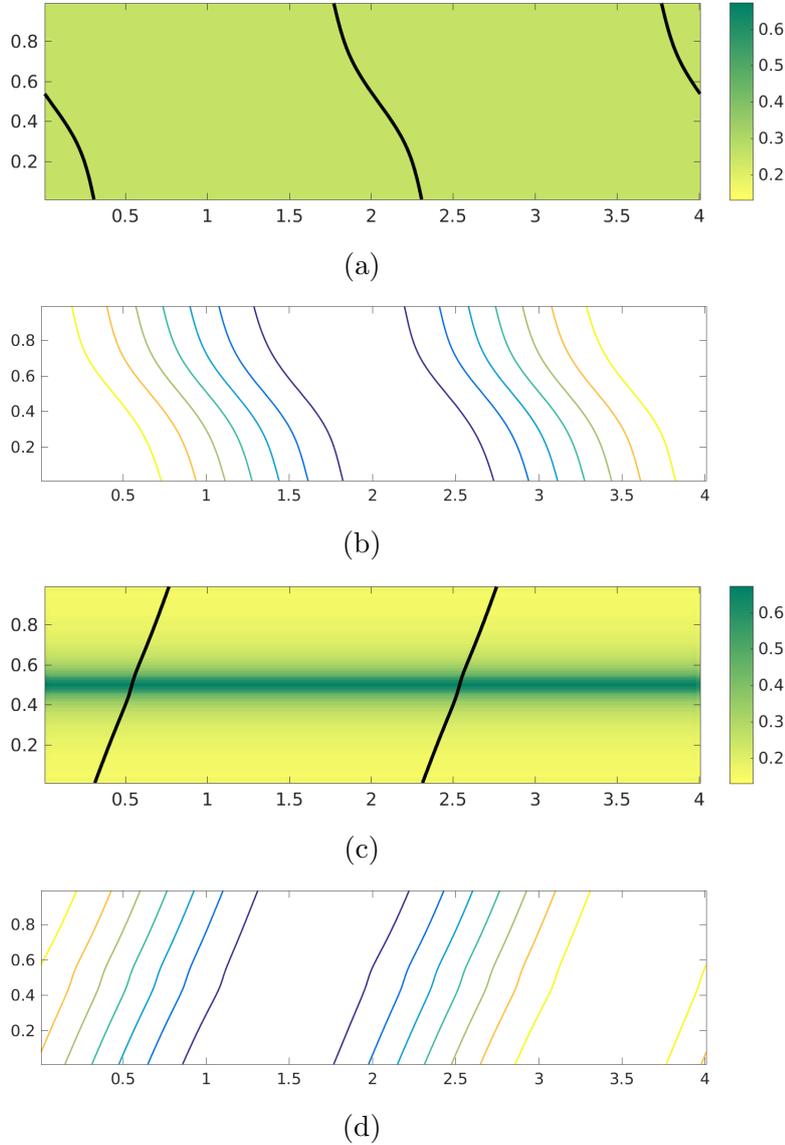


Figure 3.3: Partition of  $M$  using the eigenvector  $\phi_2$  of  $\Delta^{dyn}$  for the nonlinear shear  $T_2$  given by (3.57). (a) Colours are values of  $h_\mu$ , and black lines are the level surface  $\Gamma^{\beta_0} = \{\phi_2 = -3.4232 \times 10^{-5}\}$ . (b) The level surfaces of  $\phi_2$ . (c) Colours are the values of  $h_\nu$ , and black lines are the level surface  $T_2\Gamma^{\beta_0}$ . (d) The level surfaces of  $\mathcal{H}\phi_2$ .

It was found that the hypersurface  $\Gamma^{\beta_0}$  is  $H_M^{dyn}$  *minimising* for  $\beta_0 = -3.4232 \times 10^{-5}$ ; see figures 3.3a and 3.3c. The co-dimension 1 mass of  $\Gamma^{\beta_0}$  and its image under  $T_2$  are  $\mu_1(\Gamma^{\beta_0}) = 0.5682$  and  $\nu_1(T_2\Gamma^{\beta_0}) = 0.5435$ , respectively. Recall that the action of  $T_2$  on  $M$  has the effect of compressing the mass distribution towards the horizontal

line  $x_2 = 0.5$ . To avoid a large  $\nu_1$ -mass of  $T_2\Gamma^{\beta_0}$  on  $T_2M$ , one makes appropriate compromises on the  $\mu_1$ -mass of  $\Gamma^{\beta_0}$  in  $M$ . For example, in Figure 3.3c as the black curves approach the dark green, high density region, they become more vertical so as to traverse this high density region using a shorter curve length and reducing their  $\nu_1$ -mass. This necessitates  $\Gamma^{\beta_0}$  being slightly curved and having slightly greater  $\mu_1$ -mass. Once again the partition  $M = M_1^{\beta_0} \cup \Gamma^{\beta_0} \cup M_2^{\beta_0}$  has a perfectly balanced  $\mu_2$ -mass distribution of  $M_1^{\beta_0}$  and  $M_2^{\beta_0}$ . One has  $H_M^{dyn}(\Gamma^{\beta_0}) = 1.1117$ .

### 3.5.4 Case study 2: dynamics on a torus

Next we demonstrate our technique on a weighted 2-dimensional torus  $(\mathbb{T}^2, e, \mu_2)$ , where  $\mathbb{T}^2 = 2\pi(\mathbb{R}/\mathbb{Z}) \times 2\pi(\mathbb{R}/\mathbb{Z})$  and  $h_\mu(x_1, x_2) = \frac{1}{8\pi^2}(\sin(x_2 - \pi/2) + 2)$ . We set our computational resolution for  $M = \mathbb{T}^2$ , to be  $K \times L = 128 \times 128$  square grid boxes  $B_{k,l}$  of side length  $b = 1/64$ , and select the number of test points in each grid box to be  $Q = 400$ . We consider the transformation  $T := T_4 \circ T_3$  acting on  $M$ , where

$$T_3(x_1, x_2) = (x_1 + 0.3 \cos(2x_1), x_2), \quad (3.58)$$

$$T_4(x_1, x_2) = (x_1 + x_2, x_2 + 8 \sin(x_1 + x_2)), \quad (3.59)$$

computed modulo  $2\pi$ . The map  $T_3$  distorts the area of  $\mathbb{T}^2$  in the horizontal direction, and  $T_4$  is the ‘‘standard map’’.

We optimally partition  $M$  using Algorithm 3.1. The leading numerical eigenvalues  $\lambda_1, \lambda_2, \dots, \lambda_7$  of  $\Delta^{dyn}$  are 0,  $-0.3584$ ,  $-0.3751$ ,  $-1.0750$ ,  $-1.1349$ ,  $-1.4358$ ,  $-1.4966$ . We generate partitions of  $M$  using the level surfaces of the second eigenfunction  $\phi_2$  corresponding to  $\lambda_2 = -0.3584$ ; the results are shown in Figure 3.4.

It was found that the hypersurface  $\Gamma^{\beta_0}$  is  $H_M^{dyn}$  *minimising* for  $\beta_0 = -4.5492 \times 10^{-5}$ ; see figures 3.4a and 3.4c. The  $\mu_1$ -mass on  $\Gamma^{\beta_0}$  and the  $\nu_1$ -mass of its image under  $T$  are  $\mu_1(\Gamma^{\beta_0}) = 0.4584$  and  $\nu_1(T\Gamma^{\beta_0}) = 0.2375$ , respectively. Similar to the results of the previous case study of  $T_2$  acting on  $M$ , one makes appropriate compromises on the  $\mu_1$ -mass of  $\Gamma^{\beta_0}$  in  $\mathbb{T}^2$ , to ensure that the  $\nu_1$ -mass of  $T\Gamma^{\beta_0}$  in  $\mathbb{T}^2$  remains small. For example, in Figure 3.4c the black (almost straight) curves attempt to follow the yellow, low density regions, but when they have to cross the dark green, high density regions, the curves briefly turn to cross these high density regions at a sharper angle. While slightly increasing the curve length, this behaviour reduces the  $\nu_1$ -mass of the curves.

The partition  $\mathbb{T}^2 = M_1^{\beta_0} \cup \Gamma^{\beta_0} \cup M_2^{\beta_0}$  is almost perfectly balanced, with  $\mu_2(M_1) = 0.49994$  and  $\mu_2(M_2) = 0.50006$ . One has  $H_M^{dyn}(\Gamma^{\beta_0}) = 0.6968$ . Despite the highly non-linear nature of  $T$ , evident in the distribution of  $h_\nu$ , shown in Figure 3.4c, one can

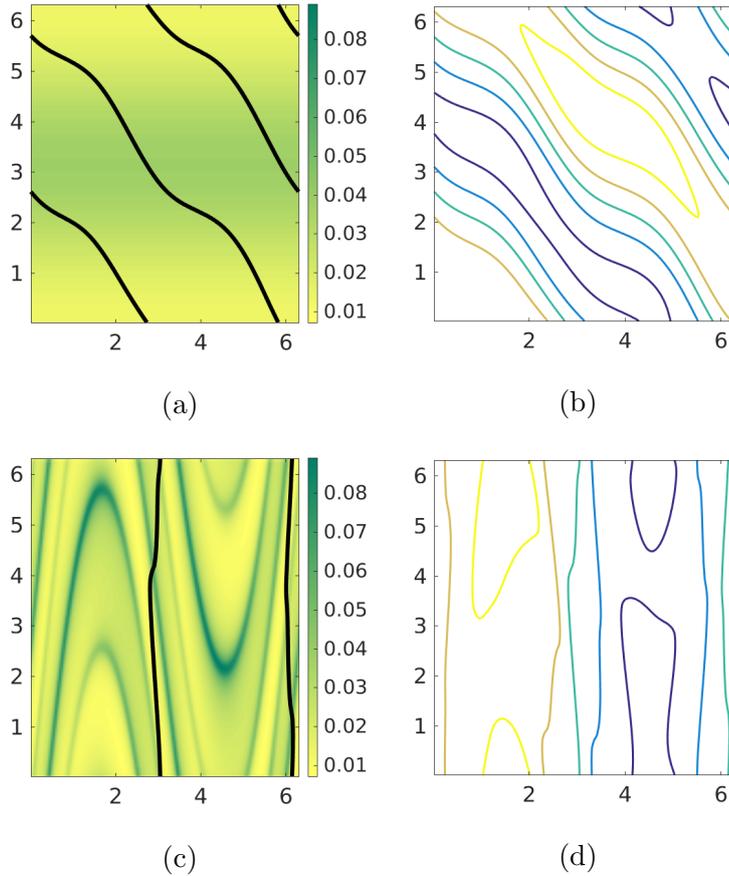


Figure 3.4: Partition of  $\mathbb{T}^2$  using the eigenvector  $\phi_2$  of  $\Delta^{dyn}$  under  $T = T_4 \circ T_3$  given by (3.58) and (3.59). (a) Colours are values of  $h_\mu$ , and black lines are the level surface  $\Gamma^{\beta_0} = \{\phi_2 = -4.5492 \times 10^{-5}\}$ . (b) The level surfaces of  $\phi_2$ . (c) Colours are the values of  $h_\nu$ , and black lines are the level surface  $T\Gamma^{\beta_0}$ . (d) The level surfaces of  $\mathcal{H}\phi_2$ .

find curves that are rather short according to  $\nu_1$ , both before and after the application of  $T$ . Therefore this solution is a suitable candidate for LCSs on  $(\mathbb{T}^2, e, \mu_2)$ , for the finite-time (single) application of  $T$ .

We remark that the curves  $\Gamma^{\beta_0}$  and  $T\Gamma^{\beta_0}$  in Figure 3.4 are qualitatively similar to those found in [49], where an analogous computation was performed on the original volume-preserving standard map with  $\mu_2 = \ell_2$ . In the current setting, these curves are additionally optimised to take into account the nonuniform  $\mu_2$  and non-volume preserving nature of  $T$ .

### 3.6 Conclusions to the chapter

The dynamic isoperimetric theory initiated in [49] was concerned with identifying subsets of  $\mathbb{R}^r$  with persistently least boundary size to volume ratio under general nonlinear volume-preserving dynamics. The motivation for this theory was that the boundaries of such sets have optimality properties desired in Lagrangian coherent structures. In the present work we have extended the constructions and theoretical results of [49] to weighted, non-flat Riemannian manifolds and to possibly non-volume preserving dynamics. This entailed developing a nontrivial generalisation of the dynamic isoperimetric problem to weighted manifolds and allowing for non-volume preserving dynamics. We proved a new dynamic version of the classical Federer-Fleming theorem in this setting, which very tightly links the (geometric) dynamic isoperimetric problem with a (functional) minimisation of a new dynamic Sobolev constant.

We then constructed a weighted dynamic Laplacian, and showed that under a natural Neumann-type boundary condition, the spectrum of this weighted dynamic Laplacian can be completely characterised using variational principles tied to the finite-time dynamics of  $T$  and the geometry of the manifold. We additionally proved that a dynamic Cheeger inequality holds on weighted Riemannian manifolds, extending a result from [49] for flat, unweighted manifolds, and volume-preserving dynamics. We demonstrated numerically that the eigenfunctions of the weighted dynamic Laplacian are able to identify sets with small boundaries that remain small when transformed by general dynamics. Such persistently minimal surfaces are excellent candidates for LCSs as diffusion across their short boundaries is minimised over a finite-time duration.

Finally, we further developed the connection between two very different methods for detecting transport barriers in dynamical systems, namely the relationship between finite-time coherent sets and LCSs as defined using isoperimetric notions. The connection between these sets, explored in [49] in the flat manifold, volume-preserving dynamics setting, is that in the limit of small diffusion, regions in phase space that minimally mix (a purely probabilistic notion) are linked with sets that have persistently least boundary size (a purely geometric notion). We further enhanced this link by extending and strengthening the result of [49] to the more general setting of weighted, curved Riemannian manifolds with possibly non-volume preserving dynamics.



## Chapter 4

# A dynamic manifold learning method for approximating Lagrangian coherent structures.

Standard Laplacian-based manifold learning method [10, 11] is designed for approximating structures of unweighted manifolds, from uniformly distributed point-cloud data (see Theorem 5.2 in [12] or Theorem 2 in [28]). Dynamic extensions of the standard Laplacian-based manifold learning method were formulated in [9, 40, 60], which have enabled the analysis of trajectory datasets arising from dynamical systems. However, the dynamic manifold learning methods in [9, 40, 60] are designed for approximating dynamical features of evolving unweighted manifolds, from trajectories that are formed by volume-preserving dynamics with uniform initial distribution. Recently, improved manifold learning methods have emerged to study unweighted manifolds from non-uniformly distributed point-cloud data. In [77], a local normalisation was applied to the relative distance between the data points of the input sample, before proceeding with the standard Laplacian-based manifold learning method [10, 11]. In [75], a robust neighbourhood selection process was used to account for the nonuniform distribution of the input sample. The dynamic manifold learning methods in [40, 60, 9] can readily incorporate the improvements of [75] or [77], thus extending their application to trajectory data with nonuniform initial distribution. However, a dynamic manifold learning method is currently missing for approximating dynamic structures of weighted manifolds, from trajectories formed by non-volume-preserving transformations.

In this chapter, we first develop an improved Laplacian-based manifold learning method for *weighted* Riemannian manifolds, which is designed to be robust to the probability distribution of the input point-cloud data. In particular, let  $(M, m, \mu_r)$

be a weighted Riemannian manifold as in Section 3.1, and let  $S^k$  be a random sample drawn from  $(M, m, \mu_r)$  according to some possibly nonuniform probability distribution in i.i.d fashion. In Section 4.1, we form a weighted graph  $G(S^k, W^{\mu, \epsilon})$  from  $S^k$  with edge weight set  $W^{\mu, \epsilon}$ , and prove that there exists a constant  $C_\rho$  and a sequence of scalars  $\{\epsilon_k\}_{k \geq 1}$  such that for every  $x \in M$  and  $f \in C^3(M, \mathbb{R})$

$$\lim_{k \rightarrow \infty} \frac{1}{\epsilon_k C_\rho} L^{\mu, k, \epsilon_k} f(x) = \Delta_\mu f(x), \quad (4.1)$$

where  $L^{\mu, k, \epsilon}$  is the continuous extension of the graph Laplacian on  $G(S^k, W^{\mu, \epsilon})$  (see (4.15)), and  $\Delta_\mu$  is the weighted Laplacian on  $(M, m, \mu_r)$  as in (3.27). The novelty of our approach for manifold learning is the introduction of a set of scaling factors for the edge weight set  $W^{\mu, \epsilon}$ , which makes  $G(S^k, W^{\mu, \epsilon})$  independent of the probability distribution of the sample  $S^k$  for sufficiently large  $k$ . More importantly, the weighted geometry of  $(M, m, \mu_r)$  is incorporated into the graph  $G(S^k, W^{\mu, \epsilon})$ . As an improvement over the robust manifold learning methods in [75] and [77], our manifold learning method is designed for approximating features of *weighted* Riemannian manifolds.

Next, we extend the above improved Laplacian-based manifold learning method to a dynamic manifold learning method, designed for approximating solutions of the *dynamic* isoperimetric problem on weighted Riemannian manifolds from sparse trajectory data. Let  $(N, n, \nu_r)$  be another weighted Riemannian manifold, where  $N$  is the image of  $M$  under a general transformation  $T : M \rightarrow N$  and  $\nu_r = \mu_r \circ T^{-1}$ . The initial point-cloud data  $S^k \in M$  is mapped to  $\hat{S}^k \in N$  under  $T$ , and we form another weighted graph  $G(\hat{S}^k, W^{\nu, \epsilon})$  from  $\hat{S}^k$  with edge weight set  $W^{\nu, \epsilon}$ . The dynamic generalisation of the operator  $L^{\mu, k, \epsilon}$  is given by

$$L^{dyn, k, \epsilon} := \frac{1}{2} (L^{\mu, k, \epsilon} + L^{\nu, k, \epsilon}),$$

where  $L^{\nu, k, \epsilon}$  is the continuous extension of the graph Laplacian on  $G(\hat{S}^k, W^{\nu, \epsilon})$ . The dynamic version of (4.1) is

$$\lim_{k \rightarrow \infty} \frac{1}{\epsilon_k C_\rho} L^{dyn, k, \epsilon_k} f(x) = \Delta^{dyn} f(x), \quad (4.2)$$

where  $\Delta^{dyn}$  is the dynamic Laplacian as in Definition 3.3.1. For situations where the trajectory dataset arises from a system with multiple time-step dynamics, the suitable version of (4.2) is given by (4.38).

In addition to the statement of the convergence result (4.1), in Section 4.1 we outline the numerical algorithm associated with our robust Laplacian-based manifold learning method, and conduct numerical experiments to test this algorithm

on random samples with very different probability distributions, as well as approximating the weighted geometry of domains equipped with general measures. Section 4.2 contains statement of the convergence result (4.2), an outline of the numerical algorithm for our robust dynamic manifold learning method, and numerical experiments for testing our robust dynamic manifold learning method on both artificially generated and real-world trajectory data. The proofs of this chapter are deferred to Appendix C.

## 4.1 A robust Laplacian-based manifold learning method for weighted Riemannian manifolds

Let  $M$  be a compact, connected  $r$ -dimensional Riemannian manifold, which is embedded in a possibly higher dimensional Euclidean space  $\mathbb{R}^d$ ;  $d \geq r$ . On  $M$  we have the absolutely continuous measure  $\mu_r$  and the metric tensor  $m$ , forming the weighted Riemannian manifold  $(M, m, \mu_r)$  as in Section 3.2. On  $\mathbb{R}^d$  we have the  $d$ -dimensional Euclidean norm  $\|\cdot\|_{\mathbb{R}^d}$ . As in Section 3.1, we assume that the density  $h_\mu$  of the measure  $\mu_r$  is bounded above and uniformly away from zero.

For  $\rho > 0$ , define  $q_\rho : \mathbb{R}^+ \cup \{0\} \rightarrow \mathbb{R}^+ \cup \{0\}$  by

$$q_\rho(x) := c_\rho \exp(-x^2) \mathcal{I}_{x \leq \rho}, \quad (4.3)$$

where  $\mathcal{I} : \mathbb{R}^+ \cup \{0\} \rightarrow \{0, 1\}$  is the characteristic function, and the constant  $c_\rho$  is chosen so that  $\int q_\rho(x) dx = 1$ . For  $\epsilon > 0$ , we construct a family of kernels  $Q_{\rho, \epsilon} : M \times M \rightarrow \mathbb{R}^+ \cup \{0\}$  by

$$Q_{\rho, \epsilon}(x, y) := \epsilon^{-r/2} q_\rho \left( \frac{\|\Phi(x) - \Phi(y)\|_{\mathbb{R}^d}}{\sqrt{\epsilon}} \right), \quad (4.4)$$

where  $\Phi : M \rightarrow \mathbb{R}^d$  is an isometric embedding. The kernels  $Q_{\rho, \epsilon}$  are symmetric for all  $\epsilon > 0$ , and almost-stochastic in the sense that  $\lim_{\epsilon \rightarrow 0} Q_{\rho, \epsilon}$  is stochastic [9]; that is, for all  $x \in M$ ,  $\lim_{\epsilon \rightarrow 0} \int_M Q_{\rho, \epsilon}(x, y) \omega_m^r(y) = 1$ , where  $\omega_m^r$  is the volume form on  $M$  as in Section 3.1. The kernel  $Q_{\rho, \epsilon}$  is a modified version of the kernel used in [10, 11] for standard Laplacian-based manifold learning (here in (4.3), we have adapted the normalisation factor  $c_\rho$  and the cutoff  $\mathcal{I}_{x \leq \rho}$  from [9]).

*Remark 4.1.1.* Alternatively, if one has access to the metric tensor  $m$  of the unknown manifold  $M$ , then the expression  $\|\Phi(x) - \Phi(y)\|_{\mathbb{R}^d}$  on the RHS of (4.4) can be replaced by  $m(x, y)$ ; see [28].

### 4.1.1 Graph Laplacian for scalar weighted graphs

Let  $S^k = \{x_i\}_{i=1}^k$  be a random sample drawn from  $M$  according to a  $C^5$  probability density  $p$  in i.i.d fashion. The regularity condition for  $p$  is to ensure that  $p$  and the second order partial derivatives of  $p$  are well approximated by point-cloud data  $S^k$ ; see (4.7) and (C.8). In practice, any real probability distribution will be sufficiently smooth. However, extremely small sample size or poor sampling strategy will lead to reductions in the accuracy our manifold learning methods. In order to formulate a manifold learning method that is robust to the probability distribution of  $S^k$ , we first discuss how the kernel  $Q_{\rho,\epsilon}$  provides an approximation to  $p$  (In practice  $p$  may be unknown). For  $\epsilon > 0$ , define

$$p^{k,\epsilon}(x) := \frac{1}{k} \sum_{i=1}^k Q_{\rho,\epsilon}(x, x_i). \quad (4.5)$$

Since  $S^k$  is an i.i.d random sample drawn from  $M$  according to the probability density  $p$ , one has by an application of the law of large numbers [73]

$$\lim_{k \rightarrow \infty} p^{k,\epsilon}(x) = \lim_{k \rightarrow \infty} \frac{1}{k} \sum_{i=1}^k Q_{\rho,\epsilon}(x, x_i) = \int_M Q_{\rho,\epsilon}(x, y) p(y) \omega_m^r(y), \quad (4.6)$$

for all  $x \in M$ , where  $\omega_m^r$  is the volume form on  $M$  as in (B.10). Moreover, due to the assumption  $p \in C^3(M, \mathbb{R})$  and the properties of the kernel  $Q_{\rho,\epsilon}$  (namely almost stochastic, rotational invariance and bounded support), by a straightforward modification of Lemma B.7.3 one has

$$\int_M Q_{\rho,\epsilon}(x, y) p(y) \omega_m^r(y) = p(x) + \frac{\epsilon a_\rho}{2} \Delta_m p(x) + R(x, \epsilon^{3/2}), \quad R(x, \epsilon^{3/2}) \in \mathcal{O}(\epsilon^{3/2}), \quad (4.7)$$

where  $a_\rho > 0$  depends on the second moment of  $q_\rho$ , and  $\mathcal{O}(\epsilon^{3/2})$  denotes the class of polynomials  $a_1 \epsilon^{3/2} + a_2 \epsilon^{4/2} + a_3 \epsilon^{5/2} + \dots$ . Therefore, if we set  $\epsilon := \epsilon_k$  as a sequence of scalars, such that  $\lim_{k \rightarrow \infty} \epsilon_k = 0$ , then by (4.7) the expression  $p^{k,\epsilon_k}$  converges pointwise to  $p$  as  $k \rightarrow \infty$ .

We now introduce the key modification we have made to the standard Laplacian-based manifold learning method [10, 11]: we set

$$w_{ij}^{\mu,\epsilon} := \begin{cases} \frac{Q_{\rho,\epsilon}(x_i, x_j)}{p^{k,\epsilon}(x_j)} \sqrt{\frac{h_\mu(x_j)}{h_\mu(x_i)}} & \text{if } i \neq j \\ 0 & \text{if } i = j \end{cases}, \quad (4.8)$$

for each  $1 \leq i, j \leq k$ . As in Section 1.3, we construct a weighted graph from  $S^k$  by setting  $w_{ij}^{\mu,\epsilon}$  as the edge weight between the vertex pair  $x_i$  and  $x_j$ . We denote the weighted graph formed above by  $G(S^k, W^{\mu,\epsilon})$ , where  $W^{\mu,\epsilon} = \{w_{ij}^{\mu,\epsilon}\}_{i,j=1}^k$ . The

purpose of the term  $\sqrt{h_\mu(x_j)/h_\mu(x_i)}$  in (4.8) is to incorporate the density  $h_\mu$  into the edge weight  $w_{ij}^{\mu,\epsilon}$  of the graph  $G(S^k, W^{\mu,\epsilon})$ . For technical reasons of establishing Theorem 4.1.4 below, we have scaled  $Q_{\rho,\epsilon}$  by  $\sqrt{h_\mu(x_j)/h_\mu(x_i)}$  instead of just  $h_\mu(x_j)$  (This technical step serves the same purpose as the normalisation (1.11) in standard Laplacian-based manifold learning; see Theorem 5.1 and Theorem 5.2 in [12]).

Denote by  $\mathbf{L}^{\mu,\epsilon}$  the graph Laplacian for  $G(S^k, W^{\mu,\epsilon})$  as in 1.3.3; i.e.  $\mathbf{L}^{\mu,\epsilon}$  has entries

$$\mathbf{L}_{ij}^{\mu,\epsilon} := \begin{cases} w_{ij}^{\mu,\epsilon} & \text{if } i \neq j \\ -\sum_{l=1}^k w_{il}^{\mu,\epsilon} & \text{if } i = j \end{cases}, \quad (4.9)$$

for  $1 \leq i, j \leq k$ . We note that  $w_{ij}^{\mu,\epsilon}$  is not symmetric, hence  $\mathbf{L}^{\mu,\epsilon}$  is not self-adjoint with respect to the standard inner product on  $\mathbb{R}^k$ . However, we can define a *weighted inner product*  $\langle \cdot, \cdot \rangle_\mu : \mathbb{R}^k \times \mathbb{R}^k \rightarrow \mathbb{R}$  by

$$\langle \mathbf{g}, \mathbf{f} \rangle_\mu := \sum_{i=1}^k g_i f_i \frac{h_\mu(x_i)}{p^{k,\epsilon}(x_i)}, \quad (4.10)$$

with respect to which  $\mathbf{L}^{\mu,\epsilon}$  is self-adjoint; see Corollary 4.1.3 below. We define the *weighted eigenproblem* of  $\mathbf{L}^{\mu,\epsilon}$  with respect to the weighted inner product (4.10) by

$$\langle \mathbf{f}, \mathbf{L}^{\mu,\epsilon} \mathbf{f} \rangle_\mu = \lambda \langle \mathbf{f}, \mathbf{f} \rangle_\mu. \quad (4.11)$$

Let  $\mathbf{B}^{\mu,\epsilon}$  be the  $k \times k$  matrix with entries

$$\mathbf{B}_{ij}^{\mu,\epsilon} = \frac{h_\mu(x_i)}{p^{k,\epsilon}(x_i)} \mathbf{L}_{ij}^{\mu,\epsilon}. \quad (4.12)$$

Solution pairs  $(\lambda, \mathbf{f})$  to the standard matrix eigenproblem  $\mathbf{B}^{\mu,\epsilon} \mathbf{f} = \lambda \mathbf{f}$  are also solutions to (4.11), since

$$\begin{aligned} \langle \mathbf{f}, \mathbf{L}^{\mu,\epsilon} \mathbf{f} \rangle_\mu &= \sum_{i=1}^k f_i \cdot \frac{h_\mu(x_i)}{p^{k,\epsilon}(x_i)} \cdot (\mathbf{L}^{\mu,\epsilon} \mathbf{f})_i \quad \text{by (4.14)} \\ &= \sum_{i=1}^k f_i \cdot \frac{h_\mu(x_i)}{p^{k,\epsilon}(x_i)} \cdot \sum_{j=1}^k \mathbf{L}_{ij}^{\mu,\epsilon} f_j \\ &= \sum_{i=1}^k f_i \cdot \sum_{j=1}^k \mathbf{B}_{ij}^{\mu,\epsilon} f_j \quad \text{by (4.12)} \\ &= \langle \mathbf{f}, \mathbf{B}^{\mu,\epsilon} \mathbf{f} \rangle \\ &= \lambda \langle \mathbf{f}, \mathbf{f} \rangle_\mu. \end{aligned} \quad (4.13)$$

The matrix  $\mathbf{B}^{\mu,\epsilon}$  has the following spectral properties:

**Theorem 4.1.2.** *Let  $G(S^k, W^{\mu,\epsilon})$  be a weighted graph, where  $W^{\mu,\epsilon} = \{w_{ij}^{\mu,\epsilon}\}_{i,j=1}^k$  with  $w_{ij}^{\mu,\epsilon}$  as in (4.8). Define  $\mathbf{B}^{\mu,\epsilon}$  as in (4.12). One has*

1. The matrix  $\mathbf{B}^{\mu,\epsilon}$  is self-adjoint with respect to the standard inner product on  $\mathbb{R}^k$ .
2. The eigenvalues  $0 = \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_k$  of  $\mathbf{B}^{\mu,\epsilon}$  are nonpositive and real.
3. If  $G(S^k, W^{\mu,\epsilon})$  is connected, then  $\lambda_1$  is unit multiplicity and the eigenvector corresponding to  $\lambda_1$  is  $\mathbf{1}$ .

*Proof.* See Appendix C. □

**Corollary 4.1.3.** Let  $\mathbf{L}^{\mu,\epsilon}$  be as in Theorem 4.1.2, and let  $\langle \cdot, \cdot \rangle_\mu$  be the weighted inner product as in (4.11). The matrix  $\mathbf{L}^{\mu,\epsilon}$  is self-adjoint with respect to  $\langle \cdot, \cdot \rangle_\mu$ ; that is,  $\langle \mathbf{g}, \mathbf{L}^{\mu,\epsilon} \mathbf{f} \rangle_\mu = \langle \mathbf{L}^{\mu,\epsilon} \mathbf{g}, \mathbf{f} \rangle_\mu$  for all  $\mathbf{g}, \mathbf{f} \in \mathbb{R}^k$ .

*Proof.* This follows immediately by part 1. in Theorem 4.1.2 and a straightforward modification of (4.13). □

## 4.1.2 From graph Laplacian to Laplace-Beltrami operator

Denote by  $\Delta_\mu$  the weighted Laplacian on  $(M, m, \mu_r)$  as in (3.27). The connection between  $\mathbf{L}^{\mu,\epsilon}$  and  $\Delta_\mu$  is as follows: Let  $f \in C^\infty(M, \mathbb{R})$ . Denote  $f_i = f(x_i)$  for each  $1 \leq i \leq k$ , and set  $\mathbf{f} = \{f_1, f_2, \dots, f_k\}$ . Since

$$(\mathbf{L}^{\mu,\epsilon} \mathbf{f})_i = \sum_{j=1}^k \mathbf{L}_{ij}^{\mu,\epsilon} f_j = \sum_{\substack{j=1 \\ j \neq i}}^k w_{ij}^{\mu,\epsilon} f_j - \sum_{l=1}^k w_{il}^{\mu,\epsilon} f_l = \sum_{j=1}^k w_{ij}^{\mu,\epsilon} (f_j - f_i), \quad (4.14)$$

for each  $1 \leq i \leq k$ , one extends the linear action  $\mathbf{L}^{\mu,\epsilon} : \mathbb{R}^k \rightarrow \mathbb{R}^k$  to  $L^{\mu,k,\epsilon} : C^\infty(M, \mathbb{R}) \rightarrow C^\infty(M, \mathbb{R})$  by defining

$$L^{\mu,k,\epsilon} f(x) := \frac{1}{k} \sum_{j=1}^k \frac{Q_{\rho,\epsilon}(x, x_j)}{p^{k,\epsilon}(x_j)} \sqrt{\frac{h_\mu(x_j)}{h_\mu(x)}} (f(x_j) - f(x)), \quad (4.15)$$

for all  $x \in M$ . We have

**Theorem 4.1.4.** Let  $S^k$  be a random sample drawn from the weighted Riemannian manifold  $(M, m, \mu_r)$  according to some  $C^5$  probability density. Assume the density  $h_\mu$  of  $\mu_r$  is in  $C^3$ , bounded above and uniformly away from zero. Define  $L^{\mu,k,\epsilon}$  and  $\Delta_\mu$  as in (4.15) and (3.27), respectively. If  $0 < \rho \leq \mathcal{S}_M$ , where  $\mathcal{S}_M$  is the scalar curvature<sup>1</sup> of  $M$ , then there exists a constant  $C_\rho > 0$  and a sequence of scalars  $\{\epsilon_k\}_{k \geq 1}$  with  $\lim_{k \rightarrow \infty} \epsilon_k = 0$  such that

$$\lim_{k \rightarrow \infty} \left( \sup_{\|f\|_{C^3(M, \mathbb{R})} \leq 1} \left\| \frac{1}{\epsilon_k C_\rho} L^{\mu,k,\epsilon_k} f(x) - \Delta_\mu f(x) \right\|_{C^0(M, \mathbb{R})} \right) = 0, \quad (4.16)$$

<sup>1</sup>The scalar curvature is defined as the trace of the Ricci curvature tensor; for more details see e.g. p.117 in [7].

for all  $x \in M$  and  $f \in C^\infty(M, \mathbb{R})$ .

*Proof.* See Appendix C. □

The spectral properties of  $\Delta_\mu$  were studied in Chapter 3: By setting the dynamics acting on  $M$  as the identity map in Theorem 3.3.4, one has

$$\inf_{\beta \in (-\infty, \infty)} H_M(\{\phi_2 = \beta\}) \leq 2\sqrt{\lambda_2}, \quad (4.17)$$

where  $H_M$  is the weighted version of the classical Cheeger ratio (see (1.8)), and  $\lambda_2$  is the first nontrivial eigenvalue of  $\Delta_\mu$  with corresponding eigenvector  $\phi_2$ . Hence, among the level surfaces of  $\phi_2$ , there exists a level surface  $\Gamma_0$  in  $M$  that yields a small Cheeger ratio  $H_M(\Gamma_0)$ . Let  $\{M_1, M_2\}$  be full dimensional, connected components of  $M$ , which are separated by the interface  $\Gamma_0$ . Using Theorem 4.1.4 and (4.17), we can attempt to approximate  $\{M_1, M_2\}$  using the first nontrivial eigenvector  $\mathbf{f}_2$  of  $\mathbf{L}^{\mu, \epsilon}$  with respect to (4.11); that is, we use the components  $f_{2,i}$  of  $\mathbf{f}_2$ , to search for  $\beta \in \mathbb{R}$  such that sets  $S_1^k = \{x_i \in S^k : f_{2,i} < \beta\}$  and  $S_2^k = \{x_i \in S^k : f_{2,i} > \beta\}$  are good representation of  $\{M_1, M_2\}$ . Our Laplacian-based manifold learning method is a heuristic that builds on this idea (see also [10] for the classical treatment), which we now discuss.

### 4.1.3 Algorithm for weighted Laplacian eigenmap

Let  $S^k$  be an i.i.d random sample drawn from  $M \subseteq \mathbb{R}^d$  according to a  $C^5$  probability density  $p$ , and recall the induced co-dimension 1 measure  $\mu_{r-1}$  on  $(M, m, \mu_r)$  given by (3.8). Let  $\Gamma$  be a piecewise continuous hypersurface in  $M$ , that partitions  $M$  into full dimensional submanifolds  $M_1, M_2, \dots, M_K$ . Suppose we are given the following information: (1) The  $\mathbb{R}^d$ -coordinates of each  $x_i \in S^k$ , and (2) the value of the density  $h_\mu(x_i)$  for each  $1 \leq i \leq k$ . The aim of our weighted Laplacian-based manifold learning method, is to find a coordinate map  $\varphi_\mu : S^k \rightarrow \mathbb{R}^s$  independent of  $p$  with  $s \ll k$ , such that certain low-dimensional structures of  $(M, m, \mu_r)$  are encapsulated by the set  $\varphi_\mu(S^k) \in \mathbb{R}^s$ . In particular, geometrically close clusters  $\{S_1^k, S_2^k, \dots, S_K^k\}$  of  $S^k$  with respect to  $\mathbb{R}^s$ -distances of  $\varphi_\mu(S^k)$ , represent full dimensional submanifolds  $\{M_1, M_2, \dots, M_K\}$ , which partitions  $M$  so that the ratio

$$H_{M,K}(\Gamma) := \frac{\sum_{i=1}^{K-1} \mu_{r-1}(\Gamma)}{\min\{\mu_r(M_1), \mu_r(M_2), \dots, \mu_r(M_K)\}}, \quad (4.18)$$

is small; we call the coordinate map  $\varphi_\mu$  the *weighted Laplacian eigenmap*.

We now present the algorithmic procedure for forming the weighted Laplacian eigenmap  $\varphi_\mu$ . This algorithm is an extension of the standard Laplacian eigenmaps

algorithm in [20], and we have incorporated the computational processes for setting  $\epsilon$  and  $\rho$  in [9] for the kernel  $Q_{\rho,\epsilon}$ .

**Algorithm 4.1:** Weighted Laplacian eigenmap

- 1 Let  $S^k = \{x_i\}_{i=1}^k$  and  $\{h_\mu(x_i)\}_{i=1}^k$  be input data, and  $w_{ij}^{\mu,\epsilon}$  be as in (4.8). Define  $\epsilon_{\min}$  to be the smallest real number such that the graph  $G(S^k, W^{\mu,\epsilon})$  is connected, and define  $\epsilon_{\max}$  to be the smallest real number such that  $w_{ij}^{\mu,\epsilon_{\max}} \neq 0$  for each  $1 \leq i, j \leq k$ .
- 2 Fix  $0 < \rho < \mathcal{S}_M$ , where  $\mathcal{S}_M$  is the scalar curvature of  $M$ . For a sequence of  $\epsilon_l \in [\epsilon_{\min}, \epsilon_{\max}]$ , construct the  $k \times k$  matrix  $\mathbf{B}^{\mu,\epsilon_l}$  according to (4.12).
- 3 Solve the eigenproblem  $\frac{1}{\epsilon} \mathbf{B}^{\mu,\epsilon_l} \mathbf{f} = \lambda^{\epsilon_l} \mathbf{f}$  for each  $\epsilon_l \in [\epsilon_{\min}, \epsilon_{\max}]$ , and find  $\epsilon_L$  amongst the sequence of  $\epsilon_l$  such that  $|\lambda_2^{\epsilon_l} - \lambda_2^{\epsilon_{l+1}}|$  is minimal, where  $\lambda_2^{\epsilon_l}$  denotes the first nontrivial eigenvalue of  $\frac{1}{\epsilon} \mathbf{B}^{\mu,\epsilon_l}$ .
- 4 Denote the eigenvalues of  $\frac{1}{\epsilon} \mathbf{B}^{\mu,\epsilon_L}$  by  $\lambda_1 \leq \lambda_2 \leq \dots$ . Find the smallest integer  $1 \leq s \ll k$ , such that the magnitude of  $|\lambda_{s+2} - \lambda_{s+1}|$  is comparable to  $2|\lambda_2 - \lambda_1|$ . If no such  $s$  exists, return  $s = 1$ .
- 5 Let  $\mathbf{f}_2, \mathbf{f}_3, \dots$  be the eigenvectors corresponding to  $\lambda_2, \lambda_3, \dots$ . Define  $\varphi_\mu(x_i) := \{\mathbf{f}_2(x_i), \mathbf{f}_3(x_i), \dots, \mathbf{f}_{s+1}(x_i)\}$  for each  $1 \leq i \leq k$ .
- 6 Return  $\varphi_\mu$ .

*Remark 4.1.5.* In step 2 of Algorithm 4.1, the parameter  $\rho$  was bounded above by the scalar curvature  $\mathcal{S}_M$  of  $M$ . In situations where  $\mathcal{S}_M$  is unknown (a common occurrence in many situations), we set  $\rho = \sqrt{2}$  to ensure that the point of inflection of the function  $q_\rho$  (4.4) is at the midpoint of  $[0, \rho]$ , so that the corresponding graph  $G(S^k, W^{\mu,\epsilon})$  has edge weights that are most capable at discriminating distances between any pairs of data points  $x_i \neq x_j$  that are in the support of  $Q_{\rho,\epsilon}$  (see also [9]).

*Remark 4.1.6.* In practice, one can save computational resources by reducing the range of  $\epsilon_l \in [\epsilon_{\min}, \epsilon_{\max}]$ . We suggest setting  $\epsilon_{\max}$  at 30% sparsity<sup>2</sup> of  $\mathbf{B}^{\mu,\epsilon_{\max}}$ , which should adequately capture the pointwise geometric structures of  $S^k$ .

#### 4.1.4 Numerical experiments for weighted Laplacian eigenmap

We detail numerical examples of Algorithm 4.1 for weighted Riemannian manifold learning. To highlight the advantages of our weighted Laplacian eigenmap  $\varphi_\mu$  formed by Algorithm 4.1 compared to the standard Laplacian eigenmap  $\varphi$  described in Section 1.3.3, we apply both manifold learning methods to random samples that are

<sup>2</sup>The sparsity of a matrix is number of nonzero entries divided by the total number of entries.

distributed very differently. We then attempt to capture the geometry of weighted Riemannian manifolds from random samples using  $\varphi_\mu$ .

Let  $\mathbb{S}^2 := [0, 4) \times [0, 1) \setminus \sim$  be a two dimensional torus, where  $\sim$  denotes identification at interval ends. The manifold  $\mathbb{S}^2$  is embedded in  $\mathbb{R}^3$ , and the metric tensor  $e$  on  $\mathbb{S}^2$  is given by the Kronecker delta; that is, the components of  $e$  satisfies  $e_{ij} = 1$  for  $i = j$  and  $e_{ij} = 0$  for  $i \neq j$ . If a random sample  $S^k = \{x_i\}_{i=1}^k$  is drawn from  $\mathbb{S}^2$  according to the  $C^5$  probability density  $p : \mathbb{S}^2 \rightarrow \mathbb{R}^+$ , then for all measurable  $U \subseteq \mathbb{S}^2$ , the probability that  $x_i \in U$  is  $\int_U p \, dx dy$ . We form a weighted 2-dimensional torus  $(\mathbb{S}^2, e, \mu_2)$ , where  $\mu_2$  is an absolutely continuous measure on  $\mathbb{S}^2$ .

### Randomness of input data

To demonstrate how the eigenmap  $\phi_\mu$  is robust to the randomness of the input data, we fix the density  $h_\mu$  of  $\mu_2$  to be 1 on all of  $\mathbb{S}^2$ , and artificially generate a few random samples of the form  $S^k = \{(x_i, y_i)\}_{i=1}^k$  from  $\mathbb{S}^2$ , with each sample having a different probability density. Although, one can calculate the scalar curvature of  $\mathbb{S}^2$  to bound the parameter  $\rho$  in step 2 of Algorithm 3.1, we simulate the situation where nothing is known about  $\mathbb{S}^2$ . Thus we set  $\rho = \sqrt{2}$  in  $Q_{\rho, \epsilon}$  in accordance to Remark 4.1.5.

To maintain some predictability of the randomness of our input data, we manipulate the probability density of each sample as follows. Partition  $\mathbb{S}^2$  into two half-tori  $\mathbb{S}_1^2 := [0, 2) \times [0, 1) \setminus \sim$  and  $\mathbb{S}_2^2 := [2, 4) \times [0, 1) \setminus \sim$ , then draw a fixed number of data points from each of  $\mathbb{S}_1^2$  and  $\mathbb{S}_2^2$  uniformly in i.i.d fashion. One has control over the ratio of the number of data points between the regions  $\mathbb{S}_1^2$  and  $\mathbb{S}_2^2$ . For our numerical experiments, we fixed  $k = 1000$  and set the ratios of points between  $\mathbb{S}_1^2$  and  $\mathbb{S}_2^2$  to be 1 : 1, 1 : 4 and 1 : 9; we denote these sets by  $S_{1:1}^k$ ,  $S_{1:4}^k$  and  $S_{1:9}^k$ , respectively. The distributions of  $S_{1:1}^k$ ,  $S_{1:4}^k$  and  $S_{1:9}^k$  embedded in  $\mathbb{R}^3$  are shown by the first column of Figure 4.1.

We start with the formulation of the unweighted Laplacian eigenmap  $\varphi$ , formed by applying Algorithm 4.1 with the following modifications:

1. In steps 2 - 4 of Algorithm 4.1, replace  $\mathbf{B}^{\mu, \epsilon}$  with  $\mathcal{L}^\epsilon$  (1.11).
2. Let  $\mathbf{D}^\epsilon$  be as Section 1.3.3. In steps 5 and 6 of Algorithm 4.1, replace  $\varphi_\mu$  with  $\varphi(x_i) = \{(\mathbf{D}^\epsilon)^{-\frac{1}{2}} \mathbf{f}_2(x_i), (\mathbf{D}^\epsilon)^{-\frac{1}{2}} \mathbf{f}_3(x_i), \dots, (\mathbf{D}^\epsilon)^{-\frac{1}{2}} \mathbf{f}_{s+1}(x_i)\}$ .

The key difference between the use of  $\mathcal{L}^\epsilon$  for standard Laplacian-based manifold learning [10, 11] and the use of  $\mathbf{B}^{\mu, \epsilon}$  in Algorithm 4.1, is the application of the scaling factors

$$\frac{1}{p^{k, \epsilon}(x_j)} \sqrt{\frac{h_\mu(x_i)}{h_\mu(x_j)}} \quad (4.19)$$

in (4.8) to the diffusion kernel  $Q_{\rho,\epsilon}$ ; the purpose of this scaling factor (4.19) was discussed in Section 4.1.1. To form the standard Laplacian-eigenmap  $\varphi$ , the reason for the left multiplication of  $\mathbf{f}_2, \mathbf{f}_3, \dots, \mathbf{f}_{s+1}$  by  $(\mathbf{D}^\epsilon)^{-\frac{1}{2}}$ , is that  $\mathcal{L}^\epsilon$  is the normalised graph Laplacian (see step 3 of Algorithm 2.1 where this multiplication was applied for spectral graph partition).

Denote by  $\mathcal{L}_{1:1}^\epsilon, \mathcal{L}_{1:4}^\epsilon$  and  $\mathcal{L}_{1:9}^\epsilon$  the normalised graph Laplacians formed from  $S_{1:1}^k, S_{1:4}^k$  and  $S_{1:9}^k$ , respectively. First we choose the parameter  $\epsilon$  in  $\mathcal{L}_{1:1}^\epsilon, \mathcal{L}_{1:4}^\epsilon$  and  $\mathcal{L}_{1:9}^\epsilon$  using steps 1 - 3 of Algorithm 4.1. For the set  $S_{1:1}^k$  we have  $\epsilon = 0.055$  with the first 7 eigenvalues of  $\frac{1}{\epsilon}\mathcal{L}_{1:1}^\epsilon$  given by 0,  $-0.1994$ ,  $-0.2371$ ,  $-0.7754$ ,  $-0.9354$ ,  $-1.7633$  and  $-1.8310$ . For the set  $S_{1:4}^k$ , we have  $\epsilon = 0.065$  with the first 7 eigenvalues of  $\frac{1}{\epsilon}\mathcal{L}_{1:4}^\epsilon$  given by 0,  $-0.2139$ ,  $-0.2281$ ,  $-0.9267$ ,  $-1.0257$ ,  $-1.5606$  and  $-2.0103$ . Finally, for the set  $S_{1:9}^k$ , we have  $\epsilon = 0.065$  with the first 7 eigenvalues of  $\frac{1}{\epsilon}\mathcal{L}_{1:9}^\epsilon$  given by 0,  $-0.2431$ ,  $-0.2704$ ,  $-0.6847$ ,  $-0.9671$ ,  $-1.9913$  and  $-2.2876$ . For each operator  $\mathcal{L}_{1:1}^\epsilon, \mathcal{L}_{1:4}^\epsilon$  and  $\mathcal{L}_{1:9}^\epsilon$ , there is a gap in the spectrum between the 3<sup>rd</sup> and 4<sup>th</sup> eigenvalues. Thus, by steps 4 and 5 of Algorithm 4.1, the eigenmap  $\varphi$  for each of the random sample  $S_{1:1}^k, S_{1:4}^k$  and  $S_{1:9}^k$  is the 2-dimensional embedding given by  $\varphi = \{\mathbf{g}_2, \mathbf{g}_3\} = \{(\mathbf{D}^\epsilon)^{-\frac{1}{2}}\mathbf{f}_2, (\mathbf{D}^\epsilon)^{-\frac{1}{2}}\mathbf{f}_3\}$ , where  $\mathbf{f}_2, \mathbf{f}_3$  are the first 2 nontrivial eigenvalues of the corresponding  $\mathcal{L}_{1:1}^\epsilon, \mathcal{L}_{1:4}^\epsilon$  or  $\mathcal{L}_{1:9}^\epsilon$ ; the images  $S_{1:1}^k, S_{1:4}^k$  and  $S_{1:9}^k$  under their eigenmap are shown in the second column of Figure 4.1.

The most notable feature of the 2-torus is its translational symmetries. Therefore, to assess the quality of the eigenmap  $\varphi$  for capturing the geometry of  $\mathbb{S}^2$ , we search for symmetries of the point-cloud data  $S_{1:1}^k, S_{1:4}^k$  and  $S_{1:9}^k$  embedded in  $\mathbb{R}^2$  via  $\varphi$ . For the set  $S_{1:1}^k$ , the eigenmap  $\varphi$  embeds  $S_{1:1}^k$  into a circular point-cloud loop in  $\mathbb{R}^2$ ; as shown in the first plot in the second column of Figure 4.1. We note that the symmetrical structures of the circular point-cloud loop  $\varphi(S_{1:1}^k)$  reflect the symmetries of  $\mathbb{S}^2$ . Therefore, we were able to use the standard Laplacian eigenmap [10, 11] to gain insight into the symmetrical structures of  $\mathbb{S}^2$  from a uniformly distributed point-cloud dataset. On the other hand, we note that the point-cloud geometries of  $\varphi(S_{1:4}^k)$  and  $\varphi(S_{1:9}^k)$  in  $\mathbb{R}^2$  do not reveal the symmetrical structures of  $\mathbb{S}^2$ , because the shapes of  $\varphi(S_{1:4}^k)$  and  $\varphi(S_{1:9}^k)$  can broadly be described as cones; as shown in the second and third plots in the second column of Figure 4.1. We speculate that these cones  $\varphi(S_{1:4}^k)$  and  $\varphi(S_{1:9}^k)$  are formed as elongated circles of the circular point-cloud loop  $\varphi(S_{1:1}^k)$ . The elongation of  $\varphi(S_{1:4}^k)$  and  $\varphi(S_{1:9}^k)$  can be explained by the point distributions of  $S_{1:4}^k$  and  $S_{1:9}^k$ , the unbalanced number of data points between the half toruses  $[0, 2) \times [0, 1) \setminus \sim$  and  $[2, 4) \times [0, 1) \setminus \sim$  of  $S_{1:4}^k$  and  $S_{1:9}^k$ , are treated as a geometrical features by the eigenmap  $\varphi$  (with the cone  $\varphi(S_{1:9}^k)$  being significantly more elongated due to the distribution of the set  $S_{1:9}^k$  being more unbalanced).

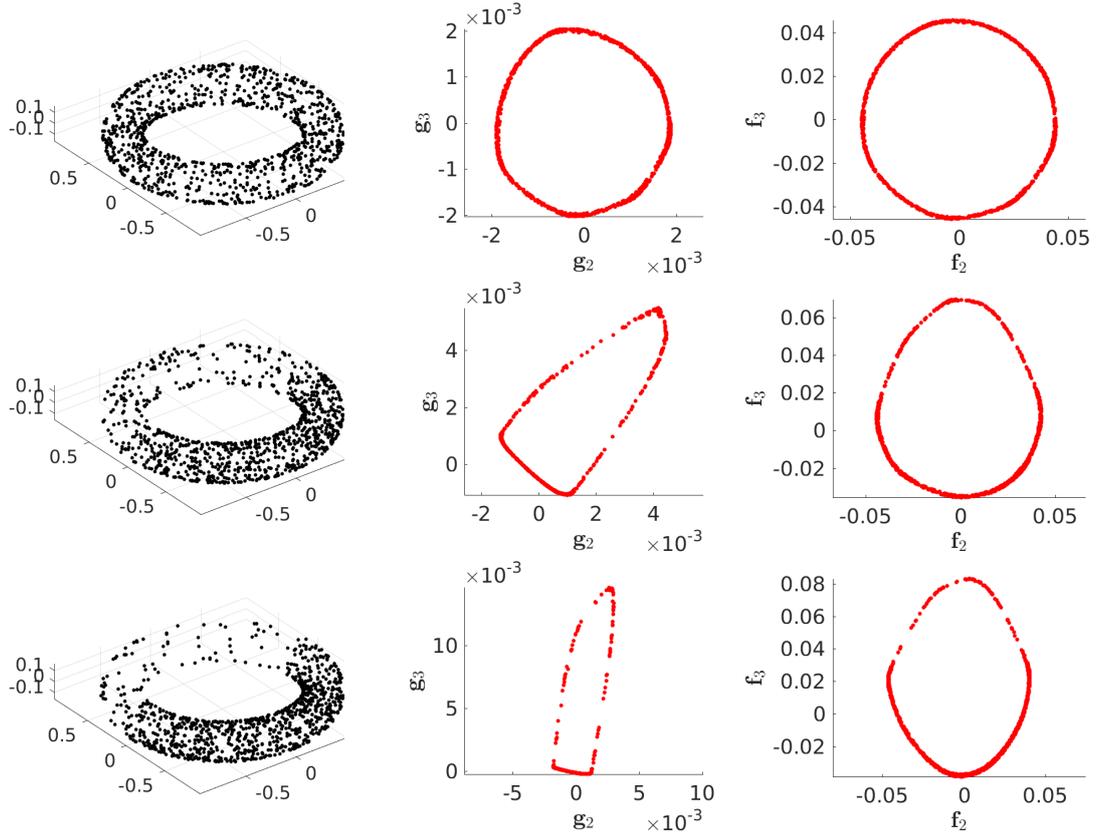


Figure 4.1: Comparison between standard and weighted Laplacian eigenmaps. Rows represent different input data, ordered from top to bottom  $S_{1:1}^k$ ,  $S_{1:4}^k$  and  $S_{1:9}^k$ . First column is the point-cloud data embedded in  $\mathbb{R}^3$ . Second and third columns are the eigenmaps  $\varphi$  and  $\varphi_\mu$ , respectively.

Next, we repeat the above numerical experiments conducted on the input data  $S_{1:1}^k$ ,  $S_{1:4}^k$  and  $S_{1:9}^k$ , but using Algorithm 4.1 without modifications. Let  $\mathbf{B}_{1:1}^{\mu,\epsilon}$ ,  $\mathbf{B}_{1:4}^{\mu,\epsilon}$  and  $\mathbf{B}_{1:9}^{\mu,\epsilon}$  denote the matrix formed from step 2 of Algorithm 4.1 from the data sets  $S_{1:1}^k, S_{1:4}^k$  and  $S_{1:9}^k$ , respectively. We set the parameter  $\epsilon$  in  $\mathbf{B}_{1:1}^{\mu,\epsilon}$ ,  $\mathbf{B}_{1:4}^{\mu,\epsilon}$  and  $\mathbf{B}_{1:9}^{\mu,\epsilon}$  using steps 1 - 3 of Algorithm 4.1. For the set  $S_{1:1}^k$  we have  $\epsilon = 0.085$  with the first 7 eigenvalues of  $\frac{1}{\epsilon}\mathbf{B}_{1:1}^{\mu,\epsilon}$  given by 0,  $-0.3262$ ,  $-0.3462$ ,  $-1.2654$ ,  $-1.3492$ ,  $2.7474$  and  $-2.8294$ . For the set  $S_{1:4}^k$ , we have  $\epsilon = 0.085$  with the first 7 eigenvalues of  $\frac{1}{\epsilon}\mathbf{B}_{1:4}^{\mu,\epsilon}$  given by 0,  $-0.2950$ ,  $-0.4286$ ,  $-1.2398$ ,  $-1.5413$ ,  $-2.8071$  and  $-3.0029$ . Finally for the set  $S_{1:9}^k$ , we have  $\epsilon = 0.085$  with the first 7 eigenvalues of  $\frac{1}{\epsilon}\mathbf{B}_{1:9}^{\mu,\epsilon}$  given by 0,  $-0.2646$ ,  $-0.5286$ ,  $-1.1575$ ,  $-1.7126$ ,  $-2.9238$  and  $-3.3729$ . Again, there is a gap in the spectrum between the 3<sup>rd</sup> and 4<sup>th</sup> eigenvalues of  $\mathbf{B}_{1:1}^{\mu,\epsilon}$ ,  $\mathbf{B}_{1:4}^{\mu,\epsilon}$  and  $\mathbf{B}_{1:9}^{\mu,\epsilon}$ . Thus, by steps 4 and 5 of Algorithm 4.1, the eigenmap for each  $S_{1:1}^k, S_{1:4}^k$  and  $S_{1:9}^k$  a 2-dimensional embedding given by  $\varphi_\mu = \{\mathbf{f}_2, \mathbf{f}_3\}$ , where  $\mathbf{f}_2, \mathbf{f}_3$  are the dominant 2 eigenvectors of the corresponding  $\mathbf{B}_{1:1}^{\mu,\epsilon}$ ,  $\mathbf{B}_{1:4}^{\mu,\epsilon}$  or  $\mathbf{B}_{1:9}^{\mu,\epsilon}$ ; the images of  $S_{1:1}^k$ ,  $S_{1:4}^k$  and  $S_{1:9}^k$  under their eigenmap are shown in the third column of Figure 4.1. Comparing the

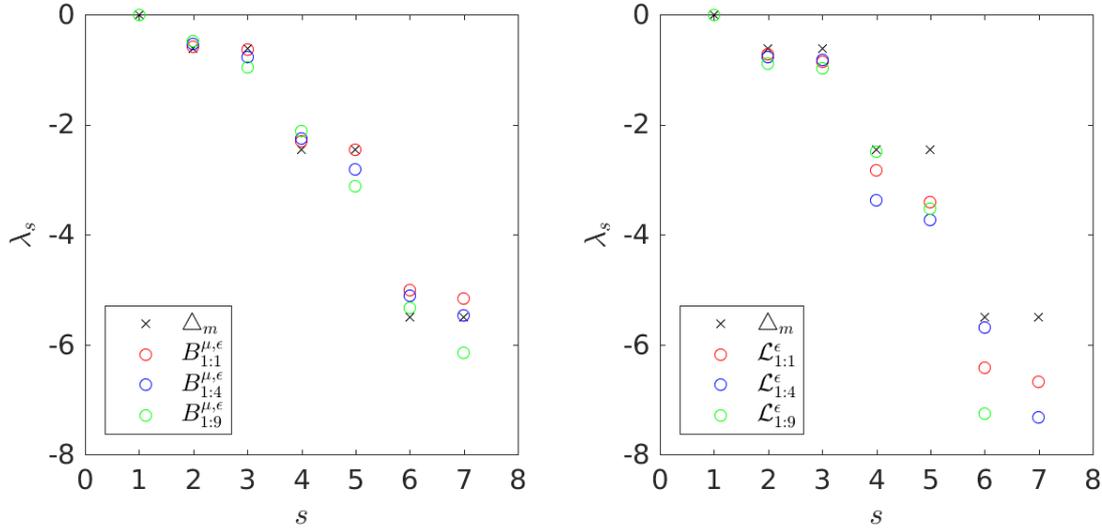


Figure 4.2: The spectra of the first 7 eigenvalues of  $\Delta_m$ ,  $\frac{1}{\epsilon C_\rho} \mathbf{B}_{1:j}^{\mu, \epsilon}$  and  $\frac{1}{2\epsilon C_\rho} \mathcal{L}_{1:j}^\epsilon$ , for  $j = 1, 4, 9$ . Left: colours represent eigenvalues for  $\frac{1}{\epsilon C_\rho} \mathbf{B}_{1:j}^{\mu, \epsilon}$  for  $j = 1, 4, 9$ , and crosses represents the eigenvalues of  $\Delta_m$ . Right: colours represent eigenvalues of  $\Delta_m$  and  $\frac{1}{2\epsilon C_\rho} \mathcal{L}_{1:j}^\epsilon$  for  $j = 1, 4, 9$ , and crosses represents the eigenvalues of  $\Delta_m$ .

eigenmaps  $\varphi_\mu$  and  $\varphi$  on each of  $S_{1:1}^k, S_{1:4}^k$  and  $S_{1:9}^k$ , we observe that  $\varphi_\mu$  is significantly more robust to the distribution the input sample compare to  $\varphi$ .

One can calculate the eigenvalues of  $\Delta_m$  on  $\mathbb{S}^2 = [0, 4) \times [0, 1) \setminus \sim$  analytically: the eigenfunctions  $\phi$  of  $\Delta_m$  are given by

$$\phi(x) = \sin((ax + by)\pi) + \cos((ax + by)\pi), \quad (4.20)$$

with corresponding eigenvalues  $-(a^2 + b^2)\pi^2$ . Periodicity of  $\mathbb{S}^2$  enforces  $a = 0, \pm\frac{1}{4}, \pm\frac{2}{4}, \pm\frac{3}{4}, \dots$  and  $b = 0, \pm 1, \pm 2, \pm 3, \dots$ . Thus the first 7 eigenvalues of  $\Delta_m$  are  $0, -0.6169, -0.6169, -2.4674, -2.4674, -5.5517$  and  $-5.5517$ . To compare the first 7 eigenvalues of  $\Delta_m$  to those of  $\mathbf{B}_{1:1}^{\mu, \epsilon}$ ,  $\mathbf{B}_{1:4}^{\mu, \epsilon}$  and  $\mathbf{B}_{1:9}^{\mu, \epsilon}$ , we calculate the constant  $C_\rho$  in (4.16) by numerically integrating

$$\begin{aligned} C_\rho &= \frac{1}{2} \sum_{i,j=1}^2 \int v_i v_j q_\rho(|\mathbf{v}|) d\ell(\mathbf{v}) \\ &= \frac{c_\rho}{2} \int_0^{\sqrt{2}} \int_0^{\sqrt{2}} (x^2 + 2xy + y^2) \exp(x^2 + y^2) dx dy \quad \text{by (4.3)} \\ &\approx 0.5482, \end{aligned}$$

where the first equality is due to technical calculations done in (C.12) and (B.139) of the appendix, and the  $c_\rho = 0.8459$  was numerically calculated from

$$c_\rho \left( \int_0^{\sqrt{2}} \exp(-x^2) dx \right) = 1.$$

To provide numerical verification of Theorem 4.1.4, we plot the spectra of the first 7 eigenvalues of  $\Delta_m$ ,  $\frac{1}{\epsilon C_\rho} \mathbf{B}_{1:1}^{\mu,\epsilon}$ ,  $\frac{1}{\epsilon C_\rho} \mathbf{B}_{1:4}^{\mu,\epsilon}$  and  $\frac{1}{\epsilon C_\rho} \mathbf{B}_{1:9}^{\mu,\epsilon}$  onto a single plot; as shown in the left side Figure 4.2. Similarly, we compare the first 7 eigenvalues of  $\frac{1}{2\epsilon C_\rho} \mathcal{L}_{1:1}^\epsilon$ ,  $\frac{1}{2\epsilon C_\rho} \mathcal{L}_{1:4}^\epsilon$  and  $\frac{1}{2\epsilon C_\rho} \mathcal{L}_{1:9}^\epsilon$  to those of  $\Delta_m$ ; the results are shown in the right side plot of Figure 4.2. For  $j = 1, 4$  and  $9$ , we note the first 3 eigenvalues of  $\frac{1}{\epsilon C_\rho} \mathcal{L}_{1:j}^\epsilon$  and the first 5 eigenvalues of  $\frac{1}{\epsilon C_\rho} \mathbf{B}_{1:j}^{\mu,\epsilon}$  all approximate well the corresponding eigenvalues of  $\Delta_m$ , see Figure 4.2. However, the higher order eigenvalues of both  $\frac{1}{\epsilon C_\rho} \mathbf{B}_{1:j}^{\mu,\epsilon}$  and  $\frac{1}{\epsilon C_\rho} \mathcal{L}_{1:j}^\epsilon$  are much poorer at approximating the higher order eigenvalues of  $\Delta_m$ , because the construction of both  $\mathbf{B}_{1:j}^{\mu,\epsilon}$  and  $\mathcal{L}_{1:j}^\epsilon$  are designed to optimise the stability of  $\lambda_2$  in step 3 of Algorithm 4.1.

### Weighted domains

For our second numerical experiment, we test the ability of the eigenmap  $\varphi_\mu$  produced by Algorithm 4.1 for learning the weighted geometry of a weighted manifold. Let  $(\mathbb{S}^2, e, \mu_2)$  and  $(\mathbb{S}^2, e, \hat{\mu}_2)$  be weighted Riemannian manifolds, where  $\mu_2$  and  $\hat{\mu}_2$  are absolutely continuous measures. Respectively, the densities  $h_\mu, h_{\hat{\mu}}$  of  $\mu_2, \hat{\mu}_2$  are given by

$$h_\mu(x, y) = \sin(\pi x) + 1.2, \quad (4.21)$$

and

$$h_{\hat{\mu}}(x, y) = \sin(2\pi x) + 1.2. \quad (4.22)$$

Let  $H_{M,K}$  be as in (4.18). Due to the periodicity of  $h_\mu$  along the  $x$ -direction, one could analytically find a co-dimensional 1 surface  $\Gamma_\mu$  that yields a minimal ratio  $H_{M,K}(\Gamma_\mu)$  for  $K = 2$ ; that is, the vertical lines at  $x = 1.5$  and  $x = 3.5$ . Similar, due to the periodicity of  $h_{\hat{\mu}}$  along the  $x$ -direction, one could find a co-dimensional 1 surface  $\Gamma_{\hat{\mu}}$  that yields a minimal ratio  $H_{M,K}(\Gamma_{\hat{\mu}})$  for  $K = 4$ ; that is, the vertical lines at  $x = 3/4, x = 7/4, x = 11/4$  and  $x = 15/4$ .

Our objective is to use the eigenmap to approximate the above  $H_{M,K}$ -minimising surface  $\Gamma_\mu$  (resp.  $\Gamma_{\hat{\mu}}$ ) from point-cloud data. To initiate these numerical experiments, we draw a fixed random sample  $S^k$  from  $\mathbb{S}^2$  uniformly in i.i.d fashion, then extract the values of  $h_\mu(x_i)$  and  $h_{\hat{\mu}}(x_i)$  for each  $x_i \in S^k$ ; the random sample  $S^k$  and the heat maps of  $h_\mu$  and  $h_{\hat{\mu}}$  are shown by the first column of Figure 4.3

We apply algorithm 4.1 to the point-cloud data  $S^k$  with input measurements  $h_\mu(S^k)$ . By steps 1 - 3 of Algorithm 4.1, we set the parameter  $\epsilon = 0.055$ , with the first 7 eigenvalues of  $\frac{1}{\epsilon} \mathbf{B}^{\mu,\epsilon}$  given by  $0, -0.1448, -0.7383, -1.4026, -1.6984, -2.9513$  and  $-3.4292$ . There is a gap in the spectrum of  $\mathbf{B}^{\mu,\epsilon}$  between the  $2^{nd}$  and  $3^{rd}$  eigenvalues. Hence by step 4 of Algorithm 4.1, we set the eigenmap  $\varphi_\mu$  to

be the 1-dimensional embedding  $\varphi_\mu = \mathbf{f}_2$ , where  $\mathbf{f}_2$  is the dominant eigenvector of  $\mathbf{B}^{\mu,\epsilon}$ . To provide a meaningful display of the eigenmap  $\varphi_\mu$ , we relabel the dataset  $S^k$  according to the value of  $\varphi_\mu(S^k)$  in increasing order, and plot  $(x_i, \varphi_\mu(x_i))$  for  $1 \leq i \leq k$ ; the result is shown by top-right plot of Figure 4.3.

Next we apply algorithm 4.1 to the point-cloud data  $S^k$  with input measurements  $h_{\hat{\mu}}(S^k)$ . By steps 1 - 3 of Algorithm 4.1, we set the parameter  $\epsilon = 0.055$ , with the first 7 eigenvalues of  $\frac{1}{k\epsilon}\mathbf{B}^{k,\hat{\mu},\epsilon}$  given by 0,  $-0.2476$ ,  $-0.2591$ ,  $-0.6166$ ,  $-2.7948$ ,  $-3.6149$  and  $-3.6187$ . There is a gap in the spectrum of  $\mathbf{B}^{\hat{\mu},\epsilon}$  between the 4<sup>th</sup> and 5<sup>th</sup> eigenvalues. Hence by step 4 of Algorithm 4.1, we set the eigenmap  $\varphi_{\hat{\mu}}$  to be the 3-dimensional embedding  $\varphi_{\hat{\mu}} = \{\mathbf{f}_2, \mathbf{f}_3, \mathbf{f}_4\}$ , where  $\mathbf{f}_2, \mathbf{f}_3$  and  $\mathbf{f}_4$  are the dominant 3 eigenvectors of  $\mathbf{B}^{k,\hat{\mu},\epsilon}$ . The image of  $S^k$  under  $\varphi_{\hat{\mu}}$  is shown in bottom-right plot of Figure 4.3.

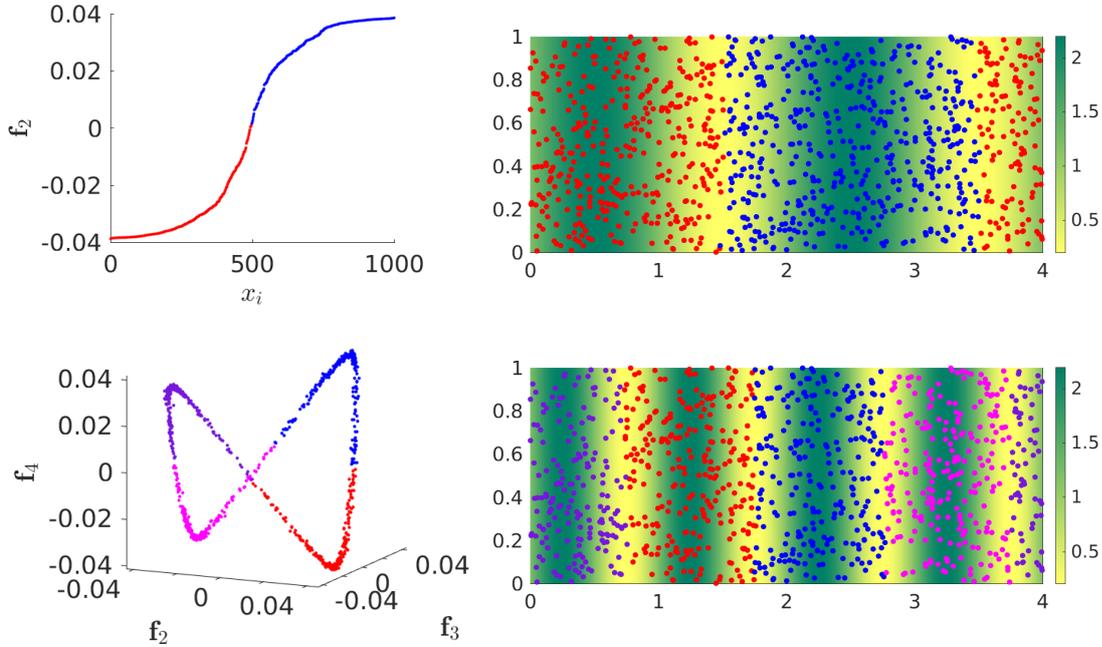


Figure 4.3: Laplacian-based manifold learning on the weighted Riemannian manifolds: First row shows  $S^k$  in  $(\mathbb{S}^2, e, \mu_2)$  on left and  $\{\mathbf{f}_2(S^k), \mathbf{f}_3(S^k)\}$  in  $\mathbb{R}^2$  on right, colours denote the clusters of  $\varphi_\mu(S^k)$ . Second row shows  $S^k$  in  $(\mathbb{S}^2, e, \hat{\mu}_2)$  on left and eigenmap  $\varphi_{\hat{\mu}}(S^k)$  in  $\mathbb{R}^2$  on right, colours denote the clusters of  $\varphi_{\hat{\mu}}(S^k)$ .

First we illustrate how the eigenmap  $\varphi_\mu = \mathbf{f}_2$  can be used to approximate the hypersurface  $\Gamma_\mu$  to minimise  $H_{M,K}(\Gamma_\mu)$ . We search for clusters of  $\varphi_\mu(S^k)$  on  $\mathbb{R}^1$ , using the  $x_i$  vs  $\varphi_\mu$  plot in the top-left of Figure 4.3: Recall that the indices of  $S^k = \{x_i\}_{i=1}^k$  is arranged in increasing order in terms of  $\varphi_\mu(S^k)$ . Hence, a small gradient on the  $x_i$  vs  $\varphi_\mu$  plot between a set of consecutive points from  $x_l$  to  $x_L$ ,  $L > l \geq 1$  shows that the set  $\{x_i\}_{i=l}^L$  is a cluster in  $\varphi_\mu(S^k)$ ; one has  $K = 2$  distinctive clusters given

by the sets  $S_1^k = \{x_i \in S^k : \mathbf{f}_2(x_i) \leq 0\}$  and  $S_2^k = \{x_i \in S^k : \mathbf{f}_2(x_i) > 0\}$ . To extract  $\Gamma_\mu$  from  $S_1^k$  and  $S_2^k$ , we colour coordinate the clusters  $S_1^k$  and  $S_2^k$  and plot their positions on  $(\mathbb{S}^2, m, \mu_2)$ , as shown in the top-right of Figure 4.3. One can see by the colour coordination that the partition  $S_1^k$  and  $S_2^k$  are good representation of the submanifolds  $\{M_1, M_2\}$  separated by the interface  $\Gamma_\mu$  (vertical lines at  $x = 1.5$  and  $x = 3.5$ ). Similar, the eigenmap  $\varphi_{\hat{\mu}} = \{\mathbf{f}_2, \mathbf{f}_3, \mathbf{f}_4\}$  can be used to approximate the hypersurface  $\Gamma_{\hat{\mu}}$  to minimise  $H_{M,K}(\Gamma_{\hat{\mu}})$ . The set  $\varphi_{\hat{\mu}}(S^k)$  in  $\mathbb{R}^3$  forms a 3-dimensional point-cloud loop; as shown in the bottom-left of Figure 4.3. We note that the 4 turning points of this point-cloud loop  $\varphi_{\hat{\mu}}(S^k)$  contains a higher concentration of data points compared to the rest of the loop. Hence, there are  $K = 4$  distinctive clusters of  $\varphi_{\hat{\mu}}(S^k)$ . Again, to extract  $\Gamma_{\hat{\mu}}$  from the 4 clusters of  $\varphi_{\hat{\mu}}(S^k)$ , we colour coordinate the clusters and plot them on  $(\mathbb{S}^2, m, \hat{\mu}_2)$ ; as show in the bottom-right of Figure 4.3. We see that the 4 clusters of  $\varphi_{\hat{\mu}}(S^k)$  provides a good representation of the submanifolds  $\{M_1, M_2, M_3, M_4\}$  separated by the interface  $\Gamma_{\hat{\mu}}$  (vertical lines at  $x = 3/4, x = 7/4, x = 11/4$  and  $x = 15/4$ ).

## 4.2 A Dynamical manifold learning for dynamics on weighted Riemannian manifolds

The weighted Laplacian eigenmap for manifold learning discussed in Section 4.1 provides the foundation for the core objective of this chapter: the formulation of a robust dynamic manifold learning method for transport barrier detection on weighted Riemannian manifolds. First we develop our dynamic manifold learning method for systems formed by a single iteration of dynamics; we then extend our techniques to situations where the dynamics arise from multiple time-step transformations. As in Section 3.1, let  $(M, m, \mu_r)$  be a weighted Riemannian manifold, and let  $T : M \rightarrow N$  be a possibly nonlinear and/or non-volume-preserving transformation acting on  $M$ , where  $N$  is another Riemannian manifold. On  $N$  we have the metric tensor  $n$  and the absolutely continuous measure  $\nu_r := \mu_r \circ T^{-1}$ , forming another weighted Riemannian manifold  $(N, n, \nu_r)$ . We assume  $M \cup N \subset \mathbb{R}^d$ , thus we have an isometric embedding map  $\Phi : (M \cup N) \rightarrow \mathbb{R}^d$ . By using the embedding map  $\Phi : (M \cup N) \rightarrow \mathbb{R}^d$ , one readily extends the domain of the kernel  $Q_{\rho,\epsilon}$  given by (4.4) to  $Q_{\rho,\epsilon} : (M \cup N) \times (M \cup N) \rightarrow \mathbb{R}^+$ .

Let  $S^k$  be a random sample drawn from  $(M, m, \mu_r)$  according to the probability density  $p$  in i.i.d fashion. The transformation  $T : M \rightarrow N$  maps  $S^k$  into  $\hat{S}^k = \{T(x_i)\}_{i=1}^k$ . As in Section 4.1, from  $S^k$ ,  $h_\mu$  and  $Q_{\rho,\epsilon}$  we form the weighted graph  $G(S^k, W^{\mu,\epsilon})$ , where  $W^{\mu,\epsilon} = \{w_{ij}^{\mu,\epsilon}\}_{i,j=1}^k$  with  $w_{ij}^{\mu,\epsilon}$  is as in (4.8). Similarly, we form a

weighted graph from  $\hat{S}^k$ ,  $h_\nu$ , and  $Q_{\rho,\epsilon}$  as follows: Since  $Q_{\rho,\epsilon} : (M \cup N) \times (M \cup N) \rightarrow \mathbb{R}^+$ , analogous to (4.5) we define  $\hat{p}^{k,\epsilon} : N \rightarrow \mathbb{R}$  by

$$\hat{p}^{k,\epsilon}(y) := \frac{1}{k} \sum_{i=1}^k Q_{\rho,\epsilon}(y, Tx_i). \quad (4.23)$$

Let  $\hat{p}$  be the probability density of the random sample  $\hat{S}^k$ . Then analogous to (4.6) and (4.7), there exists a sequence of scalars  $\{\epsilon_k\}_{k \geq 1}$  with  $\lim_{k \rightarrow \infty} \epsilon_k \rightarrow 0$  such that

$$\begin{aligned} \lim_{k \rightarrow \infty} \hat{p}^{k,\epsilon_k}(y) &= \lim_{k \rightarrow \infty} \frac{1}{k} \sum_{i=1}^k Q_{\rho,\epsilon}(y, y_i) \\ &= \lim_{\epsilon \rightarrow \infty} \int_N Q_{\rho,\epsilon}(y, z) \hat{p}(z) \omega_n^r(z) \\ &= \hat{p}(y), \end{aligned} \quad (4.24)$$

for every  $y \in N$ . Thus, analogous to (4.8) we define

$$w_{ij}^{\nu,\epsilon} := \begin{cases} \frac{Q_{\rho,\epsilon}(Tx_i, Tx_j)}{\hat{p}^{k,\epsilon}(Tx_j)} \sqrt{\frac{h_\nu(Tx_j)}{h_\nu(Tx_i)}} & \text{if } i \neq j \\ 0 & \text{if } i = j \end{cases}, \quad (4.25)$$

as the edge weights between the vertex pair  $T(x_i)$  and  $T(x_j)$ . One has the weighted graph  $G(\hat{S}^k, W^{\nu,\epsilon})$ , where  $W^{\nu,\epsilon} = \{w_{ij}^{\nu,\epsilon}\}_{i,j=1}^k$ . In this way, the transformation  $T : M \rightarrow N$  induces a dynamics of the graph  $T_G : G(S^k, W^{\mu,\epsilon}) \rightarrow G(\hat{S}^k, W^{\nu,\epsilon})$ .

Let  $\mathbf{L}^{\nu,\epsilon}$  be the graph Laplacian on  $G(\hat{S}^k, W^{\nu,\epsilon})$ . We define the *dynamic graph Laplacian* for the dynamics of graphs  $T_G : G(S^k, W^{\mu,\epsilon}) \rightarrow G(\hat{S}^k, W^{\nu,\epsilon})$  by averaging  $\mathbf{L}^{\mu,\epsilon}$  and  $\mathbf{L}^{\nu,\epsilon}$ ; that is

$$\mathbf{L}^{dyn,\epsilon} := \frac{1}{2} (\mathbf{L}^{\mu,\epsilon} + \mathbf{L}^{\nu,\epsilon}). \quad (4.26)$$

As in Section 4.1, the matrix  $\mathbf{L}^{dyn,\epsilon}$  is not symmetric, hence not self-adjoint with respect to the standard inner product on  $\mathbb{R}^k$ . Thus, similar to (4.12) we define the  $k \times k$  matrix  $\mathbf{B}^{\nu,\epsilon}$  with entries

$$\mathbf{B}_{ij}^{\nu,\epsilon} = \frac{h_\nu(Tx_i)}{\hat{p}^{k,\epsilon}(Tx_i)} \mathbf{L}_{ij}^{\nu,\epsilon}, \quad (4.27)$$

and then form

$$\mathbf{B}^{dyn,\epsilon} = \frac{1}{2} (\mathbf{B}^{\mu,\epsilon} + \mathbf{B}^{\nu,\epsilon}), \quad (4.28)$$

where  $\mathbf{B}^{\mu,\epsilon}$  is as in (4.12).

The matrix  $\mathbf{B}^{dyn,\epsilon}$  has the following spectral properties:

**Theorem 4.2.1.** *Let  $G(S^k, W^{\mu,\epsilon})$  and  $G(\hat{S}^k, W^{\nu,\epsilon})$  be weighted graphs, where  $W^{\mu,\epsilon} = \{w_{ij}^{\mu,\epsilon}\}_{i,j=1}^k$  with  $w_{ij}^{\mu,\epsilon}$  as in (4.8) and  $W^{\nu,\epsilon} = \{w_{ij}^{\nu,\epsilon}\}_{i,j=1}^k$  with  $w_{ij}^{\nu,\epsilon}$  as in (4.25). Define  $\mathbf{B}^{dyn,\epsilon}$  as in (4.28). One has*

1. The matrix  $\mathbf{B}^{dyn,\epsilon}$  is self-adjoint with respect to the standard inner product on  $\mathbb{R}^k$ .
2. The eigenvalues  $0 = \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_k$  of  $\mathbf{B}^{dyn,\epsilon}$  are nonpositive and real.
3. If each row of  $\mathbf{B}^{dyn,\epsilon}$  contains at least one nonzero entry, then  $\lambda_1$  is unit multiplicity. The eigenvector corresponding to  $\lambda_1$  is  $\mathbf{1}$ .

*Proof.* Due to the linearity of the construction 4.28, the above listed spectral properties for  $\mathbf{B}^{dyn,\epsilon}$  follows immediately from the spectral properties of  $\mathbf{B}^{\mu,\epsilon}$  and  $\mathbf{B}^{\nu,\epsilon}$  in Theorem 4.1.2.  $\square$

In Theorem 4.1.4, we have formed a connection between the graph Laplacian  $\mathbf{L}^{\mu,\epsilon}$  for  $G(S^k, W^{\mu,\epsilon})$  and the weighted Laplacian  $\Delta_\mu$  for  $(M, m, \mu_r)$ . In the dynamic setting, the analogous connection is between the dynamic graph Laplacian  $\mathbf{L}^{dyn,\epsilon}$  and the dynamic Laplacian on weighted Riemannian manifolds.

### 4.2.1 From dynamic graph Laplacian to dynamic Laplacian

Let  $\Delta^{dyn} : C^2(M, \mathbb{R}) \rightarrow C^0(M, \mathbb{R})$  be the dynamic Laplacian as in (3.28). Analogous to (4.14)-(4.15), we extend the linear operator  $\mathbf{L}^{dyn,\epsilon} : \mathbb{R}^k \rightarrow \mathbb{R}^k$  to an operator  $L^{dyn,k,\epsilon} : C^\infty(M, \mathbb{R}) \rightarrow C^\infty(M, \mathbb{R})$  as follows: Let  $f \in C^\infty(M, \mathbb{R})$ . Denote  $f_i = f(x_i)$  for each  $1 \leq i \leq k$ , then set  $\mathbf{f} = \{f_1, f_2, \dots, f_k\}$ . One has

$$\begin{aligned} (\mathbf{L}^{dyn,\epsilon} \mathbf{f})_i &= \frac{1}{2} \sum_{\substack{j=1 \\ j \neq i}}^k (w^{\mu,\epsilon} + w^{\nu,\epsilon}) (f(x_j) - f(x_i)) \\ &= \frac{1}{2} \sum_{\substack{j=1 \\ j \neq i}}^k \left( \frac{Q_{\rho,\epsilon}(x_i, x_j)}{p^{k,\epsilon}(x_j)} \sqrt{\frac{h_\mu(x_j)}{h_\mu(x_i)}} + \frac{Q_{\rho,\epsilon}(Tx_i, Tx_j)}{\hat{p}^{k,\epsilon}(Tx_j)} \sqrt{\frac{h_\nu(Tx_j)}{h_\nu(Tx_i)}} \right) (f(x_j) - f(x_i)), \end{aligned} \quad (4.29)$$

for each  $1 \leq i \leq k$ . Due to (4.29), we define the continuous extension of  $\mathbf{L}^{dyn,\epsilon}$  by

$$\begin{aligned} L^{dyn,k,\epsilon} f(x) &:= \frac{1}{2k} \sum_{j=1}^k \left( \frac{Q_{\rho,\epsilon}(x, x_j)}{p^{k,\epsilon}(x_j)} \sqrt{\frac{h_\mu(x_j)}{h_\mu(x)}} \right. \\ &\quad \left. + \frac{Q_{\rho,\epsilon}(Tx, Tx_j)}{\hat{p}^{k,\epsilon}(Tx_j)} \sqrt{\frac{h_\nu(Tx_j)}{h_\nu(Tx)}} \right) (f(x_j) - f(x)), \end{aligned} \quad (4.30)$$

for all  $x \in M$ .

**Theorem 4.2.2.** *Let  $S^k$  be an i.i.d random sample drawn from the weighted Riemannian manifold  $(M, m, \mu_r)$  according to some  $C^5$  probability density, and let*

$T : M \rightarrow N$  be a  $C^\infty$  diffeomorphism. Assume the density  $h_\mu$  of  $\mu_r$  is in  $C^3$ , bounded above and uniformly away from zero. Define  $q_\rho$ ,  $L^{\text{dyn},k,\epsilon}$  and  $\Delta^{\text{dyn}}$  as in (4.3), (4.30) and (3.28), respectively. If  $0 < \rho < \min\{\mathcal{S}_M, \mathcal{S}_N\}$ , where  $\mathcal{S}_M$  and  $\mathcal{S}_N$  are scalar curvatures of  $M$  and  $N$ , then there exists a sequence of scalars  $\{\epsilon_k\}_{k \geq 1}$  with  $\lim_{k \rightarrow \infty} \epsilon_k = 0$  such that

$$\lim_{k \rightarrow \infty} \left( \sup_{\|f\|_{C^3(M, \mathbb{R})} \leq 1} \left\| \frac{1}{\epsilon_k C_\rho} L^{\text{dyn},k,\epsilon_k} f - \Delta^{\text{dyn}} f \right\|_{C^0(M, \mathbb{R})} \right) = 0, \quad (4.31)$$

where the constant  $C_\rho$  depends on the second moment of  $q_\rho$ .

*Proof.* See Appendix C. □

## 4.2.2 Trajectory data from multiple time step dynamics

The above dynamic Laplacian construction for trajectory data formed from a single transformation, can easily be extended to situations where the trajectory data arises from multiple time-discrete transformations. Let  $(M^1, m^1, \mu_r^1)$  be a weighted Riemannian manifold, and suppose we have a sequence of diffeomorphisms  $T_t : M^t \rightarrow M^{t+1}$  for  $1 \leq t \leq \tau - 1$ , where  $M^2, M^3, \dots, M^\tau$  are full dimensional images of  $M^1$  under sequential applications of  $T_t$ . On each  $M^t$ , we have the metric  $m^t$  and the absolute continuous measure  $\mu_r^t := \mu_r^{t+1} \circ T_t$ , forming a sequence of weighted Riemannian manifolds  $\{M^t, m^t, \mu_r^t\}_{t=1}^\tau$ . Moreover, we assume all  $M^t$  is embedded in a single, possibly higher dimensional Euclidean space  $\mathbb{R}^d$ ;  $d \geq r$ . Denote the density of  $\mu_r^t$  by  $h_\mu^t$ , and let  $\Delta_{\mu,t}$  be the weighted Laplacian on  $(M^t, m^t, \mu_r^t)$  as in (3.27). We define the time-discrete dynamic Laplacian  $\Delta^{[1,\tau]} : C^2(M, \mathbb{R}) \rightarrow C^0(M, \mathbb{R})$  by

$$\Delta^{[1,\tau]} f := \frac{1}{\tau} \sum_{t=1}^{\tau} (\mathcal{H}^{(t)})^* \Delta_{\mu,t} \mathcal{H}^{(t)} f, \quad (4.32)$$

where  $\mathcal{H}^{(t)} : L^2(M^1, m^1, \mu_r^1) \rightarrow L^2(M^t, m^t, \mu_r^t)$  is as in (3.31), with adjoint  $(\mathcal{H}^{(t)})^*$ .

Let  $T^{(t)} = T_t \circ T_{t-1} \circ \dots \circ T_1$  with  $T^{(0)} = \text{Id}$ . Given an i.i.d random sample  $S^{1,k} = \{x_1, x_2, \dots, x_k\}$  drawn from  $(M^1, m^1, \mu_r^1)$  according to some  $C^5$  probability density, the sequence of transformations  $T_t$  generates  $\tau - 1$  point-cloud data from  $S^{1,k}$  of the form  $S^{t+1,k} := \{T^{(t)}(x_1), T^{(t)}(x_2), \dots, T^{(t)}(x_k)\} \in M^{t+1}$  for  $1 \leq t \leq \tau - 1$ . Denote the trajectory data formed from  $T_1, T_2, \dots, T_{\tau-1}$  acting on  $S^{1,k}$  by  $S^{[1,\tau],k}$ . To capture the evolving pointwise geometry of  $S^{[1,\tau],k}$ , we model the trajectory data  $S^{[1,\tau],k}$  as dynamics of graphs: Due to the assumption  $\cup_{t=1}^\tau M^t \subset \mathbb{R}^d$ , we have an isometric embedding map  $\Phi : \cup_{t=1}^\tau M^t \rightarrow \mathbb{R}^d$ . Hence, one can extend the domain of the kernel (4.4) to  $Q_{\rho,\epsilon} : (\cup_{t=1}^\tau M^t) \cup (\cup_{t=1}^\tau M^t) \rightarrow \mathbb{R}^+$ . Therefore, analogous to (4.23)

we define a sequence of  $p^{t,k,\epsilon} : M^t \rightarrow \mathbb{R}$  by

$$p^{t,k,\epsilon}(x) := \frac{1}{k} \sum_{i=1}^k Q_{\rho,\epsilon}(x, T^{(t-1)}x_i), \quad (4.33)$$

for  $t \in [1, \tau]$ . Let  $x_i^t = T^{(t-1)}(x_i)$  for each  $1 \leq i \leq k$  and  $1 \leq \tau$ . Analogous to (4.25), we set

$$w_{ij}^{t,\epsilon} := \begin{cases} \frac{Q_{\rho,\epsilon}(x_i^t, x_j^t)}{p^{t,k,\epsilon}(x_i^t)} \sqrt{\frac{h_\mu^t(x_j^t)}{h_\mu^t(x_i^t)}} & \text{if } i \neq j \\ 0 & \text{if } i = j \end{cases}, \quad (4.34)$$

as the edge weight between the vertex  $x_i^t$  and  $x_j^t$  for each  $1 \leq i, j \leq k$ . From (4.34), we form a sequence of weighted graphs  $\{G(S^{t,k}, W^{t,\epsilon})\}_{t=1}^\tau$ , where  $W^{t,\epsilon} = \{w_{ij}^{t,\epsilon}\}_{i,j=1}^k$ . Therefore, associate with the transformations  $T_t : M^t \rightarrow M^{t+1}$  is the dynamics of graphs  $T_{t,G} : G(S^{t,k}, W^{t,\epsilon}) \rightarrow G(S^{t+1,k}, W^{t+1,\epsilon})$ , for  $1 \leq t \leq \tau - 1$ .

For  $1 \leq t \leq \tau$ , denote by  $\mathbf{L}^{t,\epsilon}$  the graph Laplacian for  $G(S^{t,k}, W^{t,\epsilon})$  as in (4.9) with  $w_{ij}^{\mu,\epsilon}$  replaced by  $w_{ij}^{t,\epsilon}$ . The multiple time-step generalisation of  $\mathbf{L}^{dyn,\epsilon}$  (4.26) is given by

$$\mathbf{L}^{[1,\tau],\epsilon} := \frac{1}{\tau} \sum_{t=1}^\tau \mathbf{L}^{t,\epsilon}. \quad (4.35)$$

As in (4.27), we define a  $k \times k$  matrix  $\mathbf{B}^{t,\epsilon}$  with entries

$$\mathbf{B}_{ij}^{t,\epsilon} = \frac{h_\mu^t(x_i^t)}{p^{t,k,\epsilon}(x_i^t)} \mathbf{L}_{ij}^{t,\epsilon}, \quad (4.36)$$

and then as in (4.28), we form

$$\mathbf{B}^{[1,\tau],\epsilon} = \frac{1}{\tau} \sum_{t=1}^\tau \mathbf{B}^{t,\epsilon}. \quad (4.37)$$

Due to the linearity of  $\mathbf{B}^{[1,\tau],\epsilon}$ , the spectral properties listed in Theorem 4.2.1 are valid for  $\mathbf{B}^{[1,\tau],\epsilon}$  as well. Moreover, extending the linear operator  $\mathbf{L}^{[1,\tau],\epsilon} : \mathbb{R}^k \rightarrow \mathbb{R}^k$  to the operator  $L^{[1,\tau],k,\epsilon} : C^\infty(M^1, \mathbb{R}) \rightarrow C^\infty(M^1, \mathbb{R})$  as in (4.30), then again by the linearity of  $\mathbf{L}^{[1,\tau],\epsilon}$  and  $\Delta^{[1,\tau]}$ , a multiple time-discrete version of Theorem 4.1.4 holds; that is, if  $0 < \rho \leq \min_{1 \leq t \leq \tau} \{\mathcal{S}_{M^t}\}$  where  $\mathcal{S}_{M^t}$  is the scalar curvature of  $M^t$ , then there exists  $C_\rho > 0$  and a sequence of scalars  $\{\epsilon_k\}_{k \geq 1}$  with  $\lim_{k \rightarrow \infty} \epsilon_k \rightarrow 0$  such that

$$\lim_{k \rightarrow \infty} \left( \sup_{\|f\|_{C^3(M^1, \mathbb{R})} \leq 1} \left\| \frac{1}{\epsilon_k C_\rho} L^{[1,\tau],k,\epsilon_k} f - \Delta^{[1,\tau]} f \right\|_{C^0(M^1, \mathbb{R})} \right). \quad (4.38)$$

The weighted Laplacian eigenmap in Section 4.1.3 for manifold learning is a heuristic formed from a weighted version of the classical Cheeger inequality and Theorem 4.1.4. Similarly, by using the multiple time-step dynamic Cheeger inequality in Theorem 3.3.4 and (4.38), we develop a dynamic Laplacian eigenmap as a heuristic for dynamic manifold learning. We now elaborate on the computational aspect of this dynamic Laplacian eigenmap.

### 4.2.3 Algorithm for dynamic Laplacian eigenmap

Let  $S^{[1,\tau],k}$  be trajectory data generated by  $T_t : M^t \rightarrow M^{t+1}$  for  $1 \leq t \leq \tau - 1$ , where the initial points of the trajectories  $S^{1,k} = \{x_i\}_{i=1}^k$  are drawn from  $M^1 \subseteq \mathbb{R}^d$  according to some  $C^5$  probability density in i.i.d fashion. Let  $\mu_r^1, \mu_r^2, \dots, \mu_r^\tau$  be absolutely continuous measures on  $M^1, M^2, \dots, M^\tau$  respectively, which tracks the mass distribution of the objects being transported. For each  $1 \leq t \leq \tau$ , let  $\mu_{r-1}^t$  be the induced co-dimension 1 measure on  $(M^t, m^t, \mu_r^t)$  as in (3.8) with  $h_\mu$  replaced by  $h_\mu^t$ , and let  $\Gamma$  be a piece-wise continuous hypersurface in  $M^1$ , that partitions  $M^1$  into full dimensional submanifolds  $M_1^1, M_2^1, \dots, M_K^1$ .

Suppose one can track: (1) The  $\mathbb{R}^d$ -coordinates of each trajectory  $\{T^{(t)}x_i\}_{t=0}^{\tau-1}$  and (2) the mass density  $\{h_\mu^t(T^{(t)}x_i)\}_{t=1}^\tau$ , for each  $1 \leq i \leq k$ . The aim of our dynamic manifold learning method is to formulate a coordinate map  $\varphi_{dyn} : S^{1,k} \rightarrow \mathbb{R}^s$ ,  $1 \leq s \ll k$  independent of  $p$ , such that certain dynamical structures are encoded by  $\varphi_{dyn}$ . In particular, clusters  $S_1^{1,k}, S_2^{1,k}, \dots, S_K^{1,k}$  of  $S^{1,k}$  with respect to the  $\mathbb{R}^d$ -distances of  $\varphi_{dyn}(S^{1,k})$ , represents full dimensional submanifolds  $M_1^1, M_2^1, \dots, M_K^1$ , that partition  $M^1$  so that the ratio

$$H_{M,K}^{[1,\tau]}(\Gamma) = \frac{1}{\tau} \frac{\sum_{t=0}^{\tau-1} \mu_{r-1}^t(T^{(t)}\Gamma)}{\min\{\mu_r^1(M_1^1), \mu_r^1(M_2^1), \dots, \mu_r^1(M_K^1)\}}, \quad (4.39)$$

is small; we call the coordinate map  $\varphi_{dyn}$  the *dynamic Laplacian eigenmap*.

The algorithm for dynamic Laplacian eigenmap construction is a straightforward modification of Algorithm 4.1, with the following adjustments.

1. In step 1 of Algorithm 4.1, replace  $w_{ij}^{t,\epsilon_{\min}}$  with  $\sum_{t=1}^\tau w_{ij}^{t,\epsilon_{\min}}$ , and replace  $w_{ij}^{t,\epsilon_{\max}}$  with  $\sum_{t=1}^\tau w_{ij}^{t,\epsilon_{\max}}$ .
2. In steps 2 - 4 of Algorithm 4.1, replace  $\mathbf{B}^{\mu,\epsilon_l}$  with  $\mathbf{B}^{[1,\tau],\epsilon_l}$ .
3. In steps 4 and 5 of Algorithm 4.1, Replace  $\varphi_\mu$  with  $\varphi_{dyn}$ .

**Missing data** : In many real-world applications, failures of equipment and/or errors in recording measurements may lead to the instantaneous coordinates  $T^{(t)}(x_i)$  of some trajectories to be missing or corrupted. To handle this situation, we modify the matrix  $\mathbf{B}^{[1,\tau],\epsilon}$  used in the above dynamic manifold learning algorithm as follows: Suppose at the time instant  $t$ , we have a set of missing data points  $x_i^t := T^{(t)}(x_i)$ , for each  $i \in I$  in the index set  $I$ ; that is, there are  $|I|$  data points are missing from the set  $S^{t,k}$ . Then for each  $i \in I$ , set the  $i^{th}$  row of the matrix  $\mathbf{B}^{t,\epsilon}$  to be zero so that the data points  $x_i^t$  for each  $i \in I$  have no contribution to the matrix  $\mathbf{B}^{[1,\tau],\epsilon}$ . Similar approaches for treating missing data as above were employed in [57, 9]. Also, a triangulation technique was used to treat missing data in [52].

## 4.2.4 Numerical experiments for dynamic Laplacian eigenmap

### Weak standard map on torus

For our first numerical experiment on dynamic manifold learning, we compare our new dynamic Laplacian eigenmap algorithm in Section 4.2.3 to the existing space-time diffusion map algorithm in [9]. The space-time diffusion map is created by applying the following modifications to Algorithm 4.1:

1. Let  $\mathbf{Q}^{[1,\tau],\epsilon}$  be the  $k \times k$  space-time diffusion matrix as in equation (3.15) in [9]. In step 1 of Algorithm 4.1, replace  $w_{ij}^{\mu,\epsilon}$  with  $\mathbf{Q}_{ij}^{[1,\tau],\epsilon}$ .
2. In steps 3 and 4 of Algorithm 4.1, replace  $\frac{1}{\epsilon_t} \mathbf{B}^{\mu,\epsilon_t}$  with  $\frac{1}{\epsilon_t} \mathbf{Q}^{[1,\tau],\epsilon_t}$ .
3. In steps 4 and 5 of Algorithm 4.1, replace the eigenmap  $\varphi_\mu$  with the space-time diffusion map  $\hat{\varphi}$ .

The key differences between the space-time diffusion matrix  $\mathbf{Q}^{[1,\tau],\epsilon}$  in [9] and the matrix  $\mathbf{B}^{[1,\tau],\epsilon}$  (4.37) used in our dynamic Laplacian eigenmap are: (1) we have scaled the diffusion kernel  $Q_{\rho,\epsilon}(x_i^t, x_j^t)$  in (4.34) by the factor

$$\frac{1}{p^{t,k,\epsilon}(x_j^t)} \sqrt{\frac{h_\mu^t(x_i^t)}{h_\mu^t(x_j^t)}}, \quad (4.40)$$

for each  $1 \leq i, j \leq k$  and  $1 \leq t \leq \tau$ , and (2) the matrix  $\mathbf{Q}^{[1,\tau],\epsilon}$  in [9] is modelled on a ‘double diffusion’ process based on the forward-backward transfer operator in [48]. The scaling factors (4.40) are dynamic extensions of (4.19). Importantly, as improvements over the space-time diffusion map in [9], our dynamic Laplacian eigenmap is designed to be robust to the distribution of the input trajectory data, *and* it is designed for learning dynamics acting on weighted Riemannian manifold. In the following, we illustrate the robustness of our dynamic Laplacian-based manifold learning algorithm on artificially generated trajectory data.

Let  $\mathbb{T}^2$  be a  $2\pi \times 2\pi$  torus, with metric tensor  $e$  on  $\mathbb{T}^2$  given by the Kronecker delta. Consider the ‘weak standard’ map  $T : \mathbb{T}^2 \rightarrow \mathbb{T}^2$  given by

$$T(x, y) = (x + y + \alpha \sin x, y + \alpha \sin x), \quad (4.41)$$

computed modulo  $2\pi$ , where the parameter  $\alpha = 0.971635$  is chosen so that a prominent KAM curve is destroyed, and  $\mathbb{T}^2$  exhibits both regular and chaotic motions under  $T$ ; see [51].

We consider the analysis of trajectory data that are formed by 2 applications of  $T$ . Fix  $k = 2000$ , and  $\tau = 3$ . Let  $S_a^{1,k}$  be a random sample drawn from  $\mathbb{T}^2$

uniformly in i.i.d fashion, and let  $S_b^{1,k}$  be a random sample drawn from  $\mathbb{T}^2$  according to a nonuniform distribution in i.i.d fashion, such that 40% of data points in  $S_b^{1,k}$  are concentrated in the horizontal strip  $[0, 2\pi) \times [\frac{8\pi}{5}, 2\pi] \setminus \sim$  in  $\mathbb{T}^2$ . We map the initial data  $S_a^{1,k}$  and  $S_b^{1,k}$  forward by  $T$  and  $T^2$ , forming the trajectory data  $S_a^{[1,\tau],k}$  and  $S_b^{[1,\tau],k}$ ; time slices of the trajectory data  $S_b^{[1,\tau]}$  are shown in Figure 4.4 (Since  $T$  is volume preserving, all time slices of the trajectory data  $S_a^{[1,\tau],k}$  shows a uniformly distributed point-cloud data and we do not plot this).

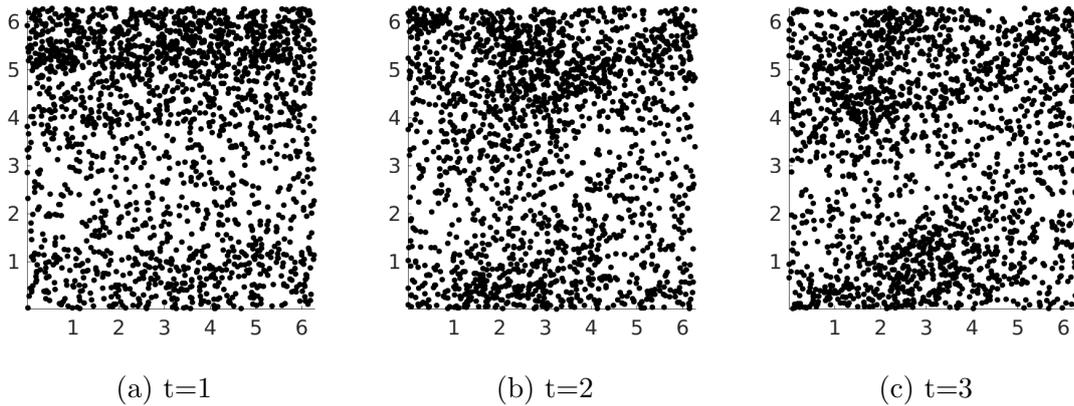


Figure 4.4: Time slices of the trajectory data  $S_b^{[1,\tau],k}$ ,  $k = 2000$  and  $\tau = 3$ .

First we form the space-time diffusion map in [9] for the trajectory data  $S_a^{[1,\tau],k}$  and  $S_b^{[1,\tau],k}$ . Let  $\mathbf{Q}_a^{[1,\tau],\epsilon}$  (resp.  $\mathbf{Q}_b^{[1,\tau],\epsilon}$ ) be the space-time diffusion matrix formed from the trajectory data  $S_a^{[1,\tau],k}$  (resp.  $S_b^{[1,\tau],k}$ ). We set the parameter  $\epsilon$  in  $\mathbf{Q}_a^{[1,\tau],\epsilon}$  and  $\mathbf{Q}_b^{[1,\tau],\epsilon}$  using step 1 - 3 of Algorithm 4.1. For the set  $S_a^{[1,\tau],k}$ , we have  $\epsilon = 0.35$  with the first 7 eigenvalues of  $\mathbf{B}_a^{[1,\tau],\epsilon}$  given by 0,  $-0.2114$ ,  $-0.2561$ ,  $-0.2742$ ,  $-0.2942$ ,  $-0.4752$  and  $-0.5310$ . For the set  $S_b^{[1,\tau],k}$ , we have  $\epsilon = 0.55$  with the first 7 eigenvalues of  $\mathbf{Q}_b^{[1,\tau],\epsilon}$  given by 0,  $-0.1970$ ,  $-0.2058$ ,  $-0.2591$ ,  $-0.2927$ ,  $-0.4139$  and  $-0.4821$ . Let  $\mathbf{g}_{2,a}$  (resp.  $\mathbf{g}_{2,b}$ ) be the first nontrivial eigenfunction of  $\mathbf{Q}_a^{[1,\tau],\epsilon}$  (resp.  $\mathbf{Q}_b^{[1,\tau],\epsilon}$ ). For  $1 \leq s \leq 5 \ll k$ , there are no gaps in the spectral of the first  $s + 2$  eigenvalues of  $\mathbf{Q}_a^{[1,\tau],\epsilon}$  and  $\mathbf{Q}_b^{[1,\tau],\epsilon}$ . Thus, by step 4 of Algorithm 4.1, we set the space-time diffusion map  $\hat{\varphi}$  to be the 1-dimensional embedding  $\hat{\varphi} = \mathbf{g}_{2,a}$  for the input trajectory  $S_a^{[1,\tau],k}$ , and  $\hat{\varphi} = \mathbf{g}_{2,b}$  for the input  $S_b^{[1,\tau],k}$ . To meaningfully display the space-time diffusion map  $\hat{\varphi}$ , we numerically interpolate the unit norm eigenfunctions  $\mathbf{g}_{2,a}/\|\mathbf{g}_{2,a}\|_2$  and  $\mathbf{g}_{2,b}/\|\mathbf{g}_{2,b}\|_2$  to  $g_{2,a} : \mathbb{T}^2 \rightarrow \mathbb{R}$  and  $g_{2,b} : \mathbb{T}^2 \rightarrow \mathbb{R}$ , respectively; the level surfaces of  $g_{2,a}$ ,  $g_{2,b}$  and their images under the pushforward operator  $\mathcal{H}^{(t)}$  (3.31) for  $t = 0, 1, 2$  are shown in Figure 4.5.

The level surface plots of  $g_{a,2}$  and its images  $\mathcal{H}^{(1)}g_{a,2}$ ,  $\mathcal{H}^{(2)}g_{a,2}$  in the first row of Figure 4.5, highlight coherent structures of the dynamical system that generated  $S_a^{[1,\tau],k}$ ; one has reflective symmetry about the horizontal line  $y = \pi$ . In contrast,

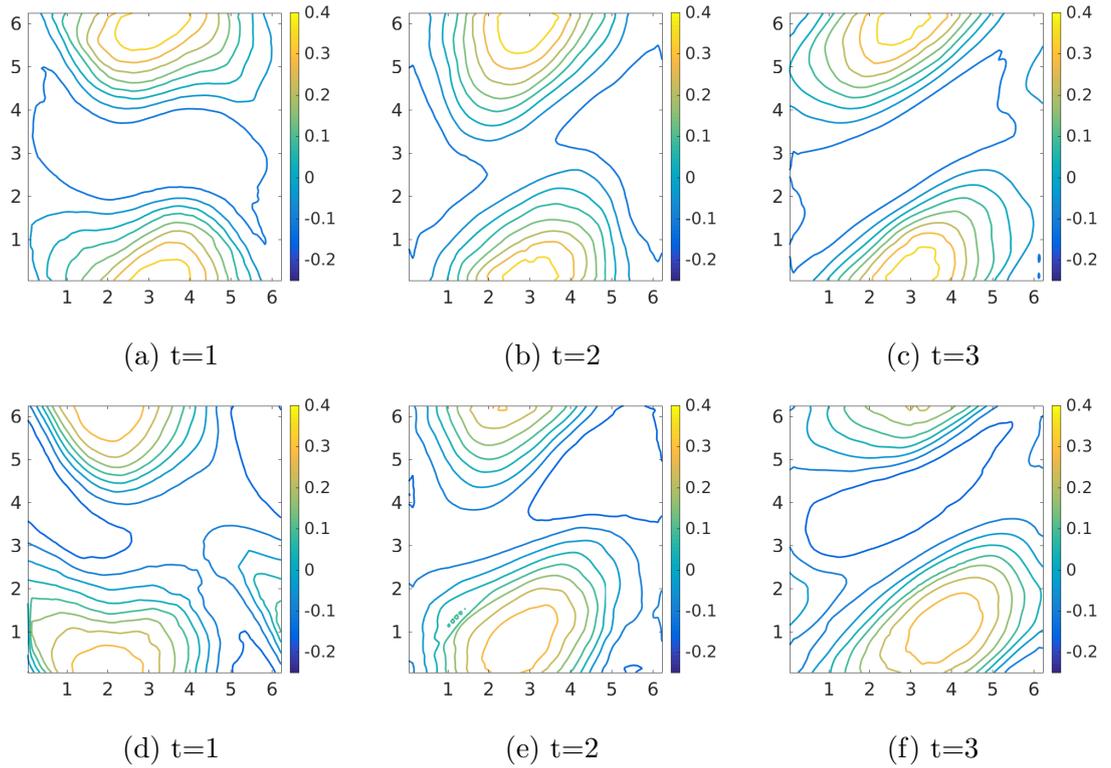


Figure 4.5: Level surfaces of the interpolated first nontrivial eigenfunctions  $g_{2,a}$  and  $g_{2,b}$  of  $\mathbf{Q}_a^{[1,\tau],\epsilon}$  and  $\mathbf{Q}_b^{[1,\tau],\epsilon}$ , respectively. (a) Level surfaces of  $g_{2,a}$ . (b) Level surfaces of  $\mathcal{H}^{(1)}g_{2,a}$ . (c) Level surfaces of  $\mathcal{H}^{(2)}g_{2,a}$ . (d) Level surfaces of  $g_{2,b}$ . (e) Level surfaces of  $\mathcal{H}^{(1)}g_{2,b}$ . (f) Level surfaces of  $\mathcal{H}^{(2)}g_{2,b}$ .

this important symmetrical feature was not present in the level surface plots of  $g_{b,2}$  and its images  $\mathcal{H}^{(1)}g_{b,2}$ ,  $\mathcal{H}^{(2)}g_{b,2}$ ; as shown in the second row of Figure 4.5. The discrepancy between  $g_{a,2}$  and  $g_{b,2}$  comes purely from the variation in distribution of the input trajectory datasets  $S_a^{[1,\tau],k}$  and  $S_b^{[1,\tau],k}$ . By construction, the trajectories of  $S_a^{[1,\tau],k}$  are uniformly distributed initially and remains uniformly distributed under  $T$ , whereas trajectories of  $S_b^{[1,\tau],k}$  are persistently more densely concentrated in the region  $[0, 2\pi) \times [\pi, 2\pi)$  compare to  $[0, 2\pi) \times [0, \pi)$ ; as shown by time slices of  $S_b^{[1,\tau],k}$  in Figure 4.4. Quantitatively, we compare the eigenfunctions  $g_{a,2}$  and  $g_{b,2}$  by computing  $\|g_{a,2} - g_{b,2}\|_2 = 0.7967$ .

We repeat the above numerical experiments conducted on  $S_a^{[1,\tau],k}$  and  $S_b^{[1,\tau],k}$  using our dynamic Laplacian eigenmap described in Section 4.2.3. Denote by  $\mathbf{B}_a^{[1,\tau],\epsilon}$  (resp.  $\mathbf{B}_b^{[1,\tau],\epsilon}$ ) the matrix formed from the trajectory data  $S_a^{[1,\tau],k}$  (resp.  $S_b^{[1,\tau],k}$ ) in step 2 of Algorithm 4.1. We set the parameter  $\epsilon$  in  $\mathbf{B}_a^{[1,\tau],\epsilon}$  and  $\mathbf{B}_b^{[1,\tau],\epsilon}$  using steps 1 - 3 of Algorithm 4.1. For the set  $S_b^{[1,\tau],k}$ , we have  $\epsilon = 0.55$  with the first 7 eigenvalues of  $\mathbf{B}_b^{[1,\tau],\epsilon}$  given by 0,  $-3.3861$ ,  $-4.0110$ ,  $-4.2742$ ,  $-4.3276$ ,  $-7.2880$  and  $-7.6287$ . For the set  $S_b^{[1,\tau],k}$ , we have  $\epsilon = 0.65$  with the first 7 eigenvalues of  $\mathbf{B}_b^{[1,\tau],\epsilon}$  given by

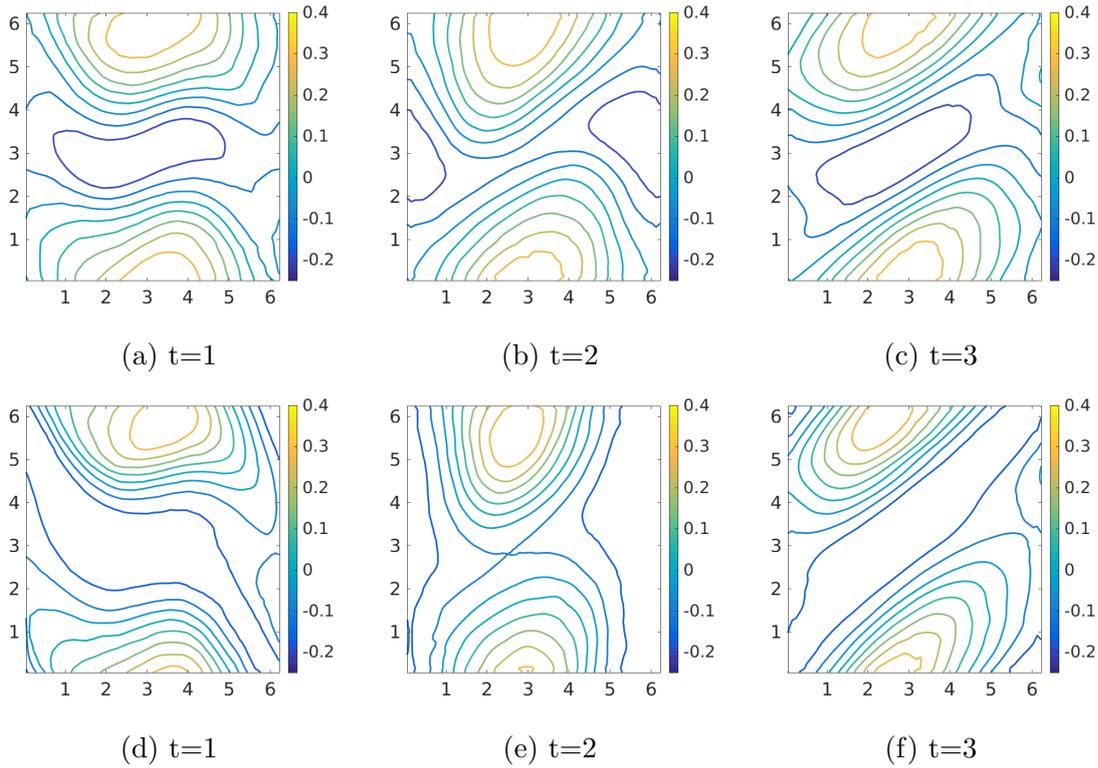


Figure 4.6: Level surfaces of the interpolate first nontrivial eigenfunctions  $f_{2,a}$  and  $f_{2,b}$  of  $\mathbf{B}_a^{[1,\tau],\epsilon}$  and  $\mathbf{B}_b^{[1,\tau],\epsilon}$ , respectively. (a) Level surfaces of  $f_{2,a}$ . (b) Level surfaces of  $\mathcal{H}^{(1)}f_{2,a}$ . (c) Level surfaces of  $\mathcal{H}^{(2)}f_{2,a}$ . (d) Level surfaces of  $f_{2,b}$ . (e) Level surfaces of  $\mathcal{H}^{(1)}f_{2,b}$ . (f) Level surfaces of  $\mathcal{H}^{(2)}f_{2,b}$ .

0,  $-3.0032$ ,  $-3.1241$ ,  $-4.3756$ ,  $-4.9048$ ,  $-6.6646$  and  $-8.9010$ . Let  $\mathbf{f}_{2,a}$  (resp.  $\mathbf{f}_{2,b}$ ) be the first nontrivial eigenfunction of  $\mathbf{B}_a^{[1,\tau],\epsilon}$  (resp.  $\mathbf{B}_b^{[1,\tau],\epsilon}$ ). For  $1 \leq s \leq 5 \ll k$ , there are no gaps in the spectrum of in the first  $s + 2$  eigenvalues of  $\mathbf{B}_a^{[1,\tau],\epsilon}$  and  $\mathbf{B}_b^{[1,\tau],\epsilon}$ . Thus, by step 4 of Algorithm 4.1, we set the dynamic Laplacian eigenmap  $\varphi_{dyn}$  to be a 1-dimensional embedding  $\varphi_{dyn} = \mathbf{f}_{2,a}$  for the input trajectory  $S_a^{[1,\tau],k}$ , or  $\hat{\varphi} = \mathbf{f}_{2,b}$  for the input  $S_b^{[1,\tau],k}$ . To meaningfully display the dynamic Laplacian eigenmap  $\varphi_{dyn}$ , we numerically interpolate the unit norm eigenfunctions  $\mathbf{f}_{2,a}/\|\mathbf{f}_{2,a}\|_2$  (resp.  $\mathbf{f}_{2,b}/\|\mathbf{f}_{2,b}\|_2$ ) to  $f_{2,a} : \mathbb{T}^2 \rightarrow \mathbb{R}$  (resp.  $f_{2,b} : \mathbb{T}^2 \rightarrow \mathbb{R}$ ); the level surfaces of the eigenfunctions  $f_{2,a}$ ,  $f_{2,b}$  and their images under the pushforward operator  $\mathcal{H}^{(t)}$  (3.31) are shown in Figure 4.5.

The level surface plots of  $f_{a,2}$  and its images  $\mathcal{H}^{(1)}f_{a,2}$ ,  $\mathcal{H}^{(2)}f_{a,2}$  shown in the first row of Figure 4.6, are consistent with the corresponding level surface plots of the space-time diffusion matrix in Figure 4.5. Moreover, despite the nonuniform distribution of the trajectory data  $S_b^{[1,\tau],k}$ , the level surface plots of  $f_{b,2}$  and its images  $\mathcal{H}^{(1)}f_{b,2}$ ,  $\mathcal{H}^{(2)}f_{b,2}$ , still highlight coherent structures of the dynamical system that generated  $S_b^{[1,\tau],k}$ ; as shown in the second row of Figure 4.5. Quantitatively, we

compare the eigenfunctions  $f_{a,2}$  and  $f_{b,2}$  by computing  $\|f_{a,2} - f_{b,2}\|_2 = 0.4895$ . Since  $\|g_{a,2} - g_{b,2}\|_2 = 0.7967$  for the space-time diffusion maps  $g_{a,2}$  and  $g_{b,2}$ , we conclude that the current dynamic Laplacian eigenmap is more robust to the distribution of the trajectory data compared to the dynamic manifold learning method [9].

### Standard map on weighted torus

Next, we test the ability of our dynamic Laplacian eigenmap to approximate dynamical structures of non-volume-preserving transformations on weighted Riemannian manifolds. Let  $M = N = \mathbb{T}^2$  be a  $2\pi \times 2\pi$  torus as in the last numerical example. We equip  $M$  with the absolutely continuous measure  $\mu_2$  with density  $h_\mu(x, y) = \frac{1}{8\pi^2}(\sin(y - \pi/2) + 2)$ , then form the weighted Riemannian manifold  $(M, e, \mu_2)$ . We consider the transformation  $T = T_4 \circ T_3$  acting on  $M$ , where

$$\begin{aligned} T_3(x, y) &= (x + 0.3x \cos(2x), y), \\ T_4(x, y) &= (x + y, y + 8 \sin(x + y)), \end{aligned}$$

both computed modulo  $2\pi$ . The weighted Riemannian manifold  $(M, e, \mu_2)$  is transformed into  $(N, e, \nu_2)$  under  $T$ , where  $\nu_2 = \mu_2 \circ T^{-1}$ . The nonlinear, non-volume-preserving transformation  $T$  on  $(M, e, \mu_2)$  was studied in Section 3.5.4: the dynamic spectral method introduced in Chapter 3 finds a partition  $\{M_1, M_2\}$  of  $M$  with small ratio  $H_{M,K}^{[1,\tau]}$  (4.39) for  $K = 2$ ; as shown in Figure 4.7, (a) and (c). The interface  $\Gamma$  that separates  $M_1$  and  $M_2$  is given by the level surface  $\phi_2 = -9.276 \times 10^{-4}$ , where  $\phi_2$  is the unit norm, first nontrivial eigenfunction of the dynamic Laplacian  $\Delta^{dyn}$  (3.28).

We now consider the situation where both the manifold  $M$  and the dynamics  $T$  are unknown; that is, the dynamic spectral method introduced in Chapter 3 cannot be applied to find the partition  $\{M_1, M_2\}$ . Suppose we have a set of trajectory data initial sampled from  $M$ , and suppose we can track the measurements  $h_\mu$  and  $h_\nu$  at the trajectories locations. We illustrate how our dynamic manifold learning method can be used to approximate the partition  $\{M_1, M_2\}$  from the given information. To initiate this numerical experiment, we start by generating a set of 2000 trajectories, by drawing 2000 initial data points  $S^{1,k}$  from  $(\mathbb{T}^2, e, \mu_2)$  uniformly in i.i.d fashion (Importantly, the distribution of  $S^{1,k}$  does not depend on  $\mu_2$ ). We then map each point in  $S^{1,k}$  forward by  $T$ . Denote by  $S^{2,k}$  the image of  $S^{1,k}$  under  $T$ , and denote the trajectory data formed from  $S^{1,k}$  and  $S^{2,k}$  by  $S^{[1,\tau],k}$ ;  $\tau = 2$ ,  $k = 2000$ .

We form the dynamic Laplacian eigenmap  $\varphi_{dyn}$  from the above trajectory data using Algorithm 4.1 together with the modifications stated in Section 4.2.3. By steps 1 - 3 of Algorithm 4.1, we have  $\epsilon = 0.5$  with first 7 eigenvalues of  $\frac{1}{\epsilon} \mathbf{B}^{[1,\tau],\epsilon}$  given by

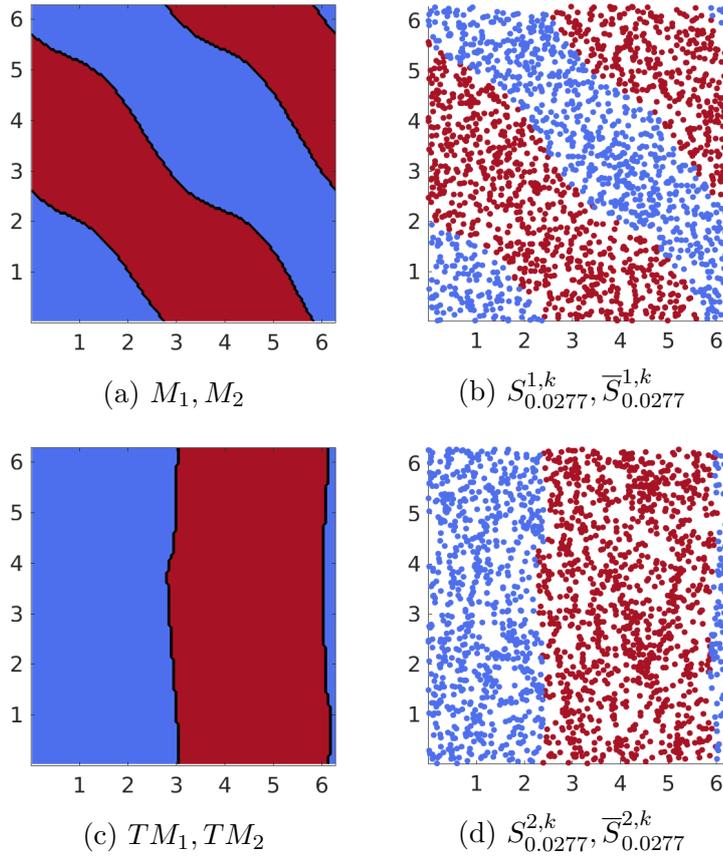


Figure 4.7: Approximation of  $\{M_1, M_2\}$  from the trajectory data  $S^{[1,\tau],k}$  using dynamic Laplacian eigenmap. (a) Colors represent partition elements  $M_1$  and  $M_2$  of  $M$ . (b) Colors represent the partition elements  $S_\beta^{1,k}$  and  $\bar{S}_\beta^{1,k}$  of  $S^{1,k}$ , with  $\beta = 0.0277$ . (c) The image of (a) under  $T$ . (d) Image of (b) under  $T$ .

0,  $-0.0768$ ,  $-0.0832$ ,  $-0.1529$ ,  $-0.1984$ ,  $-0.2414$  and  $-0.2717$ . For  $1 \leq s \leq 5 \ll k$ , there are no gaps in the spectrum of the first  $s+2$  eigenvalues of  $\mathbf{B}^{[1,\tau],\epsilon}$ . Thus, we set  $s = 1$  in accordance with step 4 of Algorithm 4.1. Therefore by step 5 of Algorithm 4.1, the dynamic Laplacian eigenmap is given by a 1-dimensional embedding map  $\varphi_{dyn} := \mathbf{f}_2$ , where  $\mathbf{f}_2$  is the first nontrivial eigenvector of  $\mathbf{B}^{[1,\tau],\epsilon}$ .

Since the dynamic Laplacian eigenmap  $\varphi_{dyn} = \mathbf{f}_2$  is 1-dimensional, for some  $\beta \in \mathbb{R}$  the subsets  $S_\beta^{1,k} := \{x_i \in S^{1,k} : \mathbf{f}_2(x_i) \leq \beta\}$  and  $\bar{S}_\beta^{1,k} := \{x_i \in S^{1,k} : \mathbf{f}_2(x_i) > \beta\}$ , are clusters of  $S^{1,k}$  with respect to the  $\mathbb{R}^1$ -distances of  $\varphi_{dyn}(S^{1,k})$ . To find  $\beta \in \mathbb{R}$  such that  $\{S_\beta^{1,k}, \bar{S}_\beta^{1,k}\}$  is a good representation of  $\{M_1, M_2\}$ , we interpolate the eigenfunction  $\mathbf{f}_2 / \|\mathbf{f}_2\|_2$  to  $f_2 : M \rightarrow \mathbb{R}$ , and numerically determine the level surface  $f_2 = \beta$  that minimises  $H_{M,2}^{[1,\tau]}(\{f_2 = \beta\})$ ; there are at most  $k - 1$  clusters of the form  $\{S_\beta^{1,k}, \bar{S}_\beta^{1,k}\}$ . We found optimally  $\beta = 0.0277$ ; the level surfaces (with the level surface  $f_2 = 0.0277$  in black) of  $f_2$  and their images under  $\mathcal{H}$  are shown in Figure 4.8, (a) and (c). To evaluate the accuracy of our dynamic Laplacian eigenmap  $f_2$  for

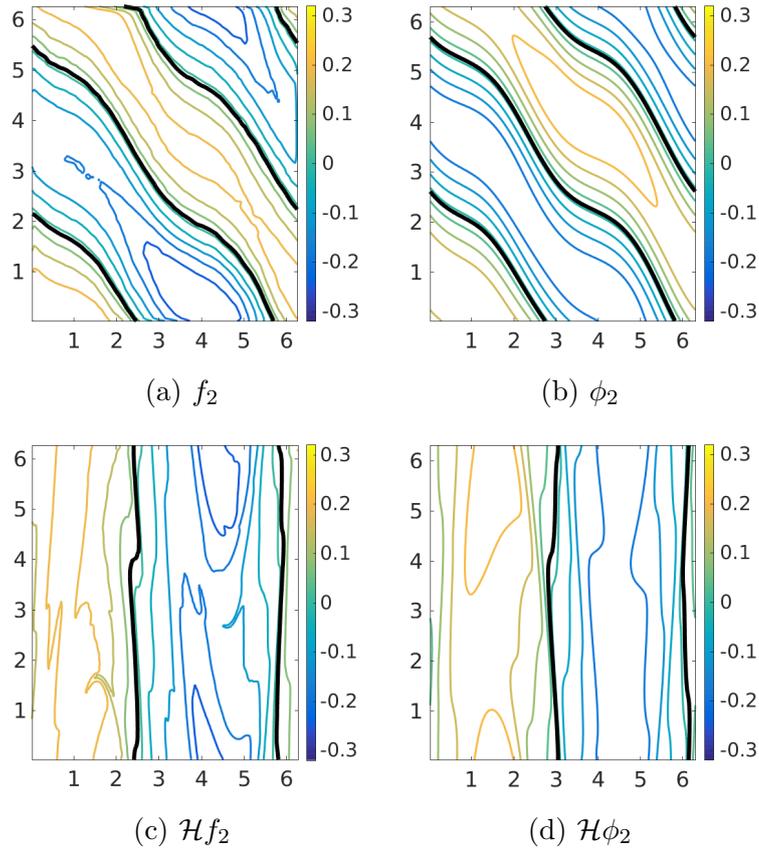


Figure 4.8: Approximation of the dominant eigenfunction  $\phi_2$  of  $\Delta^{dyn}$  from the interpolate first nontrivial eigenfunction  $f_2$  of  $\mathbf{B}^{[1,\tau],\epsilon}$ . (a) Coloured lines are level surfaces of  $f_2$ , with black line as the level surface  $f_2 = 0.0277$ . (b) Coloured lines are level surfaces of  $\phi_2$ , with black line as the level surface  $\phi_2 = -9.276 \times 10^{-4}$ . (c) Image of (a) under  $\mathcal{H}$  (d) Image of (b) under  $\mathcal{H}$ .

approximating the eigenfunction  $\phi_2$ , we compute the  $L^2$ -error  $\|\phi_2 - f_2\|_2 = 0.4208$ , and display the level surfaces (with the level surface  $\phi_2 = -9.276 \times 10^{-4}$  in black) of  $\phi_2$  in Figure 4.8, (b) and (d).

### Ocean drifters

For the last numerical example, we study geophysical fluid dynamics using our new dynamic manifold learning method on real-world data. In particular, we consider the trajectory data from the Global Ocean Drifter Program available from AOML/NOAA Drifter Data Assembly Center (<http://www.aoml.noaa.gov/envids/gld/>). We focus on the years 2005 – 2009 and restrict to those drifters that have a minimum lifetime of one year within this five-year time span; there are a total of 2267 trajectories that satisfies these conditions. We output the position of these 2267 trajectories (in longitude, latitude coordinates) every month, i.e. the length

of our trajectories is 60 months. Denote the ocean drifter data by  $S^{[1,\tau],k}$ ,  $\tau = 60$ ,  $k = 2267$ . The set  $S^{[1,\tau],k}$  was studied by [57] and [9].

We apply our new dynamic manifold learning method to  $S^{[1,\tau],k}$ , with the objective of identifying transport barriers under ocean dynamics. In particular, we partition  $S^{[1,\tau],k}$  into bundles of trajectories, such that initial points of these  $K$ -clusters represent submanifolds  $M_1^1, M_2^1, \dots, M_K^1$  of the ocean surface with small ratio (4.39). We note that typical drifters do not operate over the whole five years, and there are also gaps in observations when there is a failure in recording the drifter location (see Figure 17 in [57] for statistics of the drifters). Thus, we are dealing with highly incomplete trajectory data. Therefore, we apply the missing data treatment discussed at the end of Section 4.2.3. We also note that the distribution of the drifter locations are nonuniform; as shown by plots of the ocean drifter locations in Figure 4.10.

We form the dynamic Laplacian eigenmap  $\varphi_{dyn}$  from the trajectory data  $S^{[1,\tau],k}$  using Algorithm 4.1 together with the modifications stated in Section 4.2.3. By steps 1 - 3 of Algorithm 4.1 we have  $\epsilon = 0.105$ , with first 7 eigenvalues of  $\frac{1}{\epsilon} \mathbf{B}^{[1,\tau],\epsilon}$  given by 0,  $-0.2317$ ,  $-0.6777$ ,  $-1.4946$ ,  $-2.1596$ ,  $-2.9291$  and  $-3.3550$ . There is a gap in the spectrum between the 3<sup>rd</sup> and 4<sup>th</sup> eigenvalues of  $\mathbf{B}^{[1,\tau],\epsilon}$ . Hence by steps 4 and 5 of Algorithm 4.1, the dynamic Laplacian eigenmap  $\varphi_{dyn}$  is given by a 2-dimensional embedding  $\varphi_{dyn} := \{\mathbf{f}_2, \mathbf{f}_3\}$ , where  $\mathbf{f}_2$  and  $\mathbf{f}_3$  are the first two nontrivial eigenfunction of  $\mathbf{B}^{[1,\tau],\epsilon}$ , respectively. Let  $S^{1,k}$  denote the initial point-cloud of the trajectory data  $S^{[1,\tau],k}$ . The image of  $S^{1,k}$  under the dynamic Laplacian eigenmap  $\varphi_{dyn}$  is shown in Figure 4.9. To cluster the trajectory data  $S^{[1,\tau],k}$  into  $K = 5$  parts (we choose  $K = 5$  for comparison with [57] and [9]), we partition the initial point-cloud  $S^{1,k}$  by applying  $K$ -mean partitioning algorithm to  $\varphi_{dyn}(S^{1,k})$ . We colour the partition elements of  $\varphi_{dyn}(S^{1,k})$ , and display all drifter locations on the ocean surface at certain months as in [57] (In addition to [57], we have chosen to show the ocean drifter location for September 2008 to highlight drifters in the Arctic ocean); as shown in Figure 4.10.

In each plot of Figure 4.10, by the geographical locations of drifters that are grouped together, the 5 clusters we found can roughly be described as Northern Pacific (red), Southern Pacific (yellow), Northern Atlantic (green), Southern Atlantic/Indian Ocean (blue) and Arctic Ocean (purple). From the colour coordinations in Figure 4.10, one can see that a significant number of drifters have moved between Southern Atlantic/Indian Ocean and Southern Pacific over the 60 months time period we have tracked. The large quantity of drifter mixing between Southern Atlantic/Indian Ocean and Southern Pacific, is an indication that the strength of

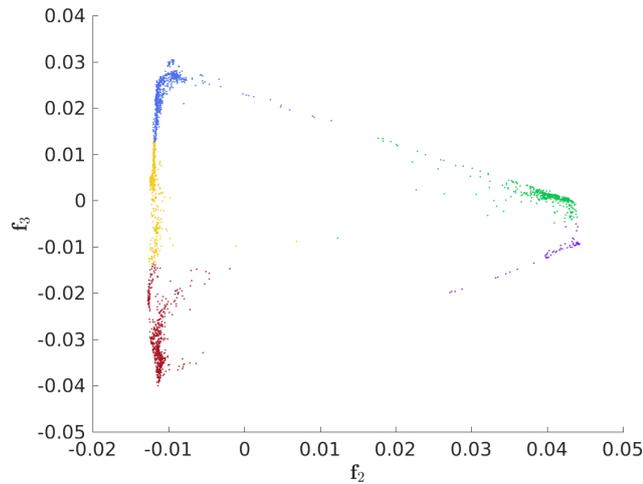


Figure 4.9: Dynamic Laplacian eigenmap  $\varphi_{dyn}$  on ocean drifter data. Colours represent partition elements produced by  $K$ -mean algorithm with  $K = 5$ .

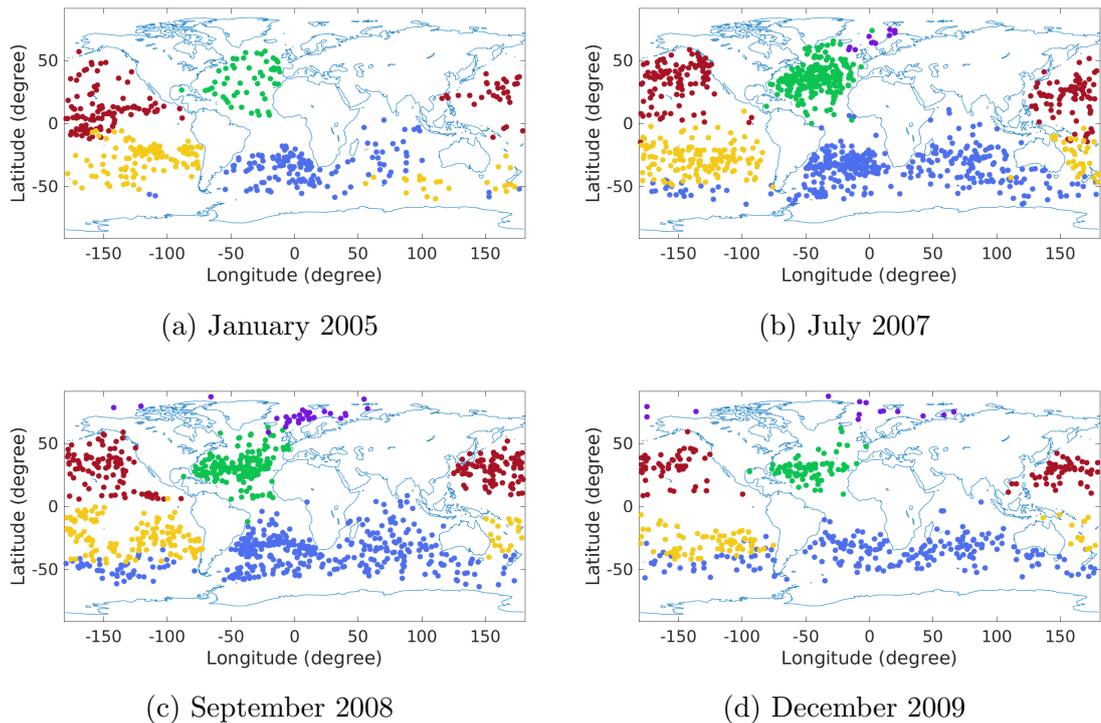


Figure 4.10: Partition of ocean drifter data using dynamic Laplacian eigenmap  $\varphi_{dyn}$ , colours represent drifters that are grouped together. Shown are location of all ocean drifters overlap in periods of 15 months. Shown are 4 time slices of the ocean drifters.

transport barrier that separates these oceans is weak. The strength of the transport barrier between Southern Atlantic/Indian Ocean and Southern Pacific is reflected by the dynamic Laplacian eigenmap in Figure 4.9; the blue and yellow dots are not well separated in terms of the coordinate  $\{\mathbf{f}_2, \mathbf{f}_3\}$ . In contrast, the blue and green

dots in Figure 4.9 are well separated, as they are concentrated at the extreme ends of the coordinate  $f_2$ . Correspondingly, there are only a few drifters that have flowed between the Southern Atlantic/Indian Ocean (blue) and Northern Atlantic (green); as shown in the plots of Figure 4.10. In a similar manner, we can make speculation on the strength of transport barriers between any oceans by studying the dynamic Laplacian eigenmap in Figure 4.9.

### 4.3 Conclusion to the chapter

In this chapter, we had considered the problem of approximating transport barriers for non-volume-preserving dynamics, from trajectory data that are randomly distributed on weighted Riemannian manifolds. We have extended the standard Laplace-based manifold learning method [10, 11], concerned with the approximation of low-dimensional feature of manifolds from uniformly distributed point-cloud data, to a situation where the manifolds are weighted, and the point-cloud data are randomly distributed (Importantly, the distribution of the point-cloud data need not coincide with the density of the manifold). We constructed a scalar weighted graph from randomly distributed point-cloud data, and show that the Laplacian of this graph converges to the weighted manifold Laplacian operator. We demonstrated numerically that the eigenfunctions of the scalar weighted graph Laplacian are (1) robust to the distributions of the point-cloud data, and (2) encode the weighted geometry of weighted Riemannian manifolds.

We then considered a dynamic extension of our robust Laplacian-based manifold learning method. We modelled trajectory data as dynamics of graphs, and constructed a dynamic Laplacian matrix for dynamics of graphs. We show that this dynamic Laplacian matrix on dynamic graphs converges to the dynamic Laplacian on weighted Riemannian manifolds (Importantly, the weights on these Riemannian manifolds need not be related to the evolving distributions of the trajectory data). We demonstrated numerically that the eigenfunctions of our dynamic Laplacian matrix are (1) robust to the distribution of the trajectory data, and (2) can be used to approximate solutions to the dynamic isoperimetric problem considered in Chapter 3.

# Appendices



# Appendix A

Let  $G = G(V, E)$  be a simple, connected graph, where  $V = \{v_1, v_2, \dots, v_k\}$  is the vertex set and  $E \in V \times V$  the edge set. On  $G$ , let  $d(v_i)$  denote the degree of the vertex  $v_i \in V$ ,  $D(V')$  the total degree of the collection of all vertices in  $V' \subseteq V$  and  $C(V_1, V_2)$  the partition boundary between  $V_1, V_2 \subseteq V$  as in Section 1.1. Define the  $k \times k$  diagonal matrix  $\mathbf{D}_{ii} = d(v_i)$ , and the graph Laplacian  $\mathbf{L}$  by (1.5).

Let  $\pi$  be a graph transformation on  $G$ , induced by the vertex permutation  $\pi_v$  on  $V$  and edge permutation  $\pi_e$  on  $E$  as in 1.1. Let  $H_G^d$  be the ratio given by (2.10), and let the matrices  $\mathbf{L}^\pi, \mathcal{L}^{dyn}$  be as in (2.15), (2.18) respectively. Define the vertex label permutation  $\pi_p$  associated with  $\pi_v$  by  $\pi_p(v_i) = v_{\pi_p(i)}$ .

## A.1 The proof of Theorem 2.2.2

To obtain the desire spectral properties for  $\mathcal{L}^{dyn}$ , we compute the Rayleigh quotient for  $\mathcal{L}^{dyn}$ . Let  $\mathbf{g}$  be a vector in  $\mathbb{R}^k$ , and  $\langle \cdot, \cdot \rangle$  the inner product on  $\mathbb{R}^k$ . Denote by  $\sum_{i \sim j}$  the summation over the set of all pairs of vertices such that  $[v_i, v_j] \in E$ . One has

$$(\mathbf{L}\mathbf{f})_i = \sum_{j=1}^k \mathbf{L}_{ij} f_j = \sum_{i \sim j} (f_j - f_i), \quad (\text{A.1})$$

for each  $1 \leq i \leq k$ . Hence

$$\begin{aligned} \langle \mathbf{f}, \mathbf{L}\mathbf{f} \rangle &= \sum_{i \sim j} f_i (f_j - f_i) \\ &= \sum_{\substack{i \sim j \\ i < j}} f_i (f_j - f_i) + \sum_{\substack{i \sim j \\ i > j}} f_i (f_j - f_i) \\ &= \sum_{\substack{i \sim j \\ i < j}} f_i (f_j - f_i) + \sum_{\substack{j \sim i \\ j > i}} f_j (f_i - f_j) \\ &= - \sum_{\substack{i \sim j \\ i < j}} (f_i - f_j)^2. \end{aligned}$$

For the remainder of Appendix A, we write  $\sum_{\substack{i \sim j \\ i < j}}$  as  $\sum_{i \sim j}$  unless otherwise stated. Note that  $(\mathbf{P}\mathbf{f})_i = f_{\pi_p(i)}$ ; thus

$$\langle \mathbf{f}, \mathbf{L}^\pi \mathbf{f} \rangle = \langle \mathbf{P}\mathbf{f}, \mathbf{L}\mathbf{P}\mathbf{f} \rangle = \sum_{i \sim j} (f_{\pi_p(i)} - f_{\pi_p(j)})^2. \quad (\text{A.2})$$

Let  $\mathbf{g} = \mathbf{D}^{1/2}\mathbf{f}$ . By (A.1) and (A.2), the Rayleigh quotient for  $\mathcal{L}^{dyn}$  is given by

$$\frac{\langle \mathbf{g}, \mathcal{L}^{dyn} \mathbf{g} \rangle}{\langle \mathbf{g}, \mathbf{g} \rangle} = \frac{\langle \mathbf{f}, (\mathbf{L} + \mathbf{L}^\pi) \mathbf{f} \rangle}{2\langle \mathbf{D}^{1/2}\mathbf{f}, \mathbf{D}^{1/2}\mathbf{f} \rangle} = \frac{\sum_{i \sim j} (f_i - f_j)^2 + (f_{\pi_p(i)} - f_{\pi_p(j)})^2}{2\sum_{i=1}^k d(v_i) f_i^2} =: R(\mathbf{f}). \quad (\text{A.3})$$

Due to the characterisation (A.3), the proof of Theorem 2.2.2 proceeds as in the static graph situation. Parts 1, 2 and 3 follow from Theorem 1 in [5] (the  $\Delta$  in [5] is our  $\mathbf{L}$ , so the obvious modifications are applied to treat  $\mathcal{L} = \mathbf{D}^{-1/2}\mathbf{L}\mathbf{D}^{-1/2}$ ), and part 4 follows from the Courant-Fischer theorem (see e.g. Theorem 4.2.11 [76]).

## A.2 The proof of Theorem 2.2.3(2)

We define the sub-vertex sets  $S^i = \{v_1, v_2, v_3, \dots, v_i\}$  and  $\overline{S^i} = \{v_{i+1}, \dots, v_k\}$ , with the set extensions  $S^0 = \overline{S^k} = \emptyset$ ;  $\{S^i, \overline{S^i}\}$  partitions  $G$  for each  $i = 0, 1, \dots, k$ . We use the abbreviations  $C(S^i) = C(S^i, \overline{S^i})$  and  $\hat{D}(S^i) = \min\{D(S^i), D(\overline{S^i})\}$ . The *ordered cut value*  $\alpha$  is defined by

$$\alpha := \min_{1 \leq i \leq k} \frac{|C(S^i)| + |C(\pi_v^{-1}(S^i))|}{\hat{D}(S^i) + \hat{D}(\pi_v^{-1}(S^i))}. \quad (\text{A.4})$$

The following Lemma forms the crucial link between the cardinality of the partition boundary  $C(\pi_v^{-1}(S^i))$  and a vector  $\mathbf{f} \in \mathbb{R}^k$ .

**Lemma A.2.1.** *Let  $G = G(V, E)$  be a simple, connected graph with  $|V| = k$ , and  $\pi = (\pi_v, \pi_e)$  a graph isomorphism. Let  $\pi_p$  be the vertex label permutation associated with  $\pi_v$ . If  $\mathbf{f} \in \mathbb{R}^k$  satisfies  $f_i \leq f_{i+1}$  for all  $i = 1, \dots, k-1$ , then*

$$\sum_{i \sim j} |f_{\pi_p(i)} - f_{\pi_p(j)}| = \sum_{i=1}^{k-1} |f_i - f_{i+1}| \cdot |C(\pi_v^{-1}(S^i))|. \quad (\text{A.5})$$

*Proof.* We perform induction on the number of vertices of  $G$ . For  $k = 2$ ,  $V = \{v_1, v_2\}$  and  $E = [v_1, v_2]$ , the vertex permutation  $\pi_v$  either fixes both vertices or interchanges them. In both cases of  $\pi_v$ , the LHS of (A.5) is

$$\sum_{i \sim j} |f_{\pi_p(i)} - f_{\pi_p(j)}| = |f_1 - f_2|,$$









*Proof.* Since  $d(\pi_v^{-1}v_r) \leq \sum_{i=r+1}^k d(\pi_v^{-1}v_i)$  and  $f_i$  is positive with  $f_i \leq f_{i+1}$  for each  $i \geq r$ , we have

$$d(\pi_v^{-1}r)f_r^2 \leq \sum_{i=r+1}^k d(\pi_v^{-1}v_i)f_i^2 \leq \sum_{\substack{i=1 \\ i \neq r}}^k d(\pi_v^{-1}v_i)f_i^2. \quad (\text{A.12})$$

Therefore

$$\begin{aligned} & \left( \sum_{i=1}^k d(v_i)f_i^2 \right) \left( \sum_{i=1}^k d(v_i)f_i^2 + d(\pi_v^{-1}v_i)f_i^2 \right) \\ = & \left( \sum_{i=1}^k d(v_i)f_i^2 \right) \left( \sum_{i=1}^k d(v_i)f_i^2 + \sum_{\substack{i=1 \\ i \neq r}}^k d(\pi_v^{-1}v_i)f_i^2 \right) + \left( \sum_{i=1}^k d(v_i)f_i^2 \right) d(\pi_v^{-1}v_r)(f_r)^2 \\ & \left( \sum_{i=1}^k d(v_i)f_i^2 + \sum_{\substack{i=1 \\ i \neq r}}^k d(\pi_v^{-1}v_i)f_i^2 \right)^2 + \left( \sum_{i=1}^k d(v_i)f_i^2 \right) d(\pi_v^{-1}r)(f_r)^2 \\ = & \left( \sum_{i=1}^k d(v_i)f_i^2 + \sum_{\substack{i=1 \\ i \neq r}}^k d(\pi_v^{-1}v_i)f_i^2 \right)^2 - \left( \sum_{\substack{i=1 \\ i \neq r}}^k d(\pi_v^{-1}v_i)f_i^2 \right) \left( \sum_{i=1}^k d(v_i)f_i^2 \right) + \left( \sum_{\substack{i=1 \\ i \neq r}}^k d(\pi_v^{-1}v_i)f_i^2 \right)^2 \\ \leq & \left( \sum_{i=1}^k d(v_i)f_i^2 + \sum_{\substack{i=1 \\ i \neq r}}^k d(\pi_v^{-1}v_i)(f_i)^2 \right)^2 - \left( \sum_{\substack{i=1 \\ i \neq r}}^k d(\pi_v^{-1}v_i)f_i^2 \right)^2 \\ \leq & \left( \sum_{i=1}^k d(v_i)f_i^2 + \sum_{\substack{i=1 \\ i \neq r}}^k d(\pi_v^{-1}v_i)f_i^2 \right)^2, \end{aligned}$$

where (A.12) has been used to obtain the inequality on the penultimate line.  $\square$

Observe that if the nonnegative vector  $\mathbf{f}$  in the statement of Lemma A.2.3, is such that  $f_i$  is decreasing for each  $i \leq r$ , for some  $1 < r \leq k$ . Then given that the vertex  $v_r$  satisfies  $d(\pi_v^{-1}v_r) \leq \sum_{i=1}^{r-1} d(\pi_v^{-1}v_i)$ , we have the following analogue for (A.12)

$$d(\pi_v^{-1}r)f_r^2 \leq \sum_{i=1}^{r-1} d(\pi_v^{-1}v_i)f_i^2 \leq \sum_{\substack{i=1 \\ i \neq r}}^k d(\pi_v^{-1}v_i)f_i^2.$$

Consequently, by performing exactly the same calculations as in Lemma A.2.3, we obtain

**Corollary A.2.4.** *Let  $G = (V, E)$  be a simple, connected graph with  $|V| = k$ , and  $\pi_v : V \rightarrow V$  be a graph isomorphism. Let  $\mathbf{f} \in \mathbb{R}^k$  be a nonnegative, such that*

$f_i$  is decreasing for each  $i \leq r$ , for some  $1 < r \leq k$ . If the vertex  $v_r$  satisfies  $d(\pi_v^{-1}v_r) \leq \sum_{i=1}^{r-1} d(\pi_v^{-1}v_i)$ , then

$$\left( \sum_{i=1}^k d(v_i) f_i^2 \right) \left( \sum_{i=1}^k d(v_i) f_i^2 + d(\pi_v^{-1}v_i) f_i^2 \right) \leq \left( \sum_{i=1}^k d(v_i) f_i^2 + \sum_{\substack{i=1 \\ i \neq r}}^k d(\pi_v^{-1}v_i) f_i^2 \right)^2.$$

**Lemma A.2.5.** Let  $G(V, E)$  be a simple, connected graph with  $|V| = k$ , and  $\pi = (\pi_v, \pi_e)$  a graph isomorphism. Let  $\pi_p$  be the vertex label permutation associated with the vertex permutation  $\pi_v$ , and  $R : \mathbb{R}^k \rightarrow \mathbb{R}$  be defined as in (A.3). One has for all  $\mathbf{f} \in \mathbb{R}^k$

$$2\sqrt{R(\mathbf{f})} \geq \frac{\sum_{i \sim j} |f_i^2 - f_j^2| + |f_{\pi_p(i)}^2 - f_{\pi_p(j)}^2|}{\sqrt{\left( \sum_{i=1}^k d(v_i) f_i^2 \right) \left( \sum_{i=1}^k d(v_i) f_i^2 + d(v_i) f_{\pi_p(i)}^2 \right)}}. \quad (\text{A.13})$$

*Proof.* For each  $1 \leq i \leq k$ ,

$$\begin{aligned} \sum_{i \sim j} (f_i + f_j)^2 &\leq \sum_{i \sim j} (f_i + f_j)^2 + (f_i - f_j)^2 \\ &= 2 \sum_{i \sim j} f_i^2 + f_j^2 \\ &= 2 \sum_{i=1}^k d(v_i) f_i^2. \end{aligned} \quad (\text{A.14})$$

Similarly

$$\sum_{i \sim j} (f_{\pi_p(i)} + f_{\pi_p(j)})^2 \leq 2 \sum_{i=1}^k d(v_i) f_{\pi_p(i)}^2. \quad (\text{A.15})$$

By adding (A.14) to (A.15) then rearranging, we arrive at the inequality

$$\frac{\sum_{i \sim j} (f_i + f_j)^2 + (f_{\pi_p(i)} + f_{\pi_p(j)})^2}{2 \left( \sum_{i=1}^k d(v_i) f_i^2 + d(v_i) f_{\pi_p(i)}^2 \right)} \leq 1. \quad (\text{A.16})$$

Therefore

$$\begin{aligned} R(\mathbf{f}) &= \frac{\sum_{i \sim j} (f_i - f_j)^2 + (f_{\pi_p(i)} - f_{\pi_p(j)})^2}{2 \sum_{i=1}^k d(v_i) f_i^2} \\ &\geq \frac{\sum_{i \sim j} (f_i - f_j)^2 + (f_{\pi_p(i)} - f_{\pi_p(j)})^2}{2 \sum_{i=1}^k d(v_i) f_i^2} \times \text{LHS of (A.16)} \\ &= \frac{(a^2 + b^2)(\hat{a}^2 + \hat{b}^2)}{4 \left( \sum_{i=1}^k d(v_i) f_i^2 \right) \left( \sum_{i=1}^k d(v_i) f_i^2 + d(v_i) f_{\pi_p(i)}^2 \right)} \end{aligned} \quad (\text{A.17})$$

where

$$\begin{aligned} a &= \left( \sum_{i \sim j} (f_i - f_j)^2 \right)^{\frac{1}{2}} & b &= \left( \sum_{i \sim j} (f_{\pi_p(i)} - f_{\pi_p(j)})^2 \right)^{\frac{1}{2}} \\ \hat{a} &= \left( \sum_{i \sim j} (f_i + f_j)^2 \right)^{\frac{1}{2}} & \hat{b} &= \left( \sum_{i \sim j} (f_{\pi_p(i)} + f_{\pi_p(j)})^2 \right)^{\frac{1}{2}}. \end{aligned}$$

Observe that each of the term  $a$ ,  $b$ ,  $\hat{a}$  and  $\hat{b}$  are positive and real, so that

$$(a^2 + b^2)(\hat{a}^2 + \hat{b}^2) \geq (a\hat{a} + b\hat{b})^2.$$

Furthermore, application of the Cauchy-Schwartz inequality on the expressions  $a\hat{a}$  and  $b\hat{b}$  yields

$$a\hat{a} = \left( \sum_{i \sim j} (f_i - f_j)^2 \right)^{\frac{1}{2}} \left( \sum_{i \sim j} (f_i + f_j)^2 \right)^{\frac{1}{2}} \geq \sum_{i \sim j} |f_i^2 - f_j^2|$$

and

$$b\hat{b} = \left( \sum_{i \sim j} (f_{\pi_p(i)} - f_{\pi_p(j)})^2 \right)^{\frac{1}{2}} \left( \sum_{i \sim j} (f_{\pi_p(i)} + f_{\pi_p(j)})^2 \right)^{\frac{1}{2}} \geq \sum_{i \sim j} |f_{\pi_p(i)}^2 - f_{\pi_p(j)}^2|.$$

Thus,

$$R(\mathbf{f}) \geq (\text{A.17}) \geq \frac{\left( \sum_{i \sim j} |f_i^2 - f_j^2| + |f_{\pi_p(i)}^2 - f_{\pi_p(j)}^2| \right)^2}{4 \left( \sum_{i=1}^k d(v_i) f_i^2 \right) \left( \sum_{i=1}^k d(v_i) f_i^2 + d(v_i) f_{\pi_p(i)}^2 \right)}, \quad (\text{A.18})$$

by taking square root on both sides of (A.18), we arrive at the required result.  $\square$

**The proof of Theorem 2.2.3(2):** Assume  $G(V, E)$  is simple and connected with  $|V| = k$ . Let  $\mathbf{g}_2 = \mathbf{D}^{1/2} \mathbf{f}$  be the eigenvector of  $\mathcal{L}^{dyn}$  corresponding to  $\lambda_2$  and  $\mathbf{1} \in \mathbb{R}^k$  be the vector with each element equal to 1. We order the vertices of  $G$  according to  $\mathbf{f}$  by  $f_i \leq f_{i+1}$ , and let  $r$  and  $q$  denote the largest integers such that  $D(S^{r-1}) < D(V)/2$ , and  $D(\pi_v^{-1}(S^q)) < D(V)/2$ . We first consider the situation of  $r \leq q$ , and define the positive and negative parts of  $f_i$  by

$$f_i^- = \begin{cases} |f_i - f_r| & \text{if } f_i < f_r \\ 0 & \text{otherwise} \end{cases}$$

and

$$f_i^+ = \begin{cases} f_i - f_r & \text{if } f_i \geq f_r \\ 0 & \text{otherwise} \end{cases};$$

i.e.  $f_i = f_i^+ - f_i^- + f_r$ , then form the vector  $\mathbf{f}^+$  and  $\mathbf{f}^-$ . First we show that  $R(\mathbf{f}) \geq \min\{R(\mathbf{f}^+), R(\mathbf{f}^-)\}$ . Due to Theorem 2.2.2,  $\sum_{i=1}^k d(v_i)f_i = \langle \mathbf{f}, \mathbf{D}\mathbf{1} \rangle = 0$ .

Thus

$$\begin{aligned}
 & \sum_{i=1}^k d(v_i)f_i^2 & (A.19) \\
 = & \sum_{i=1}^{r-1} d(v_i)f_i^2 + \sum_{i=r}^k d(v_i)f_i^2 \\
 \leq & \sum_{i=1}^{r-1} d(v_i)(f_i^2 + f_r^2) + \sum_{i=r}^k d(v_i)(f_i^2 + f_r^2) \\
 = & \sum_{i=1}^{r-1} d(v_i)(f_i^2 + f_r^2) - 2f_r \sum_{i=1}^{r-1} d(v_i)f_i - 2f_r \sum_{i=r}^k d(v_i)f_i + \sum_{i=r}^k d(v_i)(f_i^2 + f_r^2) \\
 = & \sum_{i=1}^k d(v_i)(f_i^+)^2 + d(v_i)(f_i^-)^2. & (A.20)
 \end{aligned}$$

Furthermore, by the definitions of  $f_i^-$  and  $f_i^+$ ,  $f_i^+ f_j^- = 0$  and  $f_i - f_j = (f_i^+ - f_j^+) + (f_i^- - f_j^-)$ . Therefore

$$\begin{aligned}
 (f_i - f_j)^2 &= ((f_i^+ - f_j^+) + (f_i^- - f_j^-))^2 \\
 &= (f_i^+ - f_j^+)^2 + (f_i^+ f_i^- - f_i^+ f_j^- - f_i^- f_j^+ + f_j^+ f_j^-) + (f_i^- - f_j^-)^2 \\
 &= (f_i^+ - f_j^+)^2 + (f_i^- - f_j^-)^2. & (A.21)
 \end{aligned}$$

Note that  $i, j$  can be replaced with  $\pi_p(i), \pi_p(j)$  in (A.21) to obtain analogous results for  $(f_{\pi_p(i)} - f_{\pi_p(j)})^2$ . Applying (A.21) and (A.19) to the numerator and denominator of  $R(\mathbf{f})$  respectively, then

$$\begin{aligned}
 R(\mathbf{f}) &= \frac{\sum_{i \sim j} (f_i - f_j)^2 + (f_{\pi_p(i)} - f_{\pi_p(j)})^2}{2 \sum_{i=1}^k d(v_i)f_i^2} \\
 &\geq \frac{\sum_{i \sim j} (f_i^+ - f_j^+)^2 + (f_{\pi_p(i)}^+ - f_{\pi_p(j)}^+)^2 + (f_i^- - f_j^-)^2 + (f_{\pi_p(i)}^- - f_{\pi_p(j)}^-)^2}{2 \sum_{i=1}^k d(v_i)(f_i^+)^2 + d(v_i)(f_i^-)^2} \\
 &\geq \min\{R(\mathbf{f}^+), R(\mathbf{f}^-)\} & (A.22)
 \end{aligned}$$

where we have used the fact that for any positive real numbers  $a, b, c$  and  $d$

$$\frac{a+b}{c+d} \geq \min\left\{\frac{a}{c}, \frac{b}{d}\right\}.$$

Now by Theorem 2.2.2(4), one has  $\lambda_2 = -R(\mathbf{f})$ . So if  $\min\{R(\mathbf{f}^+), R(\mathbf{f}^-)\} = R(\mathbf{f}^+)$ , then (A.22) becomes  $\lambda_2 \geq -R(\mathbf{f}^+)$ . Thus, by Lemma A.2.5

$$2\sqrt{-\lambda_2} \geq \frac{\sum_{i \sim j} |(f_i^+)^2 - (f_j^+)^2| + |(f_{\pi_p(i)}^+)^2 - (f_{\pi_p(j)}^+)^2|}{\sqrt{\left(\sum_{i=1}^k d(v_i)(f_i^+)^2\right) \left(\sum_{i=1}^k d(v_i)(f_i^+)^2 + d(v_i)(f_{\pi_p(i)}^+)^2\right)}}, \quad (A.23)$$

In the following, we consider the expression on the numerator of (A.23). By the application of Lemma A.2.1 and Corollary A.2.2

$$\begin{aligned} & \sum_{i \sim j} |(f_i^+)^2 - (f_j^+)^2| + |(f_{\pi_p(i)}^+)^2 - (f_{\pi_p(j)}^+)^2| \\ &= \sum_{i=1}^{k-1} |(f_{i+1}^+)^2 - (f_i^+)^2| \cdot (|C(S^i)| + |C(\pi_v^{-1}(S^i))|). \end{aligned} \quad (\text{A.24})$$

On substituting the ordered cut value

$$\alpha = \min_{1 \leq i \leq k} \frac{|C(S^i)| + |C(\pi_v^{-1}(S^i))|}{\hat{D}(S^i) + \hat{D}(\pi_v^{-1}(S^i))}. \quad (\text{A.25})$$

into (A.24), where  $\hat{D}(S) = \min\{D(S), D(\overline{S})\}$  for  $S \subset V$ . Then using the fact that  $\hat{D}(S^k) = \hat{D}(S^0) = \hat{D}(\pi_v^{-1}(S^k)) = \hat{D}(\pi_v^{-1}(S^0)) = 0$ , we have

$$\begin{aligned} (\text{A.24}) &\geq \alpha \sum_{i=1}^{k-1} |(f_{i+1}^+)^2 - (f_i^+)^2| \cdot (\hat{D}(S^i) + \hat{D}(\pi_v^{-1}(S^i))) \\ &= \alpha \sum_{i=1}^k (f_i^+)^2 \cdot (|\hat{D}(S^{i-1}) - \hat{D}(S^i)| + |\hat{D}(\pi_v^{-1}(S^{i-1})) - \hat{D}(\pi_v^{-1}(S^i))|). \end{aligned} \quad (\text{A.26})$$

Observe by the definition of  $f_i^+$  and the assumption that  $f_i \leq f_{i+1}$ , one has  $f_i^+ = 0$  for  $i \leq r$ . Therefore, we only need to consider the summation of (A.26) over the indices  $r < i$ . Furthermore, for each  $i > r$ ,  $\hat{D}(S^{i-1}) = D(\overline{S^{i-1}})$  because  $r$  is the largest integer such that  $D(S^{r-1}) < D(V)/2$ . So that for  $i > r$

$$\hat{D}(S^{i-1}) - \hat{D}(S^i) = \sum_{j \geq i} d(v_j) - \sum_{j \geq i+1} d(v_j) = d(v_i). \quad (\text{A.27})$$

Now for the expression  $\hat{D}(\pi_v^{-1}(S^{i-1})) - \hat{D}(\pi_v^{-1}(S^i))$ , when  $i < q$ ,  $D(\pi_v^{-1}(S^{i-1})) \leq D(\pi_v^{-1}(S^i)) < D(V)/2$ . So that for  $i < q$

$$\hat{D}(\pi_v^{-1}(S^{i-1})) - \hat{D}(\pi_v^{-1}(S^i)) = \sum_{j \leq i-1} d(\pi_v^{-1}v_j) - \sum_{j \leq i} d(\pi_v^{-1}v_j) = -d(\pi_v^{-1}v_i). \quad (\text{A.28})$$

In addition, when  $i > q$ ,  $D(\pi_v^{-1}(S^i)) \geq D(\pi_v^{-1}(S^{i-1})) > D(V)/2$ . So that for  $i > q$

$$\hat{D}(\pi_v^{-1}(S^{i-1})) - \hat{D}(\pi_v^{-1}(S^i)) = \sum_{j \geq i} d(\pi_v^{-1}v_j) - \sum_{j \geq i+1} d(\pi_v^{-1}v_j) = d(\pi_v^{-1}v_i). \quad (\text{A.29})$$

Combining (A.28) and (A.29), we conclude that  $|\hat{D}(\pi_v^{-1}(S^{i-1})) - \hat{D}(\pi_v^{-1}(S^i))| =$

$d(\pi_v^{-1}v_i)$  when  $i \neq q$ . Therefore,

$$\begin{aligned}
 & \sum_{r < i \leq k} (f_i^+)^2 |\hat{D}(\pi_v^{-1}(S^{i-1})) - \hat{D}(\pi_v^{-1}(S^i))| \\
 &= \left( \sum_{\substack{r < i \leq k \\ i \neq q}} (f_i^+)^2 d(\pi_v^{-1}v_i) \right) + (f_q^+)^2 |\hat{D}(\pi_v^{-1}(S^q)) - \hat{D}(\pi_v^{-1}(S^q))| \\
 &\geq \sum_{\substack{r < i \leq k \\ i \neq q}} (f_i^+)^2 d(\pi_v^{-1}v_i). \tag{A.30}
 \end{aligned}$$

Inserting equation (A.30) and (A.27) in (A.26) gives

$$\text{(A.24)} \geq \alpha \left( \sum_{r < i \leq k} (f_i^+)^2 d(v_i) + \sum_{\substack{r < i \leq k \\ i \neq q}} (f_i^+)^2 d(\pi_v^{-1}v_i) \right). \tag{A.31}$$

Next we treat the denominator of (A.23). Recall the theorem hypothesis  $d(\pi_v^{-1}v_q) \leq D(V)/4$ , this implies  $d(\pi_v^{-1}v_q) \leq \sum_{i=q+1}^k d(\pi_v^{-1}v_i)$ , because if otherwise

$$\begin{aligned}
 D(\pi_v^{-1}(S^q)) &= D(V) - D(\overline{\pi_v^{-1}(S^q)}) \\
 &= D(V) - d(\pi_v^{-1}v_q) - \sum_{i=q+1}^k d(\pi_v^{-1}v_i) \\
 &\geq D(V) - 2d(\pi_v^{-1}v_q) \geq D(V)/2,
 \end{aligned}$$

which contradicts the fact that  $q$  is the largest index satisfying  $D(\pi_v^{-1}(S^q)) < D(V)/2$ . Furthermore,  $f_i^+$  is positive and increasing for  $i \geq q$ , and  $q \neq k$  because  $D(\pi_v^{-1}(S^{k-1})) \geq D(V)/2$ . Hence by Lemma A.2.3

$$\begin{aligned}
 & \sqrt{\left( \sum_{i=1}^k d(v_i)(f_i^+)^2 \right) \left( \sum_{i=1}^k d(v_i)(f_i^+)^2 + d(v_i)(f_{\pi_p(i)}^+)^2 \right)} \\
 &= \sqrt{\left( \sum_{i=1}^k d(v_i)(f_i^+)^2 \right) \left( \sum_{i=1}^k d(v_i)(f_i^+)^2 + d(\pi_p^{-1}v_i)(f_i^+)^2 \right)} \\
 &\leq \sum_{i=1}^k d(v_i)(f_i^+)^2 + \sum_{\substack{i=1 \\ i \neq q}}^k d(\pi_v^{-1}v_i)(f_i^+)^2 \\
 &= \sum_{r < i \leq k} d(v_i)(f_i^+)^2 + \sum_{\substack{r < i \leq k \\ i \neq q}} d(\pi_v^{-1}v_i)(f_i^+)^2. \tag{A.32}
 \end{aligned}$$

By substituting (A.31) and (A.32) into the numerator and denominator of (A.23)

respectively, one has

$$2\sqrt{-\lambda_2} \geq \alpha \left( \frac{\sum_{r < i \leq k} d(v_i)(f_i^+)^2 + \sum_{\substack{r < i \leq k \\ i \neq q}} d(\pi_v^{-1}v_i)(f_i^+)^2}{\sum_{r < i \leq k} d(v_i)(f_i^+)^2 + \sum_{\substack{r < i \leq k \\ i \neq q}} d(\pi_v^{-1}v_i)(f_i^+)^2} \right) = \alpha. \quad (\text{A.33})$$

By comparing the definition of the ordered cut value  $\alpha$  (A.25) to the constant  $\mathbf{H}_G^{\text{dyn}}$  given by (2.1), we immediately see that  $\alpha \geq \mathbf{H}_G^{\text{dyn}}$ , so the required inequality follows from (A.33). If instead  $\min\{R(\mathbf{f}^+), R(\mathbf{f}^-)\} = R(\mathbf{f}^-)$ , then one can obtain (A.33) by following similar arguments from equation (A.23) and onwards, with the following alterations:

1. Replace  $\mathbf{f}^+$  with  $\mathbf{f}^-$ ,
2. The summation in (A.26) only needs to be considered over the indices  $i > r$ ,
3. (A.27) is replaced by

$$\hat{D}(S^{i-1}) - \hat{D}(S^i) = \sum_{j \leq i-1} d(v_j) - \sum_{j \leq i} d(v_j) = -d(v_i), \quad (\text{A.34})$$

while (A.28) and (A.29) are replaced by

$$\hat{D}(\pi_v^{-1}(S^{i-1})) - \hat{D}(\pi_v^{-1}(S^i)) = \sum_{j \leq i-1} d(\pi_v^{-1}v_j) - \sum_{j \leq i} d(\pi_v^{-1}v_j) = -d(\pi_v^{-1}v_i). \quad (\text{A.35})$$

To complete the proof, suppose we have the other situation of  $q < r$ . Then one defines the positive and negative parts of  $f_i$  as before, and follow from equation (A.19)-(A.22) to obtain  $R(\mathbf{f}) \geq \min\{R(\mathbf{f}^+), R(\mathbf{f}^-)\}$ . Now if  $\min\{R(\mathbf{f}^+), R(\mathbf{f}^-)\} = R(\mathbf{f}^-)$ , then by making the appropriate alterations from (A.23)-(A.31), one has

$$2\sqrt{-\lambda_2} \geq \alpha \left( \frac{\sum_{1 \leq i < r} (f_i^-)^2 d(v_i) + \sum_{\substack{1 \leq i < r \\ i \neq q}} (f_i^-)^2 d(\pi_v^{-1}v_i)}{\sqrt{\left(\sum_{i=1}^k d(v_i)(f_i^-)^2\right) \left(\sum_{i=1}^k d(v_i)(f_i^-)^2 + d(v_i)(f_{\pi_v^{-1}(i)}^-)^2\right)}} \right). \quad (\text{A.36})$$

To treat the denominator of (A.36), note that the theorem hypothesis  $d(\pi_v^{-1}k) \leq D(V)/4$  also implies  $d(\pi_v^{-1}v_q) \leq \sum_{i=1}^{q-1} d(\pi_v^{-1}v_i)$  as before. Furthermore,  $f_i^-$  is positive and decreasing for  $i \leq q$ , and  $q \neq 1$  because  $D(\pi_v^{-1}(S^1)) \leq D(V)/4 < D(V)/2$  (contradicts  $q-1$  being the largest index satisfying this condition). Thus, by Corollary A.2.4

$$\begin{aligned} & \sqrt{\left(\sum_{i=1}^k d(v_i)(f_i^-)^2\right) \left(\sum_{i=1}^k d(v_i)(f_i^-)^2 + d(v_i)(f_{\pi_v^{-1}(i)}^-)^2\right)} \\ & \leq \sum_{1 \leq i < r} d(v_i)(f_i^-)^2 + \sum_{\substack{1 \leq i < r \\ i \neq q}} d(\pi_v^{-1}v_i)(f_i^-)^2. \end{aligned} \quad (\text{A.37})$$

Hence, by (A.36) and (A.37) we have  $2\sqrt{-\lambda_2} \geq \alpha$  and the result follows. The case of  $\min\{R(\mathbf{f}^+), R(\mathbf{f}^-)\} = R(\mathbf{f}^+)$  is a straight forward modification to the  $r \leq q$  case, thus the details are omitted.

# Appendix B

Let  $(M, m)$  and  $(N, n)$  be compact, connected  $r$ -dimensional  $C^\infty$  Riemannian manifolds, where  $m, n$  are the Riemannian metric tensors on  $M$  and  $N$  respectively. Let  $T$  be a  $C^\infty$ -diffeomorphism of  $M$  onto  $N$ , and  $V_m$  the volume measure on  $M$ . Let  $(M, m, \mu_r)$  and  $(N, n, \nu_r)$  be weighted Riemannian manifolds, where  $\mu_r$  is absolutely continuous probability measures with respect to  $V_m$ , and  $\nu_r = \mu_r \circ T^{-1}$ . Denote by  $h_\mu$  and  $h_\nu$  the densities of the measures  $\mu_r$  and  $\nu_r$  respectively.

## B.1 Muckenhoupt weights $A_p$

For a measurable function  $f$  on  $M$ , the *essential supremum* of  $f$  is the number

$$\text{ess sup } f := \{a \in \mathbb{R} : V_m(f^{-1}(a, \infty)) = 0\}.$$

Recall from Section 3.1 that the volume form  $\omega_m^r$  on  $M$  is given in terms of the volume measure  $V_m$  on  $M$  via  $V_m(U) = \int_U \omega_m^r$  for any measurable subset  $U \subseteq M$ . Define the class of  $A_p$  weights [119]:

**Definition B.1.1.** Let  $B_\rho(x) \subset M$  denote the metric ball centered at  $x \in M$  with radius  $\rho > 0$ . The density  $h_\mu$  of the measure  $\mu_r$  is said to be an  $A_p$  weight of  $(M, m, \mu_r)$ , if there exists a constant  $C_\mu$  such that for every  $x$  and  $\rho$ ,  $h_\mu$  satisfies

$$\left( \frac{1}{V_m(B_\rho(x))} \int_{B_\rho(x)} h_\mu \cdot \omega_m^r \right) \left( \frac{1}{V_m(B_\rho(x))} \int_{B_\rho(x)} h_\mu^{-\frac{1}{p-1}} \cdot \omega_m^r \right)^{p-1} \leq C_\mu, \quad (\text{B.1})$$

for  $1 < p < \infty$ , or

$$\left( \frac{1}{V_m(B_\rho(x))} \int_{B_\rho(x)} h_\mu \cdot \omega_m^r \right) \left( \text{ess sup}_{z \in B_\rho(x)} \frac{1}{h_\mu(z)} \right) \leq C_\mu, \quad (\text{B.2})$$

for  $p = 1$ . We call  $C_\mu$  the  $A_p$  constant of  $h_\mu$ .

**Proposition B.1.2.** *Suppose the density  $h_\mu$  of the measure  $\mu_r$  is Lipschitz and uniformly bounded away from zero. Then  $h_\mu$  is an  $A_p$  weight for all  $1 \leq p < \infty$ .*

*Proof.* Let  $B_\rho(x) \subset M$ , and denote by  $\text{dist}_m$  be the Riemannian distance function with respect to the metric  $m$ . Since  $h_\mu$  is Lipschitz and nonnegative, for every  $x, z \in M$  one has  $h_\mu(z) \leq h_\mu(x) + K \text{dist}_m(x, z)$  for some  $K < \infty$ . Hence, for every  $x \in M$  and  $\rho > 0$

$$\begin{aligned} \frac{1}{V_m(B_\rho(x))} \int_{B_\rho(x)} h_\mu(z) \cdot \omega_m^r(z) &\leq \frac{h_\mu(x) + K\rho}{V_m(B_\rho(x))} \int_{B_\rho(x)} \omega_m^r(z) \\ &= h_\mu(x) + K\rho. \end{aligned} \quad (\text{B.3})$$

Since  $M$  is compact and  $B_\rho(x) \subset M$ , one has  $\rho < \infty$ . Also, since  $h_\mu$  is Lipschitz, it is bounded on  $M$ . Hence, the RHS of (B.3) is bounded above by  $\sup_{x \in M} (h_\mu(x) + K\rho)$ .

In addition, since  $h_\mu$  is uniformly bounded away from zero,  $h_\mu^{-1}$  and  $h_\mu^{-1/(p-1)}$  are bounded for  $1 < p < \infty$ . Hence, there exist constants  $\gamma_p$  and  $\gamma_1$  such that

$$\frac{1}{V_m(B_\rho(x))} \int_{B_\rho(x)} h_\mu^{-\frac{1}{p-1}} \cdot \omega_m^r \leq \gamma_p, \quad (\text{B.4})$$

$1 < p < \infty$ , and

$$\left( \text{ess sup}_{z \in B_\rho(x)} \frac{1}{h_\mu(z)} \right) \leq \gamma_1, \quad (\text{B.5})$$

Hence, by (B.3), (B.4) there are constants  $C_\mu = \gamma_p^{1/(p-1)} \cdot \sup_{x \in M} (h_\mu(x) + K\rho)$ , such that (B.1) is satisfied for  $1 < p < \infty$ . Similarly, there is a constant  $C_\mu = \gamma_1 \cdot \sup_{x \in M} (K\rho + h_\mu(x))$ , such that (B.2) is satisfied for  $p = 1$ . Hence,  $h_\mu$  is an  $A_p$  weight.  $\square$

Let  $L_{\text{loc}}^1(M, V_m)$  denote the space of locally integrable functions; that is, if  $f \in L_{\text{loc}}^1(M, V_m)$  then  $\int_{B_\rho(x)} f \cdot \omega_m^r < \infty$  for every  $x \in M$  and  $\rho \in \mathbb{R}^+$ . Given a weighted Riemannian manifold  $(M, m, \mu_r)$ , we wish to determine the condition on the density  $h_\mu$  so that  $L^p(M, \mu_r) \subset L_{\text{loc}}^1(M, V_m)$ .

**Proposition B.1.3.** *Let  $B_\rho(x) \subset M$  denote the metric ball centered at  $x \in M$  with radius  $\rho > 0$ . If  $h_\mu^{-1/(p-1)}$  is in  $L_{\text{loc}}^1(M, V_m)$  for  $p \in (0, \infty)$ , or if for every  $x$  and  $\rho$*

$$\text{ess sup}_{z \in B_\rho(x)} \frac{1}{h_\mu(z)} < \infty,$$

*for  $p = 1$ . Then  $L^p(M, m, \mu_r) \subset L_{\text{loc}}^1(M, V_m)$ .*

*Proof.* This result appeared in [119] for the case  $M = \mathbb{R}^r$ ; the arguments for the present version is identical, thus are omitted.  $\square$

## B.2 Additional notes on differential geometry

By a local chart on  $M$ , we meant a pair  $(U, \varphi)$  such that  $U$  is an open subset of  $M$  and  $\varphi : U \rightarrow \mathbb{R}^r$  a local  $C^\infty$ -diffeomorphism. The countable collection of local charts  $(U_i, \varphi_i)_{i \in I}$  such that  $\cup_{i \in I} U_i$  forms an open cover for  $M$  is called an atlas. For a fixed  $k \in I$ , one can define a set of local coordinates  $(x_1, x_2, \dots, x_r)$  on  $U_k$  as the set of smooth projections of the image of  $\varphi_k$  onto the  $j^{\text{th}}$  coordinate,  $1 \leq j \leq r$ ; that is  $x_j : U_k \rightarrow \mathbb{R}$  is a homeomorphism for each  $1 \leq j \leq r$ . Moreover, the atlas on  $M$  defines a local coordinate system for each point  $x \in M$ . In local coordinates, it is possible to carry out the operation of partial differentiation on a differentiable function  $f$  at the point  $x \in U_k$ ,  $k \in I$  as

$$\left[ \frac{\partial}{\partial x_j} \right]_x f := \frac{\partial(f \circ \varphi_k^{-1})}{\partial x_j}(\varphi_k(x)), \quad (\text{B.6})$$

for each  $1 \leq j \leq r$ . It is well known (see e.g p.7 in [36]) that the above operation is independent on the choice of  $\varphi_k$ , and therefore we use the abbreviation  $[\partial/\partial x_j]_x f = \partial f/\partial x_j(x) = \partial_j f(x)$  whenever there is no confusion on whether the partial differential is carried out on  $\mathbb{R}^r$  or  $M$ . It is straightforward to verify that the set  $\{\partial_i\}_{i=1}^r$  forms a basis for the vector fields on  $M$ . Hence, one can write the metric tensor  $m$  in coordinates as  $m_{ij}(x) = m(\partial_i, \partial_j)(x)$ .

Given a diffeomorphism  $T : M \rightarrow N$ , and local charts  $(U, \varphi)$ ,  $(TU, \vartheta)$  on  $M$ ,  $N$  respectively. Observe that  $\vartheta \circ T : M \rightarrow \mathbb{R}^r$  is smooth. Therefore, it is possible to carry out the operation of partial differentiation on  $T$  at the point  $x \in U$  as

$$\left[ \frac{\partial}{\partial x_j} \right]_x T := \frac{\partial(\vartheta \circ T \circ \varphi^{-1})}{\partial x_j}(\varphi(x)). \quad (\text{B.7})$$

One can construct the Jacobian matrix  $J_T$  in local coordinates via (B.7), as a  $r \times r$  matrix with entries  $(J_T)_{ij} := \partial_j T_i$ , where  $T_i$  is the smooth projection of the image of  $\vartheta \circ T$  onto the  $i^{\text{th}}$  coordinate, and the abbreviation  $\partial_j T_i(x) = \partial T_i/\partial x_j(x)$  had been applied.

### B.2.1 Differential forms

Let  $(x_1, x_2, \dots, x_r)$  be local coordinates on  $M$ . Denote by  $dx_i$  the differential 1-forms dual to the tangent basis  $\partial_i$ , for each  $1 \leq i \leq r$ . For  $p \leq r$ , one can express a differentiable  $p$ -form  $\eta$  in coordinates via the exterior product of 1-forms

$$\eta = \sum_{j_1 < j_2 < \dots < j_p} a_{j_1 \dots j_p} dx_{j_1} \wedge dx_{j_2} \wedge \dots \wedge dx_{j_p}, \quad (\text{B.8})$$

where  $a_{j_1 \dots j_p}$  are real-valued functions on  $M$ .

The exterior derivative on a differentiable  $f : M \rightarrow \mathbb{R}$  is a 1-form given by  $df = \sum_{i=1}^r \partial_i f dx_i$ , and the exterior derivative on the  $p$ -form  $\eta$  defined by (B.8) is a  $(p+1)$ -form satisfying

$$d\eta = \sum_{j_1 < j_2 < \dots < j_p} d(a_{j_1 \dots j_p}) \wedge dx_{j_1} \wedge dx_{j_2} \wedge \dots \wedge dx_{j_p}.$$

The interior derivative  $i(\mathcal{V})$  on a  $p$ -form  $\eta$  with respect to a vector field  $\mathcal{V}$  on  $M$ , is a  $(p-1)$ -form satisfying

$$[i(\mathcal{V})\eta](\mathcal{V}_1, \mathcal{V}_2, \dots, \mathcal{V}_{p-1}) = \eta(\mathcal{V}, \mathcal{V}_1, \mathcal{V}_2, \dots, \mathcal{V}_{p-1}),$$

for all vector fields  $\mathcal{V}_1, \mathcal{V}_2, \dots, \mathcal{V}_{p-1}$  on  $M$ .

Recall the definition of the tangent and cotangent mappings  $T_*$  and  $T^*$  associated with  $T$ , given by (3.5) and (3.6) respectively. For the differential  $p$ -form  $\eta$  given by (B.8), one has

$$(T^*\eta)(\mathcal{V}_1, \mathcal{V}_2, \dots, \mathcal{V}_p)(x) = \eta(T_*\mathcal{V}_1, T_*\mathcal{V}_2, \dots, T_*\mathcal{V}_p)(Tx),$$

for all vector fields  $\mathcal{V}_1, \mathcal{V}_2, \dots, \mathcal{V}_p$  on  $M$ . Therefore, in coordinates

$$\begin{aligned} T^*\eta &= T^* \left( \sum_{j_1 < j_2 < \dots < j_p} a_{j_1 \dots j_p} dx_{j_1} \wedge dx_{j_2} \wedge \dots \wedge dx_{j_p} \right) \\ &= \sum_{j_1 < j_2 < \dots < j_p} a_{j_1 \dots j_p} \circ T \cdot T^* dx_{j_1} \wedge T^* dx_{j_2} \wedge \dots \wedge T^* dx_{j_p} \\ &= \sum_{j_1 < j_2 < \dots < j_p} a_{j_1 \dots j_p} \circ T \cdot d(x_{j_1} \circ T) \wedge d(x_{j_2} \circ T) \wedge \dots \wedge d(x_{j_p} \circ T), \end{aligned} \tag{B.9}$$

where the last line is due to the fact that  $[T^*(df)]\mathcal{V} = \mathcal{V}(f \circ T) = [d(f \circ T)]\mathcal{V}$ , for all  $f \in C^\infty(M, \mathbb{R})$  and vector fields  $\mathcal{V}$  on  $M$ .

Let  $G_m(x)$  be a  $r \times r$  matrix with components  $m_{ij}(x)$  at the point  $x \in M$ . The volume form  $\omega_m^r$  in the local coordinates  $\{x_i\}_{i=1}^r$  is defined by

$$\omega_m^r(x) := \sqrt{\det G_m(x)} \cdot dx_1 \wedge dx_2 \wedge \dots \wedge dx_r, \tag{B.10}$$

for each point  $x \in M$ . Let  $\Gamma$  be a  $C^\infty$  co-dimension 1 subset of  $M$ . Recall from Section 3.1 that the embedding  $\Phi : \Gamma \rightarrow M$  induces a Riemannian metric on  $\Gamma$  via the pullback metric  $\Phi^*m$ ; that is  $\omega_m^{r-1} = \Phi^*\omega_m^r$ . The following is a classical result in geometric measure theory (see Theorem I.3.1 in [23]):

**Lemma B.2.1** (co-area formula). *Let  $f \in C^1(M, \mathbb{R})$ . For an open, connected  $U \subseteq M$  with compact closure, and any function  $h : M \rightarrow \mathbb{R}^+$  in  $L^1(M, V_m)$ , one has*

$$\int_U |\nabla_m f|_m h \cdot \omega_m^r = \int_{\mathbb{R}} \left( \int_{f^{-1}\{\beta\}} h \cdot \omega_m^{r-1} \right) d\beta, \tag{B.11}$$

where  $|\cdot|_m^2 = m(\cdot, \cdot)$  and  $\nabla_m f$  is the gradient of  $f$  with respect to the metric  $m$ ; defined by (3.16).

The co-area formula connects the spatial integral over the gradient of a function to the co-dimension one measure on the level sets generated by that function. If the density  $h_\mu$  of the absolutely continuous probability measure  $\mu_r$  is a positive function in  $L^1(M, V_m)$ , then one can apply the co-area formula (B.11) with  $h = h_\mu$  to obtain

$$\int_U |\nabla_m f|_m \cdot h_\mu \omega_m^r = \int_U |\nabla_m f|_m h_\mu \cdot \omega_m^r = \int_{\mathbb{R}} \mu_{r-1}(f^{-1}\{\beta\}) d\beta,$$

for all measurable  $U \subseteq M$ .

## B.2.2 Differential operators on weighted manifolds

Recall the definitions of the gradient  $\nabla_m$ , divergence  $\text{div}_m$  and weighted divergence  $\text{div}_\mu$  given by (3.16), (3.24) and (3.25) respectively. One can express  $\nabla_m f$  in local coordinates  $\{x_1, \dots, x_r\}$  on  $M$  as,

$$\nabla_m f = \sum_{i,j=1}^r m^{ij} \partial_i f \partial_j, \quad (\text{B.12})$$

for all  $f \in C^k(M, \mathbb{R})$ , and where  $m^{ij}$  is the components  $r \times r$  matrix  $G_m^{-1}$  (see p.4, equation (22) [22]). As a consequence of Stokes' theorem (see p.124, [113]), the divergence given by (3.24) can be written as,

$$\text{div}_m \mathcal{V} \cdot \omega_m^r = d[i(\mathcal{V})\omega_m^r], \quad (\text{B.13})$$

for all  $\mathcal{V} \in \mathcal{F}^k(M)$ . Since  $\{\partial_i\}_{i=1}^r$  forms a basis for the vector fields on  $M$ , one can express the vector field  $\mathcal{V}$  on  $M$  as  $\mathcal{V} = \sum_{i=1}^r \mathcal{V}^i \partial_i$ . Then (B.13) in local coordinates is (see equation (32) on p.5 in [22]),

$$\text{div}_m \mathcal{V} = \frac{1}{\sqrt{\det G_m}} \sum_{i=1}^r \partial_i \left( \sqrt{\det G_m} \mathcal{V}^i \right). \quad (\text{B.14})$$

Hence, the Laplace-Beltrami operator is given in local coordinates by

$$\Delta_m f = \frac{1}{\sqrt{\det G_m}} \sum_{i,j=1}^r \partial_i \left( m^{ij} \sqrt{\det G_m} \partial_j f \right). \quad (\text{B.15})$$

Let  $f : M \rightarrow \mathbb{R}$  be differentiable, and  $\mathcal{V}_1, \mathcal{V}_2$  vector fields on  $M$ . The standard divergence properties (see equation (12) and (13) on p.3, [22]) holds analogously for the weighted divergence (3.25); that is, for  $h_\mu \in C^1(M, \mathbb{R})$

$$\begin{aligned} \text{div}_\mu(\mathcal{V}_1 + \mathcal{V}_2) &= \frac{1}{h_\mu} \text{div}(h_\mu \mathcal{V}_1 + h_\mu \mathcal{V}_2) \\ &= \frac{1}{h_\mu} \text{div}(h_\mu \mathcal{V}_1) + \frac{1}{h_\mu} \text{div}(h_\mu \mathcal{V}_2) \\ &= \text{div}_\mu \mathcal{V}_1 + \text{div}_\mu \mathcal{V}_2, \end{aligned} \quad (\text{B.16})$$

and

$$\begin{aligned}
 \operatorname{div}_\mu(f\mathcal{V}_1) &= \frac{1}{h_\mu} \operatorname{div}(h_\mu f \mathcal{V}_1) \\
 &= \frac{f}{h_\mu} \operatorname{div}(h_\mu \mathcal{V}_1) + \frac{1}{h_\mu} m(\nabla_m f, h_\mu \mathcal{V}_1) \\
 &= f \operatorname{div}_\mu \mathcal{V}_1 + m(\nabla_m f, \mathcal{V}_1).
 \end{aligned} \tag{B.17}$$

### B.2.3 Properties of $T_*$ and $T^*$

Let  $g : N \rightarrow \mathbb{R}$  be differentiable. One can express the tangent mapping  $T_*$  given by (3.5) in local coordinates  $\{x_1, \dots, x_r\}$  as

$$(T_*\mathcal{V})g = \mathcal{V}(g \circ T) = \sum_{i=1}^r \mathcal{V}^i \frac{\partial(g \circ T)}{\partial x_i},$$

where  $\mathcal{V} = \sum_{i=1}^r \mathcal{V}^i \partial_i \in \mathcal{T}M$ . The following result computes coordinate representations of the pullback metric  $T^*n$ .

**Lemma B.2.2.** *Let  $n_{ij}$  be the local coordinates representation of the metric tensor  $n$ . Denote by  $G_n$  the  $r \times r$  matrix with components  $n_{ij}$  and  $J_T$  the Jacobian matrix of  $T$ . We have at the each point  $x \in M$*

$$(G_{T^*n})_{ij} = (J_T^\top \cdot G_n \circ T \cdot J_T)_{ij}, \tag{B.18}$$

where  $T^*n$  is the pullback metric of  $n$  given by (3.7).

*Proof.* Let  $(U, \varphi_m)$  be a local chart on  $M$ , containing the point  $x_0 \in U$  with corresponding coordinates  $\{x_i\}_{i=1}^r$ . Then the local chart  $(TU, \varphi_n)$  on  $N$  contains the point  $Tx_0 \in N$ . Let  $\{y_i\}_{i=1}^r$  denote the local coordinates on  $TU$ . Due to (3.5) and (B.6), one has for all differentiable  $g$  on  $N$

$$\begin{aligned}
 \left[ T_* \frac{\partial}{\partial x_i} \right]_{Tx_0} g &= \left[ \frac{\partial}{\partial x_i} \right]_{x_0} (g \circ T) \\
 &= \frac{\partial(g \circ T \circ \varphi_m^{-1})}{\partial x_i}(\varphi_m(x_0)) \\
 &= \frac{\partial(g \circ \varphi_n^{-1} \circ \varphi_n \circ T \circ \varphi_m^{-1})}{\partial x_i}(\varphi_m(x_0)) \\
 &= \sum_{k=1}^r \frac{\partial(\varphi_n \circ T_k \circ \varphi_m^{-1})}{\partial x_i}(\varphi_m(x_0)) \frac{\partial(g \circ \varphi_n^{-1})}{\partial y_k}(\varphi_n(Tx_0)) \\
 &= \sum_{k=1}^r \left[ \frac{\partial}{\partial x_i} \right]_{x_0} T_k \cdot \left[ \frac{\partial}{\partial y_k} \right]_{Tx_0} g,
 \end{aligned}$$

where the last equality is due to (B.7). Therefore,

$$T_* \frac{\partial}{\partial x_i} = \sum_{k=1}^r (J_T)_{ki} \circ T^{-1} \cdot \frac{\partial}{\partial y_k}, \quad (\text{B.19})$$

at the point  $Tx_0$ ; that is  $T_*(\partial/\partial x_i)$  is a tangent vector in  $\mathcal{T}_{Tx_0}(N)$  with components  $(T_*(\partial/\partial x_i))^k = (J_T)_{ki} \circ T^{-1}$ ,  $1 \leq k \leq r$ . To obtain (B.18) at the point  $x_0$ , we compute

$$\begin{aligned} (G_{T^*n})_{ij}(x_0) &= T^*n \left( \frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j} \right) (x_0) \\ &= n \left( T_* \frac{\partial}{\partial x_i}, T_* \frac{\partial}{\partial x_j} \right) (Tx_0) \\ &= n \left( \sum_{k=1}^r \left( T_* \frac{\partial}{\partial x_i} \right)^k \frac{\partial}{\partial y_k}, \sum_{l=1}^r \left( T_* \frac{\partial}{\partial x_j} \right)^l \frac{\partial}{\partial y_l} \right) (Tx_0) \\ &= \sum_{k,l=1}^r \left( n_{kl} \cdot \left( T_* \frac{\partial}{\partial x_i} \right)^k \cdot \left( T_* \frac{\partial}{\partial x_j} \right)^l \right) \Big|_{Tx_0} \\ &= \sum_{k,l=1}^r (J_T(x_0))_{ki} \cdot n_{kl}(Tx_0) \cdot (J_T(x_0))_{lj} \quad \text{by (B.19)} \\ &= \left( J_T^\top \cdot G_n \circ T \cdot J_T \right)_{ij} (x_0). \end{aligned}$$

Since  $x_0 \in U$  is arbitrary and  $(U, \varphi_m)$  is a chart for  $M$ , we conclude that the above calculations hold for all points in  $M$ .  $\square$

**Corollary B.2.3.** *Let  $n$  be the metric tensor of  $N$ , with volume form  $\omega_n^r$  given by (B.10). Define the co-tangent mapping  $T^*$  as in (3.6). One has  $T^*\omega_n^r = \omega_{T^*n}^r$ .*

*Proof.* Let  $\{x_i\}_{i=1}^r$  and  $\{y_i\}_{i=1}^r$  be local coordinates on  $M$  and  $N$  respectively. Then by (B.9), one has for each  $1 \leq i \leq r$ ,

$$T^*(dy_i) = d(y_i \circ T) = \sum_{j=1}^r \frac{\partial(y_i \circ T)}{\partial x_j} dx_j = \sum_{j,k=1}^r \frac{\partial T_k}{\partial x_j} \frac{\partial y_i}{\partial y_k} \circ T \cdot dx_j = \sum_{j,k=1}^r \frac{\partial T_k}{\partial x_j} \delta_{ik} \cdot dx_j,$$

where  $\delta_{ik}$  is the kronecker delta. Therefore,  $T^*(dy_i) = \sum_{j=1}^r \partial T_i / \partial x_j \cdot dx_j$ . It follows that

$$T^*(dy_1 \wedge dy_2 \wedge \dots \wedge dy_r) = T^*dy_1 \wedge T^*dy_2 \wedge \dots \wedge T^*dy_r = |\det J_T| \cdot dx_1 \wedge dx_2 \wedge \dots \wedge dx_r. \quad (\text{B.20})$$

Let  $G_n$ ,  $G_{T^*n}$  to be the  $r \times r$  matrices with entries  $n_{ij}$ ,  $(T^*n)_{ij}$  in coordinates  $\{y_i\}_{i=1}^r$ ,  $\{x_i\}_{i=1}^r$  respectively. Then by Lemma B.2.2, one has  $\det(G_{T^*n}) = |\det(J_T)|^2$ .

$\det(G_n) \circ T$ , which implies

$$\begin{aligned} T^*(\omega_n^r) &= \sqrt{\det G_n} \circ T \cdot T^*(dy_1 \wedge \dots \wedge dy_r) \quad \text{by (B.9)} \\ &= \sqrt{\det G_n} \circ T \cdot |\det J_T| \cdot dx_1 \wedge \dots \wedge dx_r \quad \text{by (B.20)} \\ &= \sqrt{\det G_{T^*n}} \cdot dx_1 \wedge \dots \wedge dx_r = \omega_{T^*n}^r. \end{aligned}$$

□

Recall that  $T$  is an isometry from  $(M, T^*n)$  to  $(N, n)$ . Due to Corollary B.2.3 one has

$$\int_{T(U)} \omega_n^r = \int_U \omega_{T^*n}^r = \int_U T^*(\omega_n^r), \quad (\text{B.21})$$

for all measurable  $U \subseteq M$ . Hence, by the definition of  $\mathcal{P}$  given by (3.17), one has for all  $f \in L^1(M, V_m)$

$$\begin{aligned} \int_U f \cdot \omega_m^r &= \int_{T(U)} \mathcal{P}f \cdot \omega_n^r \\ &= \int_U \mathcal{P}f \circ T \cdot \omega_{T^*n}^r \quad \text{by (B.21)} \\ &= \int_U \mathcal{P}f \circ T \cdot \sqrt{\det G_{T^*n}} \cdot dx_1 \wedge \dots \wedge dx_r \\ &= \int_U \mathcal{P}f \circ T \cdot |\det J_T| \cdot \sqrt{\det G_n} \circ T \cdot dx_1 \wedge \dots \wedge dx_r, \end{aligned}$$

where the last line is due to Lemma B.2.2. Hence, since  $T$  is a diffeomorphism and  $\omega_m^r = \sqrt{\det G_m} \cdot dx_1 \wedge \dots \wedge dx_r$  by (B.10), one has

$$\begin{aligned} \mathcal{P}f &= \frac{f \circ T^{-1}}{|\det J_T \circ T^{-1}|} \cdot \frac{\sqrt{\det G_m} \circ T^{-1}}{\sqrt{\det G_n}} \\ &= f \circ T^{-1} \cdot |\det J_{T^{-1}}| \cdot \frac{\sqrt{\det G_m} \circ T^{-1}}{\sqrt{\det G_n}}, \end{aligned} \quad (\text{B.22})$$

where we have applied the inverse function theorem to obtain the last line. Moreover, setting  $f = h_\mu$  in (B.22) and using the fact that  $\mathcal{P}h_\mu = h_\nu$  (by (3.17)) yields

$$h_\mu = h_\nu \circ T \cdot |\det J_T| \cdot \frac{\sqrt{\det G_n} \circ T}{\sqrt{\det G_m}}. \quad (\text{B.23})$$

Now by assumption,  $T$  is a diffeomorphism and the densities  $h_\mu$  and  $h_\nu$  are uniformly bounded away from zero. Therefore, by (B.23) and the nondegeneracy of the metrics  $m$  and  $n$ , the Jacobian  $|\det J_T|$  is bounded above and uniformly away from zero.

Let  $\mathcal{I}_V$  denote the characteristic function on a measurable subset  $V \subseteq N$ . One has for all  $f \in L^1(M, V_m)$

$$\int_N \mathcal{P}f \cdot \mathcal{I}_V \cdot \omega_n^r = \int_V \mathcal{P}f \cdot \omega_n^r = \int_{T^{-1}V} f \cdot \omega_m^r = \int_M f \cdot \mathcal{I}_V \circ T \cdot \omega_m^r. \quad (\text{B.24})$$

Hence, the Koopman operator  $\mathcal{K}$  adjoint to  $\mathcal{P}$  is given by  $\mathcal{K}f = f \circ T$ .

Recall the pushforward operator  $\mathcal{H} : L^2(M, m, \mu_r) \rightarrow L^2(N, n, \nu_r)$  is defined as in (3.19), with  $L^2(M, m, \mu_r)$  adjoint  $\mathcal{H}^*$ .

**Lemma B.2.4.** *The operator  $\mathcal{H} : L^2(M, m, \mu_r) \rightarrow L^2(N, n, \nu_r)$  is well defined, may be expressed as  $\mathcal{H}f = f \circ T^{-1}$ , and has adjoint  $\mathcal{H}^*g = g \circ T$ .*

*Proof.* Let  $f \in L^2(M, m, \mu_r)$ . Due to (B.22) and the fact that  $h_\mu > 0$ , one has

$$\begin{aligned}
 |\mathcal{P}(f \cdot h_\mu)|^2 &= |(f \cdot h_\mu) \circ T^{-1}|^2 \cdot |\det J_{T^{-1}}|^2 \cdot \left| \frac{\sqrt{\det G_m \circ T^{-1}}}{\sqrt{\det G_n}} \right|^2 \\
 &= \left| (f^2 \cdot h_\mu) \circ T^{-1} \cdot |\det J_{T^{-1}}| \cdot \frac{\sqrt{\det G_m \circ T^{-1}}}{\sqrt{\det G_n}} \right| \\
 &\quad \times \left| h_\mu \circ T^{-1} \cdot |\det J_{T^{-1}}| \cdot \frac{\sqrt{\det G_m \circ T^{-1}}}{\sqrt{\det G_n}} \right| \\
 &= |\mathcal{P}(f^2 \cdot h_\mu)| \cdot |\mathcal{P}h_\mu| \\
 &= \mathcal{P}(f^2 \cdot h_\mu) \cdot h_\nu.
 \end{aligned} \tag{B.25}$$

Therefore

$$\begin{aligned}
 \int_N |\mathcal{H}f|^2 d\nu_r &= \int_N \left| \frac{\mathcal{P}(f \cdot h_\mu)}{h_\nu} \right|^2 \cdot h_\nu \omega_n^r \\
 &= \int_N \frac{|\mathcal{P}(f \cdot h_\mu)|^2}{h_\nu} \cdot \omega_n^r \quad \text{since } h_\nu > 0 \\
 &= \int_N \mathcal{P}(f^2 \cdot h_\mu) \cdot \omega_n^r \\
 &= \int_M (f^2 \cdot h_\mu) \cdot \omega_m^r \\
 &= \int_M f^2 d\mu_r,
 \end{aligned} \tag{B.26}$$

where the second last line is due to (3.17). Thus, since  $f \in L^2(M, m, \mu_r)$  the RHS of (B.26) is bounded and  $\mathcal{H}$  is well defined.

To show that  $\mathcal{H}f = f \circ T^{-1}$ , we use (B.25) (without the squares) to compute  $\mathcal{P}(f \cdot h_\mu)$ , and (B.22) to compute  $h_\nu = \mathcal{P}h_\mu$ , and note that all terms in the quotient  $\mathcal{H} = \mathcal{P}(f \cdot h_\mu)/h_\nu$  not involving  $f$  cancel to leave  $\mathcal{H} = f \circ T^{-1}$ .

For all measurable  $U \subseteq M$ ,

$$\int_{T(U)} \mathcal{H}f \cdot h_\nu \omega_n^r = \int_{T(U)} \mathcal{P}(f \cdot h_\mu) \cdot \omega_n^r = \int_U f \cdot h_\mu \omega_m^r. \tag{B.27}$$

Let  $U \subseteq M$  be measurable. Since  $\mathcal{P}h_\mu = h_\nu$ , one has  $\mathcal{P}(\mathcal{I}_U \cdot h_\mu) = \mathcal{I}_{T(U)} \cdot h_\nu$ .

Therefore,

$$\begin{aligned}
 \int_{T(U)} g \, d\nu_r &= \int_N \frac{\mathcal{I}_{T(U)} \cdot h_\nu}{h_\nu} \cdot g \, d\nu_r \\
 &= \int_N \frac{\mathcal{P}(\mathcal{I}_U \cdot h_\mu)}{h_\nu} \cdot g \, d\nu_r \\
 &= \int_N \mathcal{H}(\mathcal{I}_U) \cdot g \, d\nu_r \\
 &= \int_M \mathcal{I}_U \cdot \mathcal{H}^* g \, d\mu_r \quad \text{by definition of } \mathcal{H}^* \\
 &= \int_U \mathcal{H}^* g \, d\mu_r, \tag{B.28}
 \end{aligned}$$

for all  $g \in L^2(N, n, \nu_r)$ . Therefore, using the fact that  $\nu_r = \mu_r \circ T^{-1}$ , one has for any measurable  $V \subset N$

$$\int_M \mathcal{H}^* \mathcal{I}_V \, d\mu_r = \int_N \mathcal{I}_V \, d\nu_r = \int_M \mathcal{I}_{T^{-1}V} \, d\mu_r = \int_M \mathcal{I}_V \circ T \, d\mu_r. \tag{B.29}$$

Thus,  $\mathcal{H}^* g = g \circ T$  for all  $g \in L^2(N, n, \nu_r)$ .  $\square$

**Proposition B.2.5.** *Let  $\mathcal{H} : L^2(M, m, \mu_r) \rightarrow L^2(N, n, \nu_r)$  be as in (3.19), with adjoint  $\mathcal{H}^*$ . For any  $f \in C^1(M, \mathbb{R}) \cap L^2(M, m, \mu_r)$ , one has*

$$T\{x \in M : f(x) = \beta\} = \{y \in N : \mathcal{H}f(y) = \beta\}.$$

*Proof.* This follows immediately from Lemma B.2.4  $\square$

**Lemma B.2.6.** *Let  $\mathcal{H} : L^2(M, m, \mu_r) \rightarrow L^2(N, n, \nu_r)$  be as in (3.19), with adjoint  $\mathcal{H}^*$ . One has*

1.  $\nabla_n = T_* \nabla_{T^*n} \mathcal{H}^*$ ,
2.  $\mathcal{H}^* \operatorname{div}_n T_* = \operatorname{div}_{T^*n}$ ,
3.  $\mathcal{H}^* \Delta_n \mathcal{H}f = \Delta_{T^*n}$ .

*Proof.* 1. Let  $g \in C^1(N, \mathbb{R}) \cap L^2(N, n, \mu_r)$  and  $\mathcal{V} \in \mathcal{F}^1(M)$ . One has by (3.7)

$$\begin{aligned}
 n(T_* \nabla_{T^*n} \mathcal{H}^* g, T_* \mathcal{V})(Tx) &= T^* n(\nabla_{T^*n} \mathcal{H}^* g, \mathcal{V})(x) \\
 &= \mathcal{V}(\mathcal{H}^* g)|_x \quad \text{by (3.16) with respect to } T^*n \\
 &= \mathcal{V}(g \circ T)|_x \quad \text{by (B.29)} \\
 &= (T_* \mathcal{V})g|_{Tx} \\
 &= n(\nabla_n g, T_* \mathcal{V})(Tx),
 \end{aligned}$$

for all  $x \in M$ .

2. Let  $\mathcal{V}, \mathcal{V}_1, \mathcal{V}_2, \dots, \mathcal{V}_{r-1}$  be  $r$  vector fields in  $\mathcal{F}^1(M)$ . One has at each point  $x \in M$

$$\begin{aligned}
 & [i(T_*\mathcal{V})\omega_n^r](T_*\mathcal{V}_1, T_*\mathcal{V}_2, \dots, T_*\mathcal{V}_{r-1})(Tx) \\
 &= \omega_n^r(T_*\mathcal{V}, T_*\mathcal{V}_1, T_*\mathcal{V}_2, \dots, T_*\mathcal{V}_{r-1})(Tx) \\
 &= (T^*\omega_n^r)(\mathcal{V}, \mathcal{V}_1, \mathcal{V}_2, \dots, \mathcal{V}_{r-1})(x) \\
 &= [i(\mathcal{V})\omega_{T^*n}^r](\mathcal{V}_1, \mathcal{V}_2, \dots, \mathcal{V}_{r-1})(x),
 \end{aligned}$$

where we have applied the identity  $T^*\omega_n^r = \omega_{T^*n}^r$  in Corollary B.2.3 on the last line. Hence, by the duality of  $T_*$  and  $T^*$ , one has at each point  $x \in M$

$$T^*d[i(T_*\mathcal{V})\omega_n^r] = d[i(\mathcal{V})\omega_{T^*n}^r]. \quad (\text{B.30})$$

Therefore,

$$\begin{aligned}
 \int_U \mathcal{H}^* \operatorname{div}_n(T_*\mathcal{V}) \cdot \omega_{T^*n}^r &= \int_U \operatorname{div}_n(T_*\mathcal{V}) \circ T \cdot \omega_{T^*n}^r \\
 &= \int_{TU} \operatorname{div}_n(T_*\mathcal{V}) \cdot \omega_n^r \quad \text{by (B.21)} \\
 &= \int_{TU} d[i(T_*\mathcal{V})\omega_n^r] \quad \text{by (B.13) with respect to } n \\
 &= \int_U T^*d[i(T_*\mathcal{V})\omega_n^r] \\
 &= \int_U d[i(\mathcal{V})\omega_{T^*n}^r] \quad \text{by (B.30)} \\
 &= \int_U \operatorname{div}_{T^*n}(\mathcal{V}) \cdot \omega_{T^*n}^r.
 \end{aligned}$$

3. Due to 1. and 2. and the fact that  $\mathcal{H}^*\mathcal{H}$  is the identity by B.2.4, one has  $\mathcal{H}^*\Delta_n\mathcal{H}f = \mathcal{H}^*\operatorname{div}_n(\nabla_n\mathcal{H}f) = \mathcal{H}^*\operatorname{div}_n(T_*\nabla_{T^*n}\mathcal{H}^*\mathcal{H}f) = \operatorname{div}_{T^*n}\nabla_{T^*n}f = \Delta_{T^*n}f$ , for all  $f \in C^2(M, \mathbb{R}) \cap L^2(M, m, \mu_r)$ . □

**Corollary B.2.7.** *Let  $\mathcal{H} : L^2(M, m, \mu_r) \rightarrow L^2(N, n, \nu_r)$  be as in (3.19), with adjoint  $\mathcal{H}^*$ . One has*

$$\Delta^{\operatorname{dyn}} f = \frac{1}{2} (\Delta_m + \mathcal{H}^* \Delta_n \mathcal{H}) f + \frac{1}{2} \left( \frac{m(\nabla_m h_\mu, \nabla_m f)}{h_\mu} + \frac{n(\nabla_n h_\nu, \nabla_n \mathcal{H}f) \circ T}{h_\nu \circ T} \right), \quad (\text{B.31})$$

for all  $f \in C^2(M, \mathbb{R}) \cap L^2(M, m, \mu_r)$ .

*Proof.* By definition

$$\Delta^{\operatorname{dyn}} = \Delta_\mu + \mathcal{H}^* \Delta_\nu \mathcal{H}. \quad (\text{B.32})$$

Substituting a straightforward modification of (3.27) into the second term on the RHS of (B.32), one has for all  $x \in M$  and  $f \in C^2(M, \mathbb{R}) \cap L^2(M, m, \mu_r)$ ,

$$\begin{aligned} \mathcal{H}^* \Delta_\nu \mathcal{H}f(x) &= \Delta_\nu(\mathcal{H}f)(Tx) \\ &= \Delta_n(\mathcal{H}f)(Tx) + \frac{n(\nabla_n h_\nu, \nabla_n \mathcal{H}f)(Tx)}{h_\nu(Tx)} \\ &= \mathcal{H}^* \Delta_n \mathcal{H}f(x) + \frac{n(\nabla_n h_\nu, \nabla_n \mathcal{H}f)(Tx)}{h_\nu \circ T(x)}. \end{aligned}$$

Similarly, one can expand the first term of (B.32) using (3.27) to obtain the required result.  $\square$

**Corollary B.2.8.** *Let  $\Delta^{dyn}$  and  $\Delta_\mu$  be defined by (3.28) and (3.27) respectively. One has*

$$\Delta^{dyn} = \frac{1}{2}(\Delta_\mu + \Delta_{\bar{\mu}})f, \quad (\text{B.33})$$

where  $\Delta_{\bar{\mu}}$  is given by (3.27) with respect to the metric  $T^*n$  and density  $h_\nu \circ T$ .

*Proof.* Due to Lemma B.2.6, one has  $\nabla_n = T_* \nabla_{T^*n} \mathcal{H}^*$ . Therefore, by the definition of the gradient (3.16) with respect to the metric  $n$ , one has for all  $x \in M$

$$\begin{aligned} n(\nabla_n h_\nu, \nabla_n \mathcal{H}f)_{Tx} &= (\nabla_n \mathcal{H}f)h_\nu|_{Tx} \\ &= (T_* \nabla_{T^*n} f)h_\nu|_{Tx} \quad \text{by Lemma B.2.4} \\ &= (\nabla_{T^*n} f)(h_\nu \circ T)|_x \\ &= T^*n(\nabla_{T^*n}(h_\nu \circ T), \nabla_{T^*n} f)_x, \end{aligned} \quad (\text{B.34})$$

where the equality on the last line is due to (3.16) with respect to the metric  $T^*n$ . Moreover, by Lemma B.2.6, one has the identity  $\mathcal{H}^* \Delta_n \mathcal{H} = \Delta_{T^*n}$ . Thus, by substituting (B.34) into the fourth term on the RHS of (B.31), one has

$$\begin{aligned} \Delta^{dyn} f &= \frac{1}{2}(\Delta_m + \mathcal{H}^* \Delta_n \mathcal{H})f + \frac{1}{2} \left( \frac{m(\nabla_m h_\mu, \nabla_m f)}{h_\mu} + \frac{T^*n(\nabla_{T^*n}(h_\nu \circ T), \nabla_{T^*n} f)}{h_\nu \circ T} \right) \\ &= \frac{1}{2}(\Delta_\mu + \Delta_{\bar{\mu}})f, \end{aligned} \quad (\text{B.35})$$

where the second equality is due to the definition of weighted Laplacians (3.27).  $\square$

## B.2.4 Local properties of charts

An important analytical tool for reducing a global calculation on  $M$  to local calculations on each chart of an atlas on  $M$  is the *partition of unity*.

**Definition B.2.9.** Let  $(U_i, \varphi_i)_{i \in I}$  be an atlas on  $M$ . A *partition of unity* subordinate to the covering  $\{U_i\}_{i \in I}$ , is the collection of smooth functions  $\sigma_i \in C^\infty(M, \mathbb{R})$  such that:

1.  $\text{supp}(\sigma_i) \subset U_i$ .
2. Any point  $x \in M$  has a neighbourhood  $\mathcal{O}_x$  such that  $\mathcal{O}_x \cap \text{supp}(\sigma_i) = \emptyset$  except for a finite set of  $\sigma_i$ .
3.  $0 \leq \sigma_i \leq 1$  and  $\sum_{i \in I} \sigma_i = 1$ .

It is well known that the partition of unity exist for paracompact manifolds (see e.g. Theorem 1.12 [7]). Since every compact manifold is paracompact, a partition of unity exist for  $M$ .

Furthermore, for each point  $x$  in a compact Riemannian manifold  $M$ , there exist coordinates on a neighbourhood about  $x$ , and a constant  $c > 1$  (depending on the injective radius of  $x$ , and the dimension of the sectional curvature of  $M$ ), such that

$$\frac{1}{c} \delta_{ij} \leq m_{ij} \leq c \delta_{ij}, \quad 1 \leq i, j \leq r, \quad (\text{B.36})$$

where  $\delta_{ij}$  is the Kronecker delta (see e.g p.507 in [78], or Chapter 1 of [71]). The following lemmas are consequences of (B.36).

**Lemma B.2.10.** *Let  $(U, \varphi)$  be a chart on  $(M, m)$ , set  $\Omega = \varphi(U)$ , and denote by  $d\ell$  the density with respect the Lebesgue measure. One has*

$$c^{-r/2} \int_U |f|^p d\mu_r \leq \int_{\Omega} |f \circ \varphi^{-1}|^p \cdot (h_{\mu} \circ \varphi^{-1}) d\ell \leq c^{r/2} \int_U |f|^p d\mu_r, \quad (\text{B.37})$$

for some real number  $c > 1$  and all  $f \in L^p(U, m, \mu_r)$ ,  $p \in [1, \infty)$ .

*Proof.* Let  $\delta_{ij}$  denote the Kronecker delta, and pick local coordinates on  $U$  such that the components of the metric tensor  $m$  satisfy  $\frac{1}{c} \delta_{ij} \leq m_{ij}(x) \leq c \delta_{ij}$  for all  $x \in U$  and  $1 \leq i, j \leq r$ . Due to the inequality  $m_{ij} \leq \frac{1}{c} \delta_{ij}$ , one has  $\sqrt{\det G_m(x)} \leq c^{r/2}$  for all  $x \in U$ . Furthermore, the Riemannian volume form is given by  $\omega_m^r = \sqrt{\det G_m} \cdot dx_1 \wedge dx_2 \wedge \dots \wedge dx_r$  on  $U$ , and the Lebesgue density satisfies  $d\ell = (\varphi^{-1})^*(dx_1 \wedge dx_2 \wedge \dots \wedge dx_r)$  on  $\Omega$ . Hence by the change of variable formula (B.21)

$$\begin{aligned} c^{-r/2} \int_U |f|^p d\mu_r &= c^{-r/2} \int_U |f|^p \cdot h_{\mu} \sqrt{\det G_m} \cdot dx_1 \wedge dx_2 \wedge \dots \wedge dx_r \\ &\leq \int_{\varphi^{-1}(\Omega)} |f|^p \cdot h_{\mu} \cdot dx_1 \wedge dx_2 \wedge \dots \wedge dx_r \\ &= \int_{\Omega} |f \circ \varphi^{-1}|^p \cdot (h_{\mu} \circ \varphi^{-1}) d\ell, \end{aligned}$$

where the equality is due to (B.21). The inequality  $\int_{\Omega} |f \circ \varphi^{-1}|^p \cdot (h_{\mu} \circ \varphi^{-1}) d\ell \leq c^{r/2} \int_U |f|^p d\mu_r$  is obtained analogously using  $\frac{1}{c} \delta_{ij} \leq m_{ij}$ .  $\square$

**Lemma B.2.11.** *Let  $(U, \varphi)$  be a chart on  $(M, m)$ , set  $\Omega = \varphi(U)$ , and denote by  $e$  the Euclidean metric on  $\Omega$  with respect to the Lebesgue density  $d\ell$ . One has*

$$c^{-(r/2+1)} \int_U |\nabla_m f|_m^p d\mu_r \leq \int_\Omega |\nabla_e(f \circ \varphi^{-1})|_e^p \cdot (h_\mu \circ \varphi^{-1}) d\ell \leq c^{(r/2+1)} \int_U |\nabla_m f|_m^p d\mu_r$$

for some real number  $c > 1$  and all  $\nabla_m f \in L^p(U, m, \mu_r)$ ,  $p \in [1, \infty)$ .

*Proof.* We start with the case  $p = 2$ . Let  $\delta_{ij}$  denote the Kronecker delta, and pick local coordinates on  $U$  such that the components of the metric tensor  $m$  satisfy  $\frac{1}{c}\delta_{ij} \leq m_{ij}(x) \leq c\delta_{ij}$  for all  $x \in U$  and  $1 \leq i, j \leq r$ . Denote by  $m^{ij}$  the components of the inverse matrix  $G_m^{-1}$ . One has the contraction  $\sum_k m^{ik}m_{kj} = \delta_{ij}$ , so that  $\frac{1}{c}\delta_{ij} \leq m^{ij}(x) \leq c\delta_{ij}$ . Moreover, due to Lemma B.2.10, the inequality (B.37) is valid with constant  $c$ . Hence, by writing  $\nabla_m f$  in the given local coordinates via (B.12), one has

$$\begin{aligned} & c^{-(r/2+1)} \int_U |\nabla_m f|_m^2 d\mu_r \\ & \leq c^{-1} \int_\Omega (|\nabla_m f|_m^2 \cdot h_\mu) \circ \varphi^{-1} d\ell \\ & = c^{-1} \int_\Omega m(\nabla_m f, \nabla_m f)_{\varphi^{-1}(x)} \cdot h_\mu \circ \varphi^{-1}(x) d\ell(x) \\ & = c^{-1} \int_\Omega \sum_{i,j=1}^r m_{ij} \left( \sum_{k=1}^r m^{ki} \frac{\partial(f \circ \varphi^{-1})}{\partial x_k} \right) \left( \sum_{l=1}^r m^{lj} \frac{\partial(f \circ \varphi^{-1})}{\partial x_l} \right) \cdot (h_\mu \circ \varphi^{-1}) d\ell \\ & = c^{-1} \int_\Omega \sum_{j=1}^r \left( \frac{\partial(f \circ \varphi^{-1})}{\partial x_j} \right) \left( \sum_{l=1}^r m^{lj} \frac{\partial(f \circ \varphi^{-1})}{\partial x_l} \right) \cdot (h_\mu \circ \varphi^{-1}) d\ell, \end{aligned} \quad (\text{B.38})$$

where we have contracted the index  $i$  to obtain the last line. Furthermore, using the fact that  $m^{lj} \leq c\delta_{lj}$ , one has

$$\begin{aligned} \text{RHS of (B.38)} & \leq \int_\Omega \sum_{j=1}^r \left( \frac{\partial(f \circ \varphi^{-1})}{\partial x_j} \right)^2 \cdot (h_\mu \circ \varphi^{-1}) d\ell \quad \text{since } m^{lj} \leq c\delta_{lj} \\ & = \int_\Omega |\nabla_e(f \circ \varphi^{-1})|_e^2 \cdot (h_\mu \circ \varphi^{-1}) d\ell. \end{aligned} \quad (\text{B.39})$$

The inequality  $\int_\Omega |\nabla_e(f \circ \varphi^{-1})|_e^2 \cdot (h_\mu \circ \varphi^{-1}) d\ell \leq c^{(r/2+1)} \int_U |\nabla_m f|_m^2 d\mu_r$  is obtained analogously using  $\frac{1}{c}\delta_{ij} \leq m_{ij}$ .

The general case  $p \in [1, \infty)$  is a straightforward modification of the calculation done to obtain (B.38) and (B.39).  $\square$

### B.3 Weighted Sobolev spaces

Let  $C_0^\infty(\Omega, \mathbb{R})$  be the space of smooth real-valued functions with compact support on  $\Omega \subset \mathbb{R}^r$ , and  $\ell$  the Lebesgue measure on  $\mathbb{R}^r$ . For locally integrable functions

$f, \tilde{f} \in L^1_{\text{loc}}(\Omega, \ell)$ , we say that  $\tilde{f}$  is the first order weak derivative of  $f$  if  $\int_{\Omega} f \cdot \partial_i g \, d\ell = -\int_{\Omega} \tilde{f} \cdot g \, d\ell$  for all  $g \in C_0^\infty(\Omega, \mathbb{R})$ , and each  $1 \leq i \leq r$  (see p.21 in [1]). We write  $\tilde{f} = \tilde{\partial}_i f$ , and note that  $\tilde{\partial}_i f$  is uniquely determined up to sets of measure zero.

**Definition B.3.1.** Let  $(U, \varphi)$  be a chart on  $M$  with corresponding local coordinates  $(x_1, x_2, \dots, x_r)$ . Define the first order weak gradient of  $f \in L^1_{\text{loc}}(M, V_m)$  at the point  $x \in M$  by

$$\tilde{\nabla}_m f(x) = \sum_{i,j=1}^r m^{ij}(x) \cdot \tilde{\partial}_i(f \circ \varphi^{-1}) \Big|_{\varphi(x)} \partial_j, \quad (\text{B.40})$$

where the partial derivatives appearing on the RHS exist in the weak sense.

It is straightforward to extend the operation  $T_*$  on weak gradients, and verify that Lemma B.2.6 and B.2.11 hold for weak gradients. In addition, if the density of  $\mu_r$  is an  $A_p$  weight, then by Proposition B.1.3, any  $f \in L^p(M, m, \mu_r)$  is also in  $L^1_{\text{loc}}(M, V_m)$ . Thus, one can define weak gradients on  $L^p(M, m, \mu_r)$  via the Definition B.3.1. The following proposition provides the key motivation behind the construction of the weak gradient given above.

**Proposition B.3.2.** Let  $f \in L^2(M, m, \mu_r)$ , where the density of  $\mu_r$  is an  $A_2$  weight. Assume the first order weak gradient of  $f$  defined by (B.40) exists. One has

$$\int_U f \cdot \Delta_\mu g \, d\mu_r = - \int_U m(\tilde{\nabla}_m f, \nabla_m g) \, d\mu_r, \quad (\text{B.41})$$

for all measurable  $U \subset M$  and  $g \in C_0^\infty(M, \mathbb{R}^r)$ .

*Proof.* Let  $(U_k, \varphi_k)_{k \in K}$  be an atlas on  $M$ , with corresponding local coordinates  $(x_1, x_2, \dots, x_r)$ . Due to (B.10), one has  $d\mu_r = h_\mu \sqrt{\det G_m} dx_1 \wedge dx_2 \wedge \dots \wedge dx_r$ . Additionally  $d\ell = (\varphi_k^{-1})^*(dx_1 \wedge dx_2 \wedge \dots \wedge dx_r)$ . Hence, for each  $k \in K$  and any measurable  $\Omega_k \subset \varphi_k(U_k)$ , one has by the coordinate representation of  $\Delta_m$  given by (B.15), for each  $k \in K$

$$\begin{aligned} \int_{\varphi_k^{-1}(\Omega_k)} f \cdot \Delta_\mu g \, d\mu_r &= \int_{\varphi_k^{-1}(\Omega_k)} f \cdot \sum_{i,j=1}^r \partial_i(m^{ij} h_\mu \sqrt{\det G_m}) \partial_j g \cdot dx_1 \wedge dx_2 \wedge \dots \wedge dx_r \\ &= \sum_{i,j=1}^r \int_{\Omega_k} f \circ \varphi_k^{-1} \cdot \partial_i[(m^{ij} h_\mu \sqrt{\det G_m}) \circ \varphi_k^{-1} \cdot (\partial_j g) \circ \varphi_k^{-1}] \, d\ell \\ &= \sum_{i,j=1}^r \int_{\Omega_k} \tilde{\partial}_i(f \circ \varphi_k^{-1}) \cdot [(m^{ij} h_\mu \sqrt{\det G_m}) \circ \varphi_k^{-1} \cdot (\partial_j g) \circ \varphi_k^{-1}] \, d\ell \\ &= \int_{\varphi_k^{-1}(\Omega_k)} m(\tilde{\nabla}_m f, \nabla_m g) \, d\mu_r, \end{aligned} \quad (\text{B.42})$$

where the last line is due to  $m(\tilde{\nabla}_m f, \nabla_m g) = (\tilde{\nabla}_m f)g = \sum_{i,j=1}^r m^{ij} \partial_i f \partial_j g$  and Definition B.3.1.

Now since  $M$  is compact, there exists a smooth partition of unity  $\sigma_k$  subordinate to the covering  $\{U_k\}_{k \in K}$  (see Definition B.2.9). Moreover, since  $\varphi_k$  is a diffeomorphism, for any measurable  $U \subseteq M$ , there exist  $K' \subseteq K$  and countable collection of measurable  $\Omega_k \subset \varphi_k(U_k)$ , such that  $U = \cup_{k \in K'} (\varphi_k^{-1}(\Omega_k))$ . Hence, applying (B.42) to each  $k \in K'$ , one has by setting  $\sum_{k \in K'} \sigma_k = 1$

$$\begin{aligned} \int_U f \cdot \Delta_\mu g \, d\mu_r &= \sum_{k \in K'} \int_{\varphi_k^{-1}(\Omega_k)} \sigma_k f \cdot \Delta_\mu g \, d\mu_r \\ &= \sum_{k \in K'} \int_{\varphi_k^{-1}(\Omega_k)} m(\tilde{\nabla}_m(\sigma_k f), \nabla_m g) \, d\mu_r \\ &= \int_U m \left( \tilde{\nabla}_m \left( \sum_{k \in K'} \sigma_k f \right), \nabla_m g \right) \, d\mu_r \\ &= \int_U m(\tilde{\nabla}_m f, \nabla_m g) \, d\mu_r, \end{aligned}$$

where we have used the linearity of  $\tilde{\nabla}_m$  and the fact that  $\text{supp}(\sigma_k) \subset U_k$  to obtain the penultimate line.  $\square$

We introduce the weighted Sobolev space  $W^{1,2}(M, m, \mu_r)$  of  $L^2(M, m, \mu_r)$  integrable functions, whose first order weak gradient exists in  $L^2(M, m, \mu_r)$ . We equip  $W^{1,2}(M, m, \mu_r)$  with the inner product  $\langle f, g \rangle_{W^{1,2}(M, m, \mu_r)} := \int_M (m(\tilde{\nabla}_m f, \tilde{\nabla}_m g) + fg) \, d\mu_r$  for all  $f, g \in W^{1,2}(M, m, \mu_r)$ , with the norm associated with  $\langle \cdot, \cdot \rangle_{W^{1,2}(M, m, \mu_r)}$  denoted by  $\| \cdot \|_{W^{1,2}(M, m, \mu_r)}$ .

There exist embedding theorems and the completeness property for weighted Sobolev spaces on  $\mathbb{R}^r$ , and for the unweighted Sobolev spaces on Riemannian manifolds (see [119] and [71] respectively). We develop the corresponding results for the weighted Sobolev space  $W^{1,2}(M, m, \mu_r)$  defined as above. Let  $(U, \varphi)$  be a chart on  $M$ . In the following, we first obtain the results of the desired properties in a local setting; i.e. the weighted Sobolev space  $W^{1,2}(U, m, \mu_r)$ . One can then use the fact that  $M$  is compact, and apply the standard partition of unity arguments to extend these local outcomes to global ones for  $W^{1,2}(M, m, \mu_r)$ .

Given a chart  $(U, \varphi)$  on  $M$ . Set  $\Omega = \varphi(U)$ , and let  $\ell_\mu$  be an absolutely continuous measure with density  $h_\mu \circ \varphi^{-1}$  with respect to  $\ell$ , where  $\ell$  is the Lebesgue measure on  $\mathbb{R}^r$ . One has the weighted Sobolev space  $W^{1,2}(\Omega, \ell_\mu)$  for the open subset  $\Omega \subset \mathbb{R}^r$ ; that is, the space  $W^{1,2}(\Omega, \ell_\mu)$  is equipped with the norm

$$\|f \circ \varphi^{-1}\|_{W^{1,2}(\Omega, \ell_\mu)}^2 = \int_\Omega \left( |f \circ \varphi^{-1}|^2 + |\tilde{\nabla}_e(f \circ \varphi^{-1})|_e^2 \right) \cdot (h_\mu \circ \varphi^{-1}) \, d\ell, \quad (\text{B.43})$$

for all  $f \in L^2(U, \mu_r)$ , and where  $\tilde{\nabla}_e$  is the first order weak gradient with respect to the Euclidean metric  $e$ . Suppose the density  $h_\mu$  of  $\mu_r$  is an  $A_2$  weight (i.e.  $h_\mu$

satisfies (B.1) when  $p = 2$ ). Clearly,  $h_\mu$  is an  $A_2$  weight restricted to the sub-domain  $U$ . Moreover, since  $U$  has compact closure and  $\varphi$  is a diffeomorphism, it is easy to verify that the density  $h_\mu \circ \varphi^{-1}$  of  $\ell_\mu$  is also an  $A_2$  weight. Since the density of  $\ell_\mu$  is an  $A_2$  weight, the weighted Sobolev space  $W^{1,2}(\Omega, \ell_\mu)$  is a Hilbert space and  $C^\infty(\Omega, \mathbb{R})$  is dense in  $W^{1,2}(\Omega, \ell_\mu)$  (see Theorem 1 in [62]). We show by the following lemma that if  $f \in W^{1,2}(U, m, \mu_r)$ , then  $f \circ \varphi^{-1} \in W^{1,2}(\Omega, \ell_\mu)$ .

**Lemma B.3.3.** *Let  $(U, \varphi)$  be a chart on  $M$ , and set  $\Omega = \varphi(U)$ . Denote by  $\|\cdot\|_{W^{1,2}(\Omega, \ell_\mu)}$  and  $\|\cdot\|_{W^{1,2}(U, m, \mu_r)}$  the norms on the weighted Sobolev spaces  $W^{1,2}(\Omega, \ell_\mu)$  and  $W^{1,2}(U, m, \mu_r)$  respectively. Then  $\|f \circ \varphi^{-1}\|_{W^{1,2}(\Omega, \ell_\mu)}$  and  $\|f\|_{W^{1,2}(U, m, \mu_r)}$  are equivalent for all  $f \in W^{1,2}(U, m, \mu_r)$ .*

*Proof.* This follows immediately from Lemma B.2.10 and B.2.11.  $\square$

Due to Lemma B.3.3, one now has global completeness for  $W^{1,2}(M, m, \mu_r)$ .

**Proposition B.3.4.** *Assume the density of  $\mu_r$  is an  $A_2$  weight. The weighted Sobolev space  $W^{1,2}(M, m, \mu_r)$  is complete.*

*Proof.* First we show that the Sobolev spaces on any charts on  $M$  are complete. Let  $(U, \varphi)$  be a chart on  $M$ , and  $f_j$  a Cauchy sequence in  $W^{1,2}(U, m, \mu_r)$ . Then  $f_j \circ \varphi^{-1}$  is Cauchy in  $W^{1,2}(\Omega, \ell_\mu)$  due to Lemma B.3.3, so by the completeness of  $W^{1,2}(\Omega, \ell_\mu)$ , the Cauchy sequence  $f_j \circ \varphi^{-1}$  converges to an element  $f \circ \varphi^{-1} \in W^{1,2}(\Omega, \ell_\mu)$ . Hence, the Cauchy sequence  $f_j$  converges to  $f$  in  $W^{1,2}(U, m, \mu_r)$ .

Now, let  $g_j$  be a Cauchy sequence in  $W^{1,2}(M, m, \mu_r)$ , and  $(U_i, \varphi_i)_{i \in I}$  an atlas on  $M$ . Since  $M$  is compact,  $\{U_i\}_{i \in I}$  is a finite cover for  $M$ . Hence, there exist a fixed  $s \in I$  such that  $W^{1,2}(U_s, m, \mu_r)$  contains infinitely many terms of the sequence  $g_j$ . Let  $g_{j_k}$  be a subsequence of  $g_j$  contained entirely in  $W^{1,2}(U_s, m, \mu_r)$ , then  $g_{j_k}$  is Cauchy in  $W^{1,2}(U_s, m, \mu_r)$ , so that  $g_{j_k}$  converges to an element  $g \in W^{1,2}(U_s, m, \mu_r)$  by completeness. In particular, the Cauchy sequence  $g_j$  converges to  $g$  in  $W^{1,2}(M, m, \mu_r)$ .  $\square$

We proceed to demonstrate that the space  $W^{1,2}(M, m, \mu_r)$  is approximated by smooth functions in  $C^\infty(M, \mathbb{R}) \cap W^{1,2}(M, m, \mu_r)$ . The key idea is to locally subject the functions in  $W^{1,2}(M, m, \mu_r)$  to *mollification*.

**Definition B.3.5.** Let  $\Omega$  be an open subset of  $\mathbb{R}^r$ , and  $q \in C_0^\infty(\mathbb{R}^r, \mathbb{R})$  be nonnegative such that  $\text{supp}(q) \subset E_1(0)$  and  $\int_\Omega q \, d\ell = 1$ , where  $E_1(0)$  is the open unit ball centered at the origin in  $\mathbb{R}^r$ . We define a *mollifier* by the function  $q_\epsilon := \epsilon^{-r} q(x/\epsilon)$ . For all  $f \in L^p(\Omega, \ell)$ ,  $p \in [1, \infty)$ , we call the convolution

$$q_\epsilon \star f(x) := \int_\Omega q_\epsilon(x - z) f(z) \, d\ell(z), \quad (\text{B.44})$$

the mollification of  $f$  by  $q_\epsilon$ .

One has the following weighted version of the well known result  $\tilde{\nabla}_\epsilon(q_\epsilon \star f) = q_\epsilon \star \tilde{\nabla}_\epsilon f$ , and density theorem (Lemma 7.3 and Theorem 7.9 in [61] respectively).

**Theorem B.3.6** (Theorem 2.1.4. [119]). *Let  $\Omega$  be an open subset of  $\mathbb{R}^r$ , and  $f \in L^p(\Omega, \ell_w)$ , where  $\ell_w$  is an absolutely continuous measure with respect to Lebesgue. Define  $f_\epsilon := q_\epsilon \star f$ , where the mollifier  $q_\epsilon$  and  $\star$  are as in Definition B.3.5. For  $p \in [1, \infty)$ , if the density of  $\ell_w$  is an  $A_p$  weight, then  $f_\epsilon \in C^\infty(\Omega, \mathbb{R}) \cap L^p(\Omega, \ell_w)$ ,  $\tilde{\nabla}_\epsilon f_\epsilon = q_\epsilon \star \tilde{\nabla}_\epsilon f$ , and as  $\epsilon \rightarrow 0$ ,  $f_\epsilon \rightarrow f$  in  $L^p(\Omega, \ell_w)$ .*

**Corollary B.3.7.** *Let  $W^{1,2}(M, m, \mu_r)$  be a weighted Sobolev space. Assume the density of  $\mu_r$  is an  $A_2$  weight. The space  $C^\infty(M, \mathbb{R}) \cap W^{1,2}(M, m, \mu_r)$  is dense in  $W^{1,2}(M, m, \mu_r)$ .*

*Proof.* Let  $f \in W^{1,2}(M, m, \mu_r)$  and choose some  $\gamma > 0$ . We will show that there is a  $g \in C^\infty(M, \mathbb{R}) \cap W^{1,2}(M, m, \mu_r)$ , such that  $\|f - g\|_{W^{1,2}(M, m, \mu_r)}^2 < \gamma$ . Let  $(U_i, \varphi_i)_{i \in I}$  be an atlas on  $M$ , then pick local coordinates on  $M$ , such that the components of the metric tensor satisfy  $\frac{1}{c}\delta_{jk} \leq m_{jk}(x) \leq c\delta_{jk}$  for some  $1 < c < \infty$ , and each  $x \in U_i$ ,  $1 \leq j, k \leq r$  (such coordinate exist due to the compactness of  $M$ , see (B.36)). For each  $1 \leq i \leq r$ , let  $\ell_{\mu,i}$  be an absolutely continuous measure with density  $h_\mu \circ \varphi_i^{-1}$  with respect to  $\ell$ . Since  $M$  is compact, there exists a smooth partition of unity  $\{\sigma_i\}_{i \in I}$  subordinate to the finite covering  $\{U_i\}_{i \in I}$ ; i.e.  $\sigma_i$  is given by definition B.2.9. Therefore,  $\sigma_i f$  and its first order weak gradient vanishes outside of  $U_i$ , hence  $\sigma_i f \in W^{1,2}(U_i, m, \mu_r)$ . Set  $\Omega_i = \varphi_i(U_i)$  for each  $i \in I$ , then due to Lemma B.3.3, the fact that  $\sigma_i f \in W^{1,2}(U_i, m, \mu_r)$  implies  $(\sigma_i f) \circ \varphi_i^{-1} \in W^{1,2}(\Omega_i, \ell_{\mu,i})$ . Consequently, both  $(\sigma_i f) \circ \varphi_i^{-1}$  and  $\tilde{\nabla}_\epsilon((\sigma_i f) \circ \varphi_i^{-1})$  are in  $L^2(\Omega_i, \ell_{\mu,i})$  for each  $i \in I$ . Let  $q_\epsilon$  and  $\star$  be as in definition B.3.5, then by Theorem B.3.6 applied to  $(\sigma_i f) \circ \varphi_i^{-1} \in L^2(\Omega_i, \ell_{\mu,i})$  with  $p = 2$  and  $\ell_w = \ell_{\mu,i}$  for each  $i \in I$ ,

$$\tilde{\nabla}_\epsilon(q_\epsilon \star ((\sigma_i f) \circ \varphi_i^{-1})) = q_\epsilon \star \left( \tilde{\nabla}_\epsilon((\sigma_i f) \circ \varphi_i^{-1}) \right), \quad (\text{B.45})$$

and there exist  $\epsilon_1 > 0$  such that

$$\|q_\epsilon \star ((\sigma_i f) \circ \varphi_i^{-1}) - (\sigma_i f) \circ \varphi_i^{-1}\|_{L^2(\Omega_i, \ell_{\mu,i})} < \frac{\gamma}{2c^{r/4} \cdot |I|^2}. \quad (\text{B.46})$$

for all  $i \in I$ . In addition, applying Theorem B.3.6 to  $\tilde{\nabla}_\epsilon((\sigma_i f) \circ \varphi_i^{-1}) \in L^2(\Omega_i, \ell_{\mu,i})$  with  $p = 2$  and  $\ell_\mu = \ell_{\mu,i}$ , one has  $\epsilon_2 > 0$  such that

$$\begin{aligned} & \left\| \tilde{\nabla}_\epsilon(q_\epsilon \star ((\sigma_i f) \circ \varphi_i^{-1})) - \tilde{\nabla}_\epsilon((\sigma_i f) \circ \varphi_i^{-1}) \right\|_{L^2(\Omega_i, \ell_{\mu,i})} \\ &= \left\| q_\epsilon \star \left( \tilde{\nabla}_\epsilon((\sigma_i f) \circ \varphi_i^{-1}) \right) - \tilde{\nabla}_\epsilon((\sigma_i f) \circ \varphi_i^{-1}) \right\|_{L^2(\Omega_i, \ell_{\mu,i})} \quad \text{by (B.45)} \\ &< \frac{\gamma}{2c^{r/4+2} \cdot |I|^2}, \end{aligned} \quad (\text{B.47})$$

for all  $i \in I$ .

Let  $\text{dist}_m$  denote the distance function admitted by the metric  $m$ . Set

$$\epsilon = \min \{ \epsilon_1, \epsilon_2, \text{dist}_m (\text{supp}[q_\epsilon \star ((\sigma_i f) \circ \varphi_i^{-1})], \partial\Omega_i) \}, \quad (\text{B.48})$$

and let  $f_{\epsilon,i} := q_\epsilon \star ((\sigma_i f) \circ \varphi_i^{-1}) \circ \varphi_i$ . Since  $\epsilon$  satisfies (B.48), the function  $f_{\epsilon,i}$  and its first order weak gradient  $\tilde{\nabla}_m f_{\epsilon,i}$  vanish outside of  $U_i$  for all  $i \in I$ . Moreover, since  $|I|$  is finite,  $f_\epsilon := \sum_{i \in I} f_{\epsilon,i} \in C^\infty(M, \mathbb{R}) \cap W^{1,2}(M, m, \mu_r)$ .

Set  $g = f_\epsilon$ , then by Lemma B.2.10 and the inequality (B.46)

$$\begin{aligned} \|g - f\|_{2,m,\mu} &= \left\| \sum_{i \in I} f_{\epsilon,i} - \sigma_i f \right\|_{2,m,\mu} \\ &\leq \sum_{i \in I} \left( \int_M |f_{\epsilon,i} - \sigma_i f|^2 d\mu_r \right)^{\frac{1}{2}} \quad \text{by triangle inequality} \\ &= \sum_{i \in I} \left( \int_{U_i} |f_{\epsilon,i} - \sigma_i f|^2 d\mu_r \right)^{\frac{1}{2}} \\ &\leq \sum_{i \in I} c^{r/4} \left( \int_{\Omega_i} |q_\epsilon \star ((\sigma_i f) \circ \varphi_i^{-1}) - (\sigma_i f) \circ \varphi_i^{-1}|^2 \cdot (h_\mu \circ \varphi_i^{-1}) d\ell \right)^{\frac{1}{2}} \\ &< \sqrt{\gamma/2}. \end{aligned} \quad (\text{B.49})$$

Similarly, by Lemma B.2.11, and the inequality (B.47)

$$\begin{aligned} &\|\tilde{\nabla}_m g - \tilde{\nabla}_m f\|_{2,m,\mu} \quad (\text{B.50}) \\ &= \left\| \sum_{i \in I} \tilde{\nabla}_m f_\epsilon - \sigma_i f \right\|_{2,m,\mu} \\ &\leq \sum_{i \in I} \left( \int_{U_i} |\tilde{\nabla}_m f_{\epsilon,i} - \tilde{\nabla}_m (\sigma_i f)|_m^2 d\mu_r \right)^{\frac{1}{2}} \\ &\leq \sum_{i \in I} c^{r/4+1/2} \left( \int_{\Omega_i} \left| \tilde{\nabla}_e (q_\epsilon \star ((\sigma_i f) \circ \varphi_i^{-1})) - \tilde{\nabla}_e ((\sigma_i f) \circ \varphi_i^{-1}) \right|_e^2 \cdot (h_\mu \circ \varphi_i^{-1}) d\ell \right)^{\frac{1}{2}} \\ &< \sqrt{\gamma/2}. \end{aligned} \quad (\text{B.51})$$

Thus,  $\|g - f\|_{W^{1,2}(M,m,\mu_r)}^2 < \gamma$ .  $\square$

As before, let  $\ell_\mu$  be an absolutely continuous measure with respect to  $\ell$ . The *Hardy-Littlewood maximal operator*  $\mathcal{M}$  is a non-linear operator on locally integrable functions  $f \in L^1_{\text{loc}}(\mathbb{R}^r, \ell)$  defined by

$$\mathcal{M}f(x) = \sup_{\rho>0} \frac{1}{\ell(E_\rho(x))} \int_{E_\rho(x)} |f(y)| d\ell(y), \quad (\text{B.52})$$

where  $E_\rho(x)$  is the Euclidean ball centered at  $x$  with radius  $\rho$ . If the density of the measure  $\ell_\mu$  is an  $A_p$  weight, then  $\mathcal{M}$  is bounded as an operator from  $L^p(\mathbb{R}^r, \ell_w)$  to  $L^p(\mathbb{R}^r, \ell_w)$  for  $1 < p < \infty$  (see Theorem 1, p.201 in [114], or Theorem 1.2.3 in [119]). This property of  $A_p$  weights on the operator  $\mathcal{M}$  forms an essential argument for Theorem B.3.6, and leads to the following result:

**Lemma B.3.8.** *Let  $K \subset \Omega \subset \mathbb{R}^r$ , where  $K$  is compact and  $\Omega$  is open and bounded. Let  $W^{1,2}(\Omega, \ell_w)$  be a weighted Sobolev space. Assume the density  $\ell_w$  is an  $A_2$  weight. Suppose  $f_i$  is a sequence in  $W^{1,2}(\Omega, \ell_w)$  with support  $K$ . Then there exists a subsequence  $f_{i_j}$ , and some  $f \in L^2(\Omega, \ell_w)$ , such that*

$$\int_K (f_{i_j} - f)^2 d\ell_w \rightarrow 0, \quad (\text{B.53})$$

as  $j \rightarrow \infty$ .

*Proof.* Let  $f_{\epsilon,i} = q_\epsilon \star f_i$ , where  $q_\epsilon$  and  $\star$  are as in Definition B.3.5. Denote by  $\|\cdot\|_{2,\ell_w}$  the  $L^2$ -norm associated with  $L^2(K, \ell_w)$ . First we show that

$$\|f_{\epsilon,i} - f_i\|_{2,\ell_w} \rightarrow 0, \quad (\text{B.54})$$

uniformly with respect to  $i$ . Due to Theorem B.3.6, it is sufficient to proof (B.54) for  $f_i \in C^\infty(\Omega, \mathbb{R}) \cap W^{1,2}(\Omega, \ell_w)$ . By a change of variable  $y = (x - z)/\epsilon$  and using the facts  $\int_{\mathbb{R}^r} q d\ell = 1$ ,  $\text{supp}(q) \subset E_1(0)$ , one has

$$\begin{aligned} f_{\epsilon,i}(x) - f_i(x) &= \int_{\mathbb{R}^r} q_\epsilon(x - z) f_i(z) d\ell(z) - f_i(x) \\ &= \epsilon^{-r} \int_{E_\epsilon(x)} q\left(\frac{x - z}{\epsilon}\right) f_i(z) d\ell(z) - \int_{|y| < 1} q(y) f_i(x) d\ell(y) \\ &= \int_{|y| < 1} q(y) [f_i(x - \epsilon y) - f_i(x)] d\ell(y) \\ &\leq \|q\|_\infty \cdot \int_{|y| < 1} [f_i(x - \epsilon y) - f_i(x)] d\ell(y) \quad \text{by Hölder's inequality} \\ &= \|q\|_\infty \cdot \epsilon^r \int_{E_\epsilon(x)} [f_i(z) - f_i(z + \epsilon y)] d\ell(z). \end{aligned} \quad (\text{B.55})$$

As a consequence of (B.55), one has

$$\begin{aligned} \|f_{\epsilon,i} - f_i\|_{2,\ell_w}^2 &\leq \|q\|_\infty^2 \cdot \int_K \left| \epsilon^r \int_{E_\epsilon(x)} |f_i(z) - f_i(z + \epsilon y)| d\ell(z) \right|^2 d\ell_w(x) \\ &\leq \|q\|_\infty^2 \cdot \int_{\mathbb{R}^r} \left| \epsilon^r [\ell(E_\epsilon(x))] \cdot \mathcal{H}(f_i(x) - f_i(x + \epsilon y)) \right|^2 d\ell_w(x) \\ &= \|q\|_\infty^2 \cdot \pi^2 \cdot \int_{\mathbb{R}^r} \left| \mathcal{H}(f_i(x) - f_i(x + \epsilon y)) \right|^2 d\ell_w(x), \end{aligned} \quad (\text{B.56})$$

where  $\mathcal{H}$  is defined as in (B.52). Furthermore, by using the fact that the density of  $\ell_w$  is an  $A_p$  weight, the operator  $\mathcal{H} : L^2(\mathbb{R}^r, \ell_w) \rightarrow L^2(\mathbb{R}^r, \ell_w)$  is bounded. This implies

$$\begin{aligned} \text{RHS of (B.56)} &\leq \|q\|_\infty^2 \cdot \pi^2 \cdot \int_{\mathbb{R}^r} |\mathcal{H}(f_i(x) - f_i(x + \epsilon y))|^2 d\ell_w(x) \\ &\leq \|q\|_\infty^2 \cdot \pi^2 \cdot C \int_{\mathbb{R}^r} |f_i(x) - f_i(x + \epsilon y)|^2 d\ell_w(x), \end{aligned} \quad (\text{B.57})$$

where the constant  $C$  depends only on  $r, p$  and the  $A_p$  constant of  $w$ ; the constant  $C$  is uniform with respect to  $i$ . Since  $f_i$  is continuous independent of  $i$ , by (B.56)-(B.57) one has the convergence  $f_{\epsilon,i} \rightarrow f_i$  in  $L^2(\Omega, \ell_w)$  uniformly with respect to  $i$ .

Due to (B.54), we can now pick a subsequence of  $f_i$ , so that for any fixed  $\gamma > 0$ , there exists an  $\epsilon$  sufficiently small such that

$$\|f_{\epsilon,i_j} - f_{i_j}\|_{2,\ell_w} \leq \gamma/2, \quad (\text{B.58})$$

for all  $i_j \geq 1$ . Furthermore, since the density of  $\ell_w$  is an  $A_2$  weight, by a straightforward modification of Lemma B.1.3, one has  $L^2(\Omega, \ell_w) \subset L^1_{\text{loc}}(\Omega, \ell)$ . Hence the sequence  $f_i$  belongs to  $L^1_{\text{loc}}(\Omega, \ell)$ , therefore

$$\int_{\Omega} |f_i| d\ell = \int_K |f_i| d\ell < \infty,$$

which implies

$$\begin{aligned} \sup_{x \in \Omega} |f_{\epsilon,i}(x)| &= \sup_{x \in \Omega} |q_\epsilon \star f_i(x)| \\ &= \sup_{x \in \Omega} \left| \int_{\Omega} q_\epsilon(x-z) \cdot f_i(z) d\ell(z) \right| \\ &\leq \|q_\epsilon\|_\infty \cdot \int_{\Omega} |f_i| d\ell < \infty, \end{aligned}$$

so that  $f_{\epsilon,i}$  is uniformly bounded on  $\Omega$ . Similarly, by using Leibniz's rule for differentiating under integral sign, one has

$$\begin{aligned} \sup_{x \in \Omega} |\nabla_e f_{\epsilon,i}(x)|_e &= \sup_{x \in \Omega} |\nabla_e q_\epsilon \star f_i(x)|_e \\ &= \sup_{x \in \Omega} \left| \nabla_e \left( \int_{\Omega} q_\epsilon(x-z) \cdot f_i(z) d\ell(z) \right) \right|_e \\ &\leq \sup_{x \in \Omega} \int_{\Omega} |\nabla_e q_\epsilon(x-z)|_e \cdot |f_i(z)| d\ell(z) \\ &\leq \|\nabla_e q_\epsilon\|_\infty \cdot \int_{\Omega} |f_i| d\ell < \infty, \end{aligned}$$

which implies  $f_{\epsilon,i}$  is equicontinuous on  $\Omega$ . Therefore, by the Arzela-Ascoli theorem (Theorem 11.28 in [104]), there exist a subsequence  $f_{\epsilon,i_j}$  that convergences uniformly

on every compact subset of  $\Omega$ . In particular, there is an  $f_\epsilon$  such that

$$\begin{aligned}
 & \lim_{j,k \rightarrow \infty} \|f_{i_j} - f_{i_k}\|_{2,\ell_w}^2 \\
 & \leq \lim_{i,k \rightarrow \infty} \left\{ \|f_{i_j} - f_{\epsilon,i_j}\|_{2,\ell_w}^2 + \|f_{\epsilon,i_j} - f_\epsilon\|_{2,\ell_w}^2 + \|f_\epsilon - f_{\epsilon,i_k}\|_{2,\ell_w}^2 + \|f_{\epsilon,i_k} - f_{i_k}\|_{2,\ell_w}^2 \right\} \\
 & \leq \gamma/2 + 0 + 0 + \gamma/2 = \gamma,
 \end{aligned} \tag{B.59}$$

where the convergence of the first and last term on the RHS was handled by (B.58). Hence  $f_{i_j}$  is a Cauchy sequence in  $L^2(K, \ell_\mu)$ . Therefore, by the completeness of  $L^2$  spaces, the Cauchy sequence  $f_{i_j}$  converges to some  $f$  in  $L^2(K, \ell_w)$ .  $\square$

One now has the weighted version of the well known Sobolev compactness embedding theorem for  $\mathbb{R}^r$ , which applies to  $W^{1,2}(M, m, \mu_r)$ .

**Theorem B.3.9** (Rellich Compactness). *Let  $W^{1,2}(M, m, \mu_r)$  be a weighted Sobolev space. If the density of  $\mu_r$  is an  $A_2$  weight, then the embedding  $W^{1,2}(M, m, \mu_r) \hookrightarrow L^2(M, m, \mu_r)$  is compact.*

*Proof.* Let  $f_j$  be a sequence in  $W^{1,2}(M, m, \mu_r)$ , and  $(U_i, \varphi_i)_{i \in I}$  an atlas on  $M$ . As in the proof of Corollary B.3.7, pick local coordinates on  $M$  such that the components of the metric tensor satisfy  $\frac{1}{c}\delta_{sp} \leq m_{sp}(x) \leq c\delta_{sp}$  for some  $1 < c < \infty$ , and all  $x \in U_i$ ,  $i \in I$ ,  $1 \leq s, p \leq r$ . Furthermore, let  $\{\sigma_i\}_{i \in I}$  be a partition of unity subordinate to the finite covering  $\{U_i\}_{i \in I}$ . For each  $i \in I$ , set  $\Omega_i = \varphi_i(U_i)$ . One has  $\sigma_i f_j \in W^{1,2}(U_i, m, \mu_r)$ , so by Lemma B.3.3 the sequence  $(\sigma_i f_j) \circ \varphi_i^{-1}$  belongs to  $W^{1,2}(\Omega_i, \ell_{\mu,i})$ , where  $\ell_{\mu,i}$  is an absolutely continuous measure with density  $h_\mu \circ \varphi_i^{-1}$  with respect to  $\ell$ . Moreover, the compactness of the closure of  $M$  implies  $(\sigma_i f_j) \circ \varphi_i^{-1}$  has compact support  $C_i \subset \Omega_i$ . Therefore, for each  $i \in I$  one can apply Lemma B.3.8 with  $\ell_w = \ell_{\mu,i}$  to obtain a subsequence  $(\sigma_i f_{j_k}) \circ \varphi_i^{-1}$ , and a function  $g_i$  in  $L^2(\Omega_i, \ell_{\mu,i})$ , such that for any  $\gamma > 0$ , there exist a  $K(\gamma) \in \mathbb{N}$  with

$$\left( \int_{C_i} |(\sigma_i f_{j_k}) \circ \varphi_i^{-1} - g_i|^2 d\ell_{\mu,i} \right)^{\frac{1}{2}} < \frac{\gamma}{c^{r/4} \cdot |I|^2}, \tag{B.60}$$

for all  $k \geq K(\gamma)$ .

Since each  $\varphi_i$  is a diffeomorphism and  $g_i \in L^2(\Omega_i, \ell_{\mu,i})$ , each  $g_i \circ \varphi_i$  belongs to  $L^2(U_i, m, \mu_r)$ . Extend  $g_i \circ \varphi_i$  to  $\tilde{g}_i \in L^2(M, m, \mu_r)$ , by setting

$$\tilde{g}_i(x) := \begin{cases} g_i \circ \varphi_i(x) & x \in U_i \\ 0 & x \in M \setminus U_i \end{cases},$$

for each  $i \in I$ . Then by a similar argument as in (B.49)

$$\begin{aligned} \left\| f_{j_k} - \sum_{i \in I} \tilde{g}_i \right\|_{2,m,\mu} &\leq \sum_{i \in I} \left( c^{r/2} \int_{\Omega_i} |\sigma_i f_{j_k} - \tilde{g}_i|^2 \circ \varphi_i^{-1} d\ell_{\mu,i} \right)^{\frac{1}{2}} \\ &= \sum_{i \in I} c^{r/4} \left( \int_{C_i} |(\sigma_i f_{j_k}) \circ \varphi_i^{-1} - g_i|^2 d\ell_{\mu,i} \right)^{\frac{1}{2}} \\ &< \gamma, \end{aligned}$$

where the inequality on the last line is due to (B.60). Since  $\tilde{g}_i \in L^2(U_i, m, \mu_r)$  and  $|I|$  is finite, we have  $\sum_{i \in I} \tilde{g}_i \in L^2(M, m, \mu_r)$ ; this completes the proof of the theorem.  $\square$

**Lemma B.3.10** (Poincaré inequality). *Let  $W^{1,2}(M, m, \mu_r)$  be a weighted Sobolev space, and denote by  $\alpha(f)$  the weighted mean of  $f$ ; i.e.  $\alpha(f) = \int_M f \cdot h_\mu \omega_m^r$ . Assume the density of  $\mu_r$  is an  $A_2$  weight. There is a constant  $K$  depending on  $r$  and  $M$  such that*

$$\|f - \alpha(f)\|_{2,m,\mu} \leq K \|\tilde{\nabla}_m f\|_{2,m,\mu}, \quad (\text{B.61})$$

for all  $f \in W^{1,2}(M, m, \mu_r)$ .

*Proof.* We follow a standard argument as in corollary of Theorem 5 on p.194, [88]. Suppose the inequality (B.61) is false, then due to Corollary B.3.7, there exists a sequence in  $f_k \in C^\infty(M, \mathbb{R}) \cap W^{1,2}(M, m, \mu_r)$ , such that  $\|f_k - \alpha(f_k)\|_{2,m,\mu} > k \|\nabla_m f_k\|_{2,m,\mu}$  for  $k = 1, 2, \dots$ . Define

$$g_k = \frac{f_k - \alpha(f_k)}{\|f_k - \alpha(f_k)\|_{2,m,\mu}},$$

then  $\|g_k\|_{2,m,\mu} = 1$ ,  $\alpha(g_k) = 0$  and  $\|\nabla_m g_k\|_{2,m,\mu} \leq 1/k$ . In particular,  $g_k$  is a bounded sequence in  $W^{1,2}(M, m, \mu_r)$ . Hence, by Theorem B.3.9 there exists a subsequence  $g_{k_j} \in W^{1,2}(M, m, \mu_r)$ , which converges to some  $g$  in  $L^2(M, m, \mu_r)$ . One has

$$\|g\|_{2,m,\mu} = 1, \quad (\text{B.62})$$

and

$$\alpha(g) = \int_M g \cdot h_\mu \omega_m^r = \lim_{j \rightarrow \infty} \int_M g_{k_j} \cdot h_\mu \omega_m^r = \lim_{j \rightarrow \infty} \alpha(g_{k_j}) = 0 \quad (\text{B.63})$$

and  $\lim_{j \rightarrow \infty} \|\nabla_m g_{k_j}\|_{2,m,\mu} = 0$ .

Now, for any  $\psi \in C_0^\infty(M, \mathbb{R})$ , the weak gradient of  $g$  satisfies

$$\begin{aligned}
 \int_M m(\tilde{\nabla}_m g, \nabla_m \psi) d\mu_r &= - \int_M g \Delta_\mu \psi d\mu_r \\
 &= - \lim_{j \rightarrow \infty} \int_M g_{k_j} \Delta_\mu \psi d\mu_r \\
 &= \lim_{j \rightarrow \infty} \int_M m(\nabla_m g_{k_j}, \nabla_m \psi) d\mu_r \\
 &\leq \lim_{j \rightarrow \infty} \|\nabla_m g_{k_j}\|_{2,m,\mu} \cdot \|\nabla_m \psi\|_{2,m,\mu} \\
 &= \left( \lim_{j \rightarrow \infty} \frac{1}{k_j} \right) \|\nabla_m \psi\|_{2,m,\mu} = 0.
 \end{aligned}$$

Therefore

$$\tilde{\nabla}_m g = 0. \tag{B.64}$$

But since  $M$  is connected, (B.63) and (B.64) implies  $g$  is the zero function, which contradicts (B.62).  $\square$

## B.4 The proof of Theorem 3.2.4

To obtain the inequality  $\mathbf{s}^{dyn} \leq \mathbf{H}_M^{dyn}$ , let  $\Gamma$  be a compact, connected  $C^\infty$  hypersurface in  $M$  that disconnects  $M$  into two open disjoint subsets  $M_1$  and  $M_2$ . Let  $\text{dist}_m(x_1, x_2)$  denote the Riemannian distance function with respect to the metric tensor  $m$  between the points  $x_1$  and  $x_2$  in  $M$ , then define  $U_\epsilon := \{x \in M : \text{dist}_m(x, \Gamma) < \epsilon\}$  for  $\epsilon > 0$ .

Consider the set of functions

$$f_\epsilon(x) := \begin{cases} 1, & x \in M_1 \setminus U_\epsilon \\ -1, & x \in M_2 \setminus U_\epsilon \\ (1/\epsilon)\text{dist}_m(x, \Gamma), & x \in M_1 \cap U_\epsilon \\ -(1/\epsilon)\text{dist}_m(x, \Gamma), & x \in M_2 \cap U_\epsilon \end{cases}. \tag{B.65}$$

In the following, we obtain an upper bound for  $\mathbf{s}^{dyn}$  by locally approximating functions in  $C^\infty(M, \mathbb{R})$  by  $f_\epsilon$ .

**Lemma B.4.1.** *Let  $\mathcal{H}$ ,  $\mathbf{s}^{dyn}$  and  $f_\epsilon$  be defined by (3.19), (3.20) and (B.65) respectively. If the density of  $\mu_r$  is  $C^1$ , then for  $\epsilon > 0$  sufficiently small, one has*

$$\mathbf{s}^{dyn} \leq \frac{\|\nabla_m f_\epsilon\|_{1,m,\mu} + \|\nabla_n \mathcal{H} f_\epsilon\|_{1,n,\nu}}{2 \inf_\beta \|f_\epsilon - \alpha\|_{1,m,\mu}}. \tag{B.66}$$

*Proof.* We claim the existence of  $g \in C^\infty(M, \mathbb{R})$ , such that the terms  $\nabla_m f_\epsilon$  and  $f_\epsilon - \beta$  are approximated by  $\nabla_m g$  and  $g - \alpha$  respectively in the norm  $\|\cdot\|_{1,m,\mu}$ , and

the term  $\nabla_m \mathcal{H}f$  is approximated by  $\nabla_n \mathcal{H}g$  in the norm  $\|\cdot\|_{1,n,\nu}$ . In particular, due to these smooth approximations and the definition of  $\mathbf{s}^{dyn}$ , one immediately obtains the required inequality (B.66).

Let  $(U_i, \varphi_i)_{i \in I}$  be an atlas of  $M$ , and set  $\Omega_i = \varphi_i(U_i)$  for each  $i \in I$ . For each  $i \in I$ , let  $\ell_{\mu,i}$  be an absolutely continuous measure with density  $h_\mu \circ \varphi_i^{-1}$  with respect to Lebesgue measure  $\ell$ . Since  $M$  is compact, there exist a smooth partition of unity  $\{\sigma_i\}_{i \in I}$  subordinate to the finite covering  $\{U_i\}_{i \in I}$ . Moreover, one can verify that  $f_\epsilon$  is a Lipschitz function in  $L^1(M, m, \mu_r)$ . Therefore  $(\sigma_i f_\epsilon) \circ \varphi_i^{-1}$  is Lipschitz in  $L^1(\Omega_i, \ell_{\mu,i})$  for each  $i \in I$ . It follows that the restriction of  $(\sigma_i f_\epsilon) \circ \varphi_i^{-1}$  to any line in  $\Omega_i$  is absolutely continuous, which implies all partial derivatives of  $(\sigma_i f_\epsilon) \circ \varphi_i^{-1}$  exist almost everywhere on  $\Omega_i$  (see Theorem 7.20 in [104]). Therefore, the Euclidean gradient  $\nabla_e((\sigma_i f_\epsilon) \circ \varphi_i^{-1}) \in L^1(\Omega_i, \ell_{\mu,i})$  for each  $i \in I$ .

Set  $f_{\delta,\epsilon} := q_\delta \star f_\epsilon$ , where  $q_\delta$  and  $\star$  is as in Definition B.3.5. Then by straightforward modifications to the arguments used in Corollary B.3.7 to obtain (B.49) and (B.50). One can obtain for any  $\gamma > 0$ , a  $\delta > 0$  chosen analogously to (B.48) such that  $f_{\delta,\epsilon} \in C^\infty(M, \mathbb{R})$ ,

$$\|\nabla_m f_{\delta,\epsilon} - \nabla_m f_\epsilon\|_{1,m,\mu} < \gamma \quad (\text{B.67})$$

and

$$\|(f_{\delta,\epsilon} - \alpha) - (f_\epsilon - \alpha)\|_{1,m,\mu} = \|f_{\delta,\epsilon} - f_\epsilon\|_{1,m,\mu} < \gamma.$$

Finally, since  $T$  is a diffeomorphism and  $h_\mu$  is  $C^1$ ,  $\mathcal{H}f_\epsilon$  is Lipschitz in  $L^1(N, n, \nu_r)$ . Thus, the approximation of  $\nabla_n \mathcal{H}f_\epsilon$  by  $\nabla_n \mathcal{H}f_{\delta,\epsilon}$  in the norm  $\|\cdot\|_{1,n,\nu}$  can be obtained analogously to (B.67). Thus, setting  $g = f_{\delta,\epsilon}$  proves the claim.  $\square$

To complete the proof of Theorem 3.2.4, we show that the RHS of (B.66) is bounded above by  $\mathbf{H}_M^{dyn}$  as  $\epsilon \rightarrow 0$ . In order to show such convergence holds, we require additional results concerning the connection between  $\mu_r(U_\epsilon)$  and  $\mu_{r-1}(\Gamma)$ .

Suppose  $\epsilon$  is smaller than the injectivity radius of each point  $x \in \Gamma$ , and recall that  $U_\epsilon := \{x \in M : \text{dist}_m(x, \Gamma) < \epsilon\}$  are open subsets of  $M$ . Since  $M$  is compact, the closure of  $U_\epsilon$  is a compact subset of  $M$ . Due to the compactness of  $\overline{U_\epsilon}$  and the size of  $\epsilon$ , by the Hopf-Rinow theorem  $\overline{U_\epsilon}$  is geodesically complete [35]. This implies that the signed distance function  $f : U_\epsilon \rightarrow \mathbb{R}$  defined by

$$f(x) := \begin{cases} \text{dist}_m(x, \Gamma) & x \in M_1 \\ -\text{dist}_m(x, \Gamma) & x \in M_2 \\ 0 & x \in \Gamma \end{cases}, \quad (\text{B.68})$$

is smooth and  $|\nabla_m f|_m = 1$  on  $U_\epsilon \setminus \Gamma$  (Proposition 2.1 [106]).

The following concerns the regularity of the co-dimensional one measure  $\mu_{r-1}$  on the level surfaces of  $U_\epsilon$ .

**Lemma B.4.2.** *Let  $\Gamma$  be a  $C^\infty$  hypersurface in  $M$  that disconnects  $M$  into two disjoint open subsets  $M_1$  and  $M_2$ . Define  $\Gamma^\beta := \{x \in M : \text{dist}_m(x, \Gamma) = \beta\}$ , and fix  $\epsilon$  to be smaller than the injectivity radius of each point  $x \in \Gamma$ . If the density of  $\mu_r$  is continuous, then the real valued function  $A$  given by*

$$A(\beta) := \mu_{r-1}(\Gamma^\beta),$$

*is continuous on the intervals  $[-\epsilon, 0]$  and  $[0, \epsilon]$ .*

*Proof.* Let  $f : M \rightarrow \mathbb{R}$  be the signed distance function as in (B.68), and let  $U_\epsilon := \{x \in M : \text{dist}_m(x, \Gamma) < \epsilon\}$ . Fix  $\beta_0 \in (0, \epsilon)$ , then  $\Gamma^{\beta_0}$  is in  $U_\epsilon \setminus \Gamma$ . Hence  $f$  is  $C^\infty$  restricted to  $\Gamma^{\beta_0}$ , and  $df(x) \neq 0$  for each  $x \in \Gamma^{\beta_0}$ . Therefore, by the implicit function theorem there exist open neighborhoods  $\mathcal{O}_x$  about each point  $x \in \Gamma^{\beta_0}$ , and local coordinates  $(x_1, x_2, \dots, x_{r-1})$  for  $\Gamma^{\beta_0}$ , such that  $(x_1, x_2, \dots, x_{r-1}, f)$  are local coordinates on  $\mathcal{O}_x$ . Let  $G_m$  be the  $r \times r$  matrix with entries  $m_{ij}$  in the coordinates  $(x_1, x_2, \dots, x_{r-1}, f)$ . Then the volume form on  $\mathcal{O}_x$  is given by

$$\omega_m^r = \sqrt{\det(G_m)} \cdot dx_1 \wedge dx_2 \dots \wedge dx_{r-1} \wedge df.$$

Moreover, by a combination of the Stokes' and divergence theorem (see p.122, [113] and p.7, equation (38) [22] respectively), one has

$$\int_{\Gamma^{\beta_0}} \omega_m^{r-1} = \int_{\Gamma^{\beta_0}} m(\mathbf{n}, \mathbf{n}) \cdot \omega_m^{r-1} = \int_{U_{\beta_0}} \text{div}_m \mathbf{n} \cdot \omega_m^r = \int_{U_{\beta_0}} d(i(\mathbf{n})\omega_m^r) = \int_{\Gamma^{\beta_0}} i(\mathbf{n})\omega_m^r,$$

where  $\mathbf{n}$  is the unit normal bundle along  $\Gamma^{\beta_0}$ . Hence  $\omega_m^{r-1} = i(\mathbf{n})\omega_m^r$  for all  $x \in \Gamma^{\beta_0}$ .

Now, since  $f = \beta_0$  along  $\Gamma^{\beta_0}$ , the vector  $\nabla_m f$  is normal to the hypersurface  $\Gamma^{\beta_0}$ ; which implies  $\mathbf{n} = \nabla_m f / |\nabla_m f|_m$ , and  $dx_i(\nabla_m f) = 0$  for  $i = 1, \dots, r-1$ . Therefore

$$\begin{aligned} \omega_m^{r-1} \Big|_{\Gamma^{\beta_0}} &= i(\mathbf{n})\omega_m^r \Big|_{\Gamma^{\beta_0}} \\ &= \sqrt{\det G_m} \cdot i(\mathbf{n})(dx_1 \wedge dx_2 \dots \wedge dx_{r-1} \wedge df) \Big|_{\Gamma^{\beta_0}} \\ &= (-1)^r \sqrt{\det G_m} \cdot \frac{df(\nabla_m f)}{|\nabla_m f|_m} \cdot dx_1 \wedge dx_2 \dots \wedge dx_{r-1} \Big|_{\Gamma^{\beta_0}} \\ &= (-1)^r \sqrt{\det G_m} \cdot |\nabla_m f|_m \cdot dx_1 \wedge dx_2 \dots \wedge dx_{r-1} \Big|_{\Gamma^{\beta_0}}, \end{aligned} \quad (\text{B.69})$$

where the penultimate equality is due to Leibniz rule on interior product, and the fact that  $dx_i(\nabla_m f) = 0$  for  $i = 1, \dots, r-1$ .

To complete the proof, we note that  $|\nabla_m f| = 1$  on the  $U_\epsilon \setminus \Gamma$  because  $f$  is the signed distance function,  $h_\mu$  is continuous by assumption, and  $G_m$  is smooth since  $m$  is smooth. Hence  $h_\mu \omega_m^{r-1} \Big|_{\Gamma^\beta}$  is a continuous density for all  $\beta \in (-\epsilon, 0) \cup (0, \epsilon)$ . Therefore  $A(\beta) = \mu_{r-1}(\Gamma^\beta) = \int_{\Gamma^\beta} h_\mu \omega_m^{r-1}$  is continuous on  $[-\epsilon, 0]$  and  $[0, \epsilon]$ .  $\square$

**Lemma B.4.3.** *Let  $\Gamma$  be a compact, connected  $C^\infty$  hypersurface in  $M$ . Define  $U_\epsilon := \{x \in M : \text{dist}_m(x, \Gamma) < \epsilon\}$  for some  $\epsilon > 0$ . Assume the density of  $\mu_r$  is continuous. One has*

$$\lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} \mu_r(U_\epsilon) = 2\mu_{r-1}(\Gamma). \quad (\text{B.70})$$

*Proof.* Let  $f$  be the signed distance function as in (B.68), and  $\Gamma^\beta = \{x \in M : \text{dist}_m(x, \Gamma) = \beta\}$ . Then  $|\nabla_m f|_m = 1$ , and  $f$  is  $C^\infty$  on  $U_\epsilon \setminus \Gamma$ . Hence, by the co-area formula (B.11)

$$\begin{aligned} \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} \mu_r(U_\epsilon) &= \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} \int_{U_\epsilon} |\nabla_m f|_m \cdot h_\mu \omega_m^r \\ &= \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} \int_{-\epsilon}^\epsilon \left( \int_{f^{-1}\{\beta\}} h_\mu \cdot \omega_m^{r-1} \right) d\beta \\ &= \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} \int_0^\epsilon \left( \int_{\Gamma^\beta} h_\mu \cdot \omega_m^{r-1} \right) d\beta + \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} \int_{-\epsilon}^0 \left( \int_{\Gamma^\beta} h_\mu \cdot \omega_m^{r-1} \right) d\beta \\ &= \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} \int_{\beta_0}^\epsilon \mu_{r-1}(\Gamma^\beta) d\beta + \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} \int_{-\epsilon}^0 \mu_{r-1}(\Gamma^\beta) d\beta. \end{aligned} \quad (\text{B.71})$$

Take  $\epsilon$  to be smaller than the injectivity radius of each  $x \in \Gamma$ . Since  $h_\mu$  is continuous, by Lemma B.4.2 the function  $A(\beta) := \mu_{r-1}(\Gamma^\beta)$  is continuous on the intervals  $[-\epsilon, 0]$  and  $[0, \epsilon]$ . Thus, one can apply the fundamental theorem of calculus to both terms on the last line of (B.71) to obtain

$$\text{RHS of (B.71)} = \lim_{\epsilon \rightarrow 0} \frac{a(\epsilon) - a(0)}{\epsilon} + \lim_{\epsilon \rightarrow 0} \frac{a(0) - a(-\epsilon)}{\epsilon} = 2A(0) = 2\mu_{r-1}(\Gamma),$$

where  $a(\beta)$  is the anti-derivative of  $A(\beta)$ .  $\square$

Now, to obtain the inequality  $\mathbf{s}^{\text{dyn}} \leq \mathbf{H}_M^{\text{dyn}}$  via Lemma B.4.1, we start with the term  $\|\nabla_m f_\epsilon\|_{1,m,\mu}$  on the numerator of (B.66). Note that  $f_\epsilon$  is constant on  $M \setminus U_\epsilon$ , which implies  $\nabla_m f_\epsilon(x) = 0$  for all  $x \in M \setminus U_\epsilon$ . But if  $x \in U_\epsilon$ , then  $|\nabla_m f_\epsilon|_m = \frac{1}{\epsilon} |\nabla_m(\text{dist}_m(x, \Gamma))|_m = \frac{1}{\epsilon}$  for  $\epsilon$  smaller than as in Lemma B.4.2. Therefore, by Lemma B.4.3 one has

$$\lim_{\epsilon \rightarrow 0} \|\nabla_m f_\epsilon\|_{1,m,\mu} = \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} \int_{U_\epsilon} d\mu_r = \lim_{\epsilon \rightarrow 0} \frac{\mu_r(U_\epsilon)}{\epsilon} = 2\mu_{r-1}(\Gamma). \quad (\text{B.72})$$

Next, we consider the term  $\|\nabla_n \mathcal{H} f_\epsilon\|_{1,n,\nu}$  on the numerator of (B.66). Observe that at each point  $x \in M \setminus U_\epsilon$ ,

$$\begin{aligned} |\nabla_n \mathcal{H} f(Tx)|_n^2 &= n(\nabla_n \mathcal{H} f, \nabla_n \mathcal{H} f)_{Tx} \\ &= n(T_* \nabla_{T^* n} f, T_* \nabla_{T^* n} f)_{Tx} \quad \text{by Lemma B.2.6} \\ &= T^* n(\nabla_{T^* n} f, \nabla_{T^* n} f)_x \quad \text{by (3.7)} \\ &= m(\nabla_m f, \nabla_{T^* n} f)_x = 0, \end{aligned}$$

where we have used (3.16) and the fact that  $\nabla_m f(x) = 0$  for all  $x \in M \setminus U_\epsilon$  to obtain the last line. Hence, the integral  $\int_{N \setminus T\Omega_\epsilon} |\nabla_n \mathcal{H}f_\epsilon|_n d\nu_r$  vanishes. Set  $f$  to be the signed distance function defined by (B.68), then  $f(x) = \epsilon \cdot f_\epsilon(x)$  for all  $x \in U_\epsilon$ . Thus  $\mathcal{H}f = \epsilon \cdot \mathcal{H}f_\epsilon$  on  $TU_\epsilon$ . Let  $\Gamma^\beta$  be the level surfaces of the signed distance function  $f$ ; that is  $\Gamma^\beta = \{x \in M : f(x) = \beta\}$ . Then  $T\Gamma^\beta$  are generated by the level surfaces of  $\mathcal{H}f$ ; that is  $T\Gamma^\beta = \{y \in N : \mathcal{H}f(y) = \beta\}$ . Therefore, by the co-area formula (B.11) one has,

$$\begin{aligned}
 \|\nabla_n \mathcal{H}f_\epsilon\|_{1,n,\nu} &= \int_{TU_\epsilon} |\nabla_n \mathcal{H}f_\epsilon|_n d\nu_r \\
 &= \frac{1}{\epsilon} \int_{TU_\epsilon} |\nabla_n \mathcal{H}f|_n \cdot h_\nu \omega_n^r \\
 &= \frac{1}{\epsilon} \int_{-\epsilon}^\epsilon \left( \int_{(\mathcal{H}f)^{-1}\{t\}} h_\nu \cdot \omega_n^{r-1} \right) d\beta \\
 &= \frac{2}{\epsilon} \int_0^\epsilon \left( \int_{T\Gamma^\beta} h_\nu \cdot \omega_n^{r-1} \right) d\beta \\
 &= \frac{2}{\epsilon} \int_0^\epsilon \nu_{r-1}(T\Gamma^\beta) d\beta.
 \end{aligned} \tag{B.73}$$

By a straightforward modification of Lemma B.4.2, the expression  $\nu_{r-1}(T\Gamma^\beta)$  appearing on the RHS of (B.73) is continuous as a function of  $\beta$  on the interval  $[0, \epsilon]$ . Thus by taking the limit of  $\epsilon \rightarrow 0$  on both sides of (B.73) and using the fundamental theorem of calculus (see a similar argument used in Lemma B.4.3), one arrives at

$$\lim_{\epsilon \rightarrow 0} \|\nabla_n \mathcal{H}f_\epsilon\|_{1,n,\nu} = 2\nu_{r-1}(T\Gamma). \tag{B.74}$$

Finally, for the denominator term of (B.66), we may without loss of generality assume that  $\mu_r(M_1) \leq \mu_r(M_2)$ . Then

$$\begin{aligned}
 \|f_\epsilon - \alpha\|_{1,m,\mu} &\geq \int_{M \setminus U_\epsilon} |f_\epsilon - \alpha| d\mu_r \\
 &= |1 - \alpha| \cdot (\mu_r(M_1) - \mu_r(U_\epsilon)) + |1 + \alpha| \cdot (\mu_r(M_2) - \mu_r(U_\epsilon)) \\
 &\geq 2(\mu_r(M_1) - \mu_r(U_\epsilon)),
 \end{aligned} \tag{B.75}$$

for each  $\epsilon > 0$ . Hence, by taking the limit of  $\epsilon \rightarrow 0$  on (B.75) one has  $\inf_\alpha \|f_\epsilon - \alpha\|_{1,m,\mu} \geq 2\mu_r(M_1) = 2 \min\{\mu_r(M_1), \mu_r(M_2)\}$ . Combining this inequality with (B.66), (B.72) and (B.74), we conclude that  $\mathbf{s}^{dyn} \leq \mathbf{H}_M^{dyn}$ .

## B.5 The proof of Theorem 3.3.3

In this proof we follow the work of [49] and [88], and consider a weak formulation of the eigenvalue problem for the weighted dynamic Laplacian  $\Delta^{dyn}$ . One can find

a set of weak solution pairs  $(\phi_i, \lambda_i) \in L^2(M, m, \mu_r) \times \mathbb{R}$  to the weak formulation of  $\Delta^{dyn}$  that satisfies the conclusions of Theorem 3.3.3. Moreover, we show that the operator  $\Delta^{dyn}$  has the smooth uniformly elliptic property. That is,  $\Delta^{dyn}$  can be expressed in local coordinates as  $\Delta^{dyn} = \sum_{i,j=1}^r a_{ij} \partial_i \partial_j + b_i \partial_i + c$ , where  $a_{ij}$ ,  $b_i$  and  $c$  are bounded and smooth functions on  $M$ , and there exists a constant  $\gamma > 0$  such that  $\sum_{i,j=1}^r a_{ij} \varepsilon_i \varepsilon_j \geq \gamma |\varepsilon|^2$  for all  $x \in M$  and  $\varepsilon \in \mathbb{R}^r$ . The elliptic regularity theorem (see Theorem 8.14 in [61]) gives the additional regularity of the eigenfunctions  $\phi_i$  on  $M$ . Thus, the weak solution pairs  $(\phi_i, \lambda_i)$  solve the eigenproblem

$$\Delta^{dyn} \phi_i = \lambda_i \phi_i, \quad (\text{B.76})$$

for each  $i$ .

### B.5.1 Weak formulation of the $\Delta^{dyn}$ eigenproblem

Let  $f, g \in C^\infty(M, \mathbb{R})$ , and note that the smoothness assumption on the density  $h_\mu$  implies  $f, g \in L^2(M, m, \mu_r)$ . Consider the integral  $\int_M g \cdot \Delta^{dyn} f \cdot h_\mu \omega_m^r$ . Recall by (3.26) that the weighted divergence  $\text{div}_\mu$  satisfies

$$\int_{\partial U} m(\mathcal{V}, \mathbf{n}) \cdot h_\mu \omega_m^{r-1} = \int_U \text{div}_\mu \mathcal{V} \cdot h_\mu \omega_m^r, \quad (\text{B.77})$$

for all open  $U \subseteq M$  and continuously differentiable vector fields  $\mathcal{V} \in \mathcal{F}^1(M)$ , and where  $\mathbf{n}$  is the unit normal bundle along  $\partial U$ . Since  $f, g \in C^\infty(M, \mathbb{R})$ , the vector  $g \cdot \nabla_m f \in \mathcal{F}^\infty(M)$ . Consequently, by taking  $U = M$  and  $\mathcal{V} = g \cdot \nabla_m f$  in (B.77), follow by applying the expansion rule (B.17) for the weighted divergence  $\text{div}_\mu$ , one has the following weighted Green's identity:

$$\begin{aligned} \int_{\partial M} g \cdot m(\nabla_m f, \mathbf{n}) \cdot h_\mu \omega_m^{r-1} &= \int_M \text{div}_\mu (g \cdot \nabla_m f) \cdot h_\mu \omega_m^r \\ &= \int_M g \cdot \Delta_\mu f \cdot h_\mu \omega_m^r + \int_M m(\nabla_m g, \nabla_m f) \cdot h_\mu \omega_m^r. \end{aligned} \quad (\text{B.78})$$

Rearranging (B.78) gives

$$\int_M g \cdot \Delta_\mu f \cdot h_\mu \omega_m^r = - \int_M m(\nabla_m g, \nabla_m f) \cdot h_\mu \omega_m^r + \int_{\partial M} g \cdot m(\nabla_m f, \mathbf{n}) \cdot h_\mu \omega_m^{r-1}. \quad (\text{B.79})$$

Since  $\mathcal{H}$  is the adjoint of  $\mathcal{H}^*$

$$\int_M g \cdot \mathcal{H}^* \Delta_\nu \mathcal{H} f \, d\mu_r = \int_N \mathcal{H} g \cdot \Delta_\nu \mathcal{H} f \, d\nu_r.$$

Therefore, one has analogous to (B.79)

$$\int_M g \cdot \mathcal{H}^* \Delta_\nu \mathcal{H} f \cdot h_\mu \omega_m^r = - \int_N n(\nabla_n \mathcal{H} g, \nabla_n \mathcal{H} f) \cdot h_\nu \omega_n^r + \int_{\partial N} \mathcal{H} g \cdot n(\nabla_n \mathcal{H} f, \hat{\mathbf{n}}) \cdot h_\nu \omega_n^{r-1}, \quad (\text{B.80})$$

where  $\hat{\mathbf{n}}$  is the unit normal bundle along  $\partial N$ . Combining (B.79) and (B.80), we arrive at

$$\begin{aligned}
 & 2 \int_M g \cdot \Delta^{dym} f \cdot h_\mu \omega_m^r \\
 &= \int_M g \cdot (\Delta_\mu f + \mathcal{H}^* \Delta_\nu \mathcal{H} f) \cdot h_\mu \omega_m^r \\
 &= - \int_M m(\nabla_m f, \nabla_m g) \cdot h_\mu \omega_m^r - \int_N n(\nabla_n \mathcal{H} g, \nabla_n \mathcal{H} f) \cdot h_\nu \omega_n^r \\
 & \quad + P_1(f, g, \partial M) + P_2(f, g, \partial N),
 \end{aligned} \tag{B.81}$$

where

$$P_1(f, g, \partial M) = \int_{\partial M} g \cdot m(\nabla_m f, \mathbf{n}) \cdot h_\mu \omega_m^{r-1}, \tag{B.82}$$

and

$$P_2(f, g, \partial N) = \int_{\partial N} \mathcal{H} g \cdot n(\nabla_n \mathcal{H} f, \hat{\mathbf{n}}) \cdot h_\nu \omega_n^{r-1}. \tag{B.83}$$

Next, we demonstrate that if the boundary condition (3.35) in the hypothesis of Theorem 3.3.3 is satisfied for  $f$ , then the boundary term  $P_1(f, g, \partial M) + P_2(f, g, \partial N)$  of (B.81) vanishes for all  $g \in C^\infty(M, \mathbb{R})$ .

**Proposition B.5.1.** *Let  $f, g \in C^\infty(M, \mathbb{R})$ , and define  $P_1(f, g, \partial M)$  and  $P_2(f, g, \partial N)$  by (B.82) and (B.83), , where  $\partial M$  and  $\partial N$  are the boundary of  $M$  and  $N$  respectively. If the boundary condition*

$$m([\nabla_m + \nabla_{T^*n}]f, \mathbf{n})(x) = 0,$$

*holds for all  $x \in \partial M$ , then*

$$P_1(f, g, \partial M) + P_2(f, g, \partial N) = 0. \tag{B.84}$$

*Proof.* Let the hypersurface  $\partial M$  be generated by the zero level set of  $\psi \in C^\infty(M, \mathbb{R})$ ; i.e.  $\partial M = \{x \in M : \psi(x) = 0\}$ . Due to Proposition B.2.5, the surface  $\partial N$  is generated by the zero level set of  $\mathcal{H}\psi$ . Now by Lemma B.2.6 and the fact that  $\mathcal{H}^*\mathcal{H}$  is the identity,

$$\begin{aligned}
 n(\nabla_n \mathcal{H} f, \nabla_n \mathcal{H} \psi)_{Tx} &= n(T_* \nabla_{T^*n} f, T_* \nabla_{T^*n} \psi)_{Tx} \\
 &= T^* n(\nabla_{T^*n} f, \nabla_{T^*n} \psi)_x \quad \text{by (3.7)} \\
 &= (\nabla_{T^*n} f) \psi \Big|_x \quad \text{by (3.16)} \\
 &= m(\nabla_{T^*n} f, \nabla_m \psi)_x.
 \end{aligned} \tag{B.85}$$

Hence,

$$\begin{aligned}
 & \int_{-\infty}^{\infty} \int_{\mathcal{H}\psi=\beta} \mathcal{H}g \cdot \frac{n(\nabla_n \mathcal{H}f, \nabla_n \mathcal{H}\psi)}{|\nabla_n \mathcal{H}\psi|_n} \cdot h_\nu \omega_n^{r-1} d\beta \\
 &= \int_N \mathcal{H}g \cdot n(\nabla_n \mathcal{H}f, \nabla_n \mathcal{H}\psi) \cdot h_\nu \omega_n^r \quad \text{by the co-area formula (B.11)} \\
 &= \int_M g \cdot n(\nabla_n \mathcal{H}f, \nabla_n \mathcal{H}\psi) \circ T \cdot h_\mu \omega_m^r \quad \text{by (B.27)} \\
 &= \int_M g \cdot m(\nabla_{T^*n} f, \nabla_m \psi) \cdot h_\mu \omega_m^r \quad \text{by (B.85)} \\
 &= \int_{-\infty}^{\infty} \int_{\psi=\beta} g \cdot \frac{m(\nabla_{T^*n} f, \nabla_m \psi)}{|\nabla_m \psi|_m} \cdot h_\mu \omega_m^{r-1} d\beta, \tag{B.86}
 \end{aligned}$$

where the last line is due to the application of the co-area formula (B.11). Differentiating both sides of (B.86) with respect to  $\beta$ , then at  $\beta = 0$

$$\int_{\mathcal{H}\psi=0} \mathcal{H}g \cdot \frac{n(\nabla_n \mathcal{H}f, \nabla_n \mathcal{H}\psi)}{|\nabla_n \mathcal{H}\psi|_n} \cdot h_\nu \omega_n^{r-1} = \int_{\psi=0} g \cdot \frac{m(\nabla_{T^*n} f, \nabla_m \psi)}{|\nabla_m \psi|_m} \cdot h_\mu \omega_m^{r-1}. \tag{B.87}$$

Additionally, the vector  $\nabla_m \psi$  is normal to the level surfaces of  $\psi$ . Therefore  $\mathbf{n} = \nabla_m \psi / |\nabla_m \psi|_m$ , and similarly  $\hat{\mathbf{n}} = \nabla_n \mathcal{H}\psi / |\nabla_n \mathcal{H}\psi|_n$ . Hence,

$$\begin{aligned}
 P_2(f, g, \partial N) &= \int_{\partial N} \mathcal{H}g \cdot n(\nabla_n \mathcal{H}f, \hat{\mathbf{n}}) \cdot h_\nu \omega_n^{r-1} \\
 &= \int_{\mathcal{H}\psi=0} \mathcal{H}g \cdot \frac{n(\nabla_n \mathcal{H}f, \nabla_n \mathcal{H}\psi)}{|\nabla_n \mathcal{H}\psi|_n} \cdot h_\nu \omega_n^{r-1} \\
 &= \int_{\psi=0} g \cdot \frac{m(\nabla_{T^*n} f, \nabla_m \psi)}{|\nabla_m \psi|_m} \cdot h_\mu \omega_m^{r-1} \quad \text{by (B.87)} \\
 &= \int_{\partial M} g \cdot m(\nabla_{T^*n} f, \mathbf{n}) \cdot h_\mu \omega_m^{r-1}.
 \end{aligned}$$

We conclude that

$$P_1(f, g, \partial M) + P_2(f, g, \partial N) = \int_{\partial M} g \cdot m([\nabla_m + \nabla_{T^*n}]f, \mathbf{n}) \cdot h_\mu \omega_m^{r-1},$$

which vanishes due to the theorem hypothesis of  $m([\nabla_m + \nabla_{T^*n}]f, \mathbf{n})(x) = 0$  for all  $x \in \partial M$ .  $\square$

Due to Proposition B.5.1 and (B.81), one has

$$\int_M g \cdot \Delta^{dyn} f d\mu_r = - \int_M m(\nabla_m f, \nabla_m g) d\mu_r - \int_N n(\nabla_n \mathcal{H}g, \nabla_n \mathcal{H}f) d\mu_r, \tag{B.88}$$

for all  $f, g \in C^\infty(M, \mathbb{R})$ . Note that (B.88) is symmetric in  $f$  and  $g$ , hence the operator  $\Delta^{dyn}$  is self-adjoint in  $L^2(M, m, \mu_r)$ .

Suppose the solution  $(\phi, \lambda) \in C^\infty(M, \mathbb{R}) \times \mathbb{R}$  exists for the eigenvalue problem (B.76). Then under the boundary condition (3.35), one has by (B.81) and Proposition B.5.1 the following formulation for the eigenvalue problem  $\Delta^{dyn} \phi = \lambda \phi$ :

$$\int_M m(\nabla_m g, \nabla_m \phi) d\mu_r + \int_N n(\nabla_n \mathcal{H}g, \nabla_n \mathcal{H}\phi) d\nu_r = -2\lambda \int_M g \phi d\mu_r. \tag{B.89}$$

for all  $g \in C^\infty(M, \mathbb{R})$ . Equivalently,

$$\int_M [m(\nabla_m g, \nabla_m \phi) + T^*n(\nabla_{T^*n} g, \nabla_{T^*n} \phi)] d\mu_r = -2\lambda \int_M g\phi d\mu_r. \quad (\text{B.90})$$

for all  $g \in C^\infty(M, \mathbb{R})$ .

## B.5.2 Existence of weak solution and variational characterisation of eigenvalues

Let  $S$  be a weighted Sobolev space  $W^{1,2}(M, m, \mu_r)$ , where the density of  $\mu_r$  is  $h_\mu$ . Recall from Section B.3 that the weak gradient with respect to the metric  $m$  is denoted by  $\tilde{\nabla}_m$ . Due to (B.89), the the weak formulation for the eigenproblem (B.76) is given by

$$\int_M m(\tilde{\nabla}_m g, \tilde{\nabla}_m \phi) d\mu_r + \int_N n(\tilde{\nabla}_n \mathcal{H}g, \tilde{\nabla}_n \mathcal{H}\phi) d\nu_r = -2\lambda \int_M g\phi d\mu_r. \quad (\text{B.91})$$

We show existence of solutions  $(\phi_i, \lambda_i) \in S \times \mathbb{R}$  for the above weak formulation (B.91), for all  $g \in S$ . We call such pairs  $(\phi_i, \lambda_i)$  weak solutions<sup>1</sup> for the eigenvalue problem (B.76).

Our approach to finding the weak solutions for  $\Delta^{dyn}$  is based on the construction of functionals  $F$  and  $G$ , and using the method of Lagrange multipliers. For  $f \in S$ , we define  $G(f) = 1 - \int_M f^2 d\mu_r$  and  $F(f) = (1/2)(F_1(f) + F_2(f))$ , where  $F_1(f) = \int_M |\tilde{\nabla}_m f|_m^2 d\mu_r$  and  $F_2(f) = \int_N |\tilde{\nabla}_n \mathcal{H}f|_n^2 d\nu_r$ . First we list some useful properties of the functionals  $F_1$ ,  $F_2$  and  $G$ .

**Lemma B.5.2.** *Let  $f \in S$ , and denote the linear dual of  $S$  as  $S^*$ . Define the functional  $F_2$  as above.*

- (i) *The functional  $F_2 : S \rightarrow \mathbb{R}$  is well-defined,*
- (ii) *The derivative  $F_2'(f)$  is linear and bounded (hence  $F_2'(f) \in S^*$ ),*
- (iii)  *$F_2$  is Fréchet-differentiable,*
- (iv)  *$f \rightarrow F_2'(f)$  is continuous as a map from  $S$  to  $S^*$ .*

*Proof.* (i) Let  $(U_k, \varphi_k)_{k \in K}$  be an atlas of  $M$ . Then due to the fact that  $T$  is a  $C^\infty$ -diffeomorphism, there exists a set of finite constants  $C_{ij}^k$  such that  $(T^*n)^{ij} = C_{ij}^k m^{ij}$  on  $U_k$  for each  $1 \leq i, j \leq r$  and  $k \in K$ . Hence, by writing  $\tilde{\nabla}_{T^*n}$  in

---

<sup>1</sup>The weak solution pairs  $(\phi_i, \lambda_i)$  does not necessarily solve the eigenvalue problem  $\Delta^{dyn} \phi_i = \lambda_i \phi_i$ , because  $\phi_i$  may lack sufficient regularity on  $M$ , see p.210 in [88] for a discussion.

coordinates form via (B.12) (with respect to weak partial derivatives  $\tilde{\partial}$ ), one has on all points in  $U_k$

$$\tilde{\nabla}_{T^*n}f = \sum_{ij} (T^*n)^{ij} (\tilde{\partial}_i f) \partial_j = \sum_{ij} C_{ij}^k m^{ij} (\hat{\partial}_i f) \partial_j \leq C^k \cdot \tilde{\nabla}_m f, \quad (\text{B.92})$$

for all  $k \in K$ , where  $C_k = \max_{ij} C_{ij}^k$ . Furthermore, since  $M$  is compact, there exists a partition of unity  $\sigma_k$  subordinate to the covering  $\cup_{k \in K} U_k$  (Lemma B.2.9). Therefore,

$$\begin{aligned} F_2(f) &= \int_N |\tilde{\nabla}_n \mathcal{H}f|_n^2 d\nu_r \\ &= \int_N n(\tilde{\nabla}_n \mathcal{H}f, \tilde{\nabla}_n \mathcal{H}f) d\nu_r \\ &= \int_N n(T_* \tilde{\nabla}_{T^*n} f, T_* \tilde{\nabla}_{T^*n} f) d\nu_r \quad \text{by Lemma B.2.6} \\ &= \int_N T^* n(\tilde{\nabla}_{T^*n} f, \tilde{\nabla}_{T^*n} f) \circ T^{-1} d\nu_r \quad \text{by (3.7)} \\ &= \int_M T^* n(\tilde{\nabla}_{T^*n} f, \tilde{\nabla}_{T^*n} f) d\mu_r \quad \text{by (B.28)} \\ &= \int_M m(\tilde{\nabla}_{T^*n} f, \tilde{\nabla}_m f) d\mu_r \quad \text{by (3.16)} \\ &= \sum_{k \in K} \int_{U_k} \sigma_k \cdot m(\tilde{\nabla}_{T^*n} f, \tilde{\nabla}_m f) d\mu_r \\ &\leq \sum_{k \in K} C_k \int_{U_k} \sigma_k \cdot m(\tilde{\nabla}_m f, \tilde{\nabla}_m f) d\mu_r \quad \text{by (B.92)} \\ &\leq C \cdot \int_M |\tilde{\nabla}_m f|_m^2 d\mu_r = C \cdot F_1(f), \end{aligned} \quad (\text{B.93})$$

where  $C = \max_{k \in K} C_k$ . Since  $f \in S$ ,  $\tilde{\nabla}_m f \in L^2(M, m, \mu_r)$ . It follows that  $F_2 : S \rightarrow \mathbb{R}$  is well defined.

(ii) For all  $f, g \in S$

$$\begin{aligned} &F_2'(f)g \\ &= \lim_{\epsilon \rightarrow 0} \frac{F_2(f + \epsilon g) - F_2(f)}{\epsilon} \\ &= \lim_{\epsilon \rightarrow 0} \frac{\int_N |\tilde{\nabla}_n \mathcal{H}(f + \epsilon g)|_n^2 d\nu_r - \int_N |\tilde{\nabla}_n \mathcal{H}f|_n^2 d\nu_r}{\epsilon} \\ &= \lim_{\epsilon \rightarrow 0} \frac{\int_N \left( |\tilde{\nabla}_n \mathcal{H}f|_n^2 + 2\epsilon \cdot n(\tilde{\nabla}_n \mathcal{H}f, \tilde{\nabla}_n \mathcal{H}g) + \epsilon^2 \cdot |\tilde{\nabla}_n \mathcal{H}g|_n^2 - |\tilde{\nabla}_n \mathcal{H}f|_n^2 \right) d\nu_r}{\epsilon} \\ &= 2 \int_N n(\tilde{\nabla}_n \mathcal{H}g, \tilde{\nabla}_n \mathcal{H}f) d\nu_r, \end{aligned} \quad (\text{B.94})$$

where to obtain the last line, we have used the fact that the coefficient of the  $\epsilon^2$  term on the penultimate line is finite from part (i). Clearly  $F_2'(f)$  is linear.

Furthermore, by the Cauchy-Schwarz inequality, one has

$$\begin{aligned}
 \text{RHS of (B.94)} &\leq 2\|\nabla_n \mathcal{H}f\|_{2,n,\nu} \cdot \|\tilde{\nabla}_n \mathcal{H}g\|_{2,n,\nu} \\
 &\leq 2C \cdot \left( \int_M |\tilde{\nabla}_m f|_m^2 d\mu_r \right)^{\frac{1}{2}} \cdot \left( \int_M |\tilde{\nabla}_m g|_m^2 d\mu_r \right)^{\frac{1}{2}} \quad \text{by (B.93)} \\
 &\leq 2C \cdot \left( \int_M |\tilde{\nabla}_m f|_m^2 d\mu_r \right)^{\frac{1}{2}} \cdot \|g\|_S,
 \end{aligned}$$

where  $C$  is the same constant that appeared in part (i). Therefore,  $F'_2(f)$  is bounded.

By using the results of part (i) and (ii), the proof of (iii) and (iv) is similar to the corresponding results of Lemma C.1 in [49].  $\square$

*Remark B.5.3.* One may obtain analogous results of Lemma B.5.2 for  $F_1$  by setting  $T$  as the identity map in  $F_2$ , while the corresponding results for  $G$  is a straightforward modification with

$$G'(f)g = -2 \int_M fg d\mu_r. \quad (\text{B.95})$$

An important concept associated with linear functionals is the weak convergence. Let  $f_i$  be a sequence in  $S$ . We say that  $f_i \rightharpoonup f$  weakly in  $S$ , if  $H(f_i) \rightarrow H(f)$  for all  $H \in S^*$  (where  $S^*$  is the linear dual of  $S$ ). Moreover, since  $S$  is a Hilbert space (Proposition B.3.4), by the Riez representation theorem, if  $f_i \rightharpoonup f$  weakly in  $S$  then  $\langle g, f_i \rangle_S = \langle g, f \rangle_S$ , for all  $g \in S$ . One has the following standard result (see p.174, [88])

**Lemma B.5.4.** *Every bounded sequence in a Hilbert space contains a weakly convergent subsequence.*

Recall from Section B.3 that the  $A_p$  condition on the density  $h_\mu$  has important consequences for the weighted Sobolev space  $W^{1,2}(M, m, \mu_r)$ . By assumption, the density  $h_\mu$  is smooth and uniformly bounded away from zero. Hence, by Proposition B.1.2, the density  $h_\mu$  is an  $A_2$  weight on the the space  $S$ .

**Lemma B.5.5.**  *$F$  attains its minimum on the constraint set  $\mathcal{C} = \{f \in S : G(f) = 0\}$ .*

*Proof.* Define the inner product  $\langle h, g \rangle_{S'} = \int_N (n(\tilde{\nabla}_n \mathcal{H}g, \tilde{\nabla}_n \mathcal{H}h) + \mathcal{H}(gh)) d\nu_r$  for all  $g, h \in S$ , and denote the norm associated with  $\langle \cdot, \cdot \rangle_{S'}$  by  $\|\cdot\|_{S'}$ . Set  $I = \inf\{F(g) : g \in \mathcal{C}\} \geq 0$ , and select a sequence  $f_i \in \mathcal{C}$  such that  $F(f_i) \rightarrow I$  and  $F(f_i) \leq I + 1$ .

First, we show that the sequence  $f_i$  is bounded in both  $S$  and  $S'$ . Due to Lemma B.3.10, there exists a constant  $K$  (independent of  $f_i$ ) such that  $\|f_i - \alpha(f_i)\|_{2,\mu} \leq K\|\tilde{\nabla}_m f_i\|_{2,\mu}$  for each  $i$ . Hence,

$$\begin{aligned} \|f_i\|_S^2 &= \|f_i\|_{2,m,\mu}^2 + \|\tilde{\nabla}_m f_i\|_{2,m,\mu}^2 \\ &\leq \|f_i - \alpha(f_i)\|_{2,m,\mu}^2 + |\alpha(f_i)| + \|\tilde{\nabla}_m f_i\|_{2,m,\mu}^2 \quad \text{by triangle's inequality} \\ &\leq (1 + K^2)\|\tilde{\nabla}_m f_i\|_{2,m,\mu}^2 + |\alpha(f_i)|. \end{aligned}$$

Moreover, by Cauchy-Schwarz

$$|\alpha(f_i)|^2 = \left| \int_M f_i d\mu_r \right|^2 \leq \left( \int_M f_i^2 d\mu_r \right) = 1 - G(f_i) = 1.$$

Hence

$$\|f_i\|_S \leq (1 + K^2)\|\nabla_m f_i\|_{2,m,\mu} + 1 = (1 + K^2)F_1(f_i) + 1 \leq (1 + K^2)(I + 1) + 1,$$

so that  $f_i$  is a bounded sequence in  $S$ . By applying similar arguments as (B.93) in the proof of Lemma B.5.2(i), one can verify that  $f_i$  is also a bounded sequence in  $S'$ .

Since  $f_i$  is a bounded sequence in  $S$ , and  $S$  a Hilbert space (due to Proposition B.3.4), by Lemma B.5.4, there exists a subsequence  $f_{i_k}$  such that  $f_{i_k} \rightharpoonup f$  weakly in  $S$ . Moreover, due to Lemma B.3.9, the embedding  $S \hookrightarrow L^2(M, m, \mu_r)$  is compact, which implies the existence of a subsequence  $f_{i_k}$  of  $f_{i_k}$ , such that  $f_{i_k} \rightarrow f$  in  $L^2(M, m, \mu_r)$ . The strong convergence  $f_{i_k} \rightarrow f$  in  $L^2(M, m, \mu_r)$  implies  $\mathcal{H}f_{i_k} \rightarrow \mathcal{H}f$  in  $L^2(N, n, \nu_r)$ , because by the change of variable (B.27)

$$\|f_{i_k} - f\|_{2,m,\mu}^2 = \int_M |f_{i_k} - f|^2 d\mu_r = \int_N |\mathcal{H}f_{i_k} - \mathcal{H}f|^2 d\nu_r = \|\mathcal{H}f_{i_k} - \mathcal{H}f\|_{2,n,\nu}^2.$$

Next, we use the fact that the subsequence  $f_{i_k}$  is bounded in  $S'$  together with the weak convergence of  $f_{i_k}$  in  $S$ , to show that  $f_{i_k}$  converges weakly in  $S'$ . Due to lemma (B.5.2) and Remark B.5.3, one has  $F'_2(g) \in S^*$  and  $G'(g) \in S^*$  for all  $g \in S$ . Therefore

$$\begin{aligned} \lim_{i_k \rightarrow \infty} \langle f_{i_k}, g \rangle_{S'} &= \lim_{i_k \rightarrow \infty} \int_N n(\tilde{\nabla}_n \mathcal{H}f_{i_k}, \tilde{\nabla}_n \mathcal{H}g) d\nu_r + \lim_{i_k \rightarrow \infty} \int_N \mathcal{H}(f_{i_k}g) d\nu_r \\ &= \frac{1}{2} \lim_{i_k \rightarrow \infty} F'_2(g)f_{i_k} + \lim_{i_k \rightarrow \infty} \int_M f_{i_k}g d\mu_r \quad \text{by (B.94) and (B.27)} \\ &= \frac{1}{2} \lim_{i_k \rightarrow \infty} F'_2(g)f_{i_k} + \frac{1}{2} \lim_{i_k \rightarrow \infty} G'(g)f_{i_k} \quad \text{by (B.95)} \\ &= \frac{1}{2}(F'_2(g)f + G'(g)f) \\ &= \langle f, g \rangle_{S'}, \end{aligned}$$

where the penultimate line is due to the weak convergence of  $f_{i_k} \rightharpoonup f$  in  $S$ .

Now, the weak convergence of  $f_{i_k}$  in  $S'$  implies

$$\begin{aligned} \|f\|_{S'}^2 &= \langle f, f \rangle_{S'} = \lim_{i_k \rightarrow \infty} \langle f_{i_k}, f \rangle_{S'} \\ &= \lim_{i_k \rightarrow \infty} \left\{ \langle \tilde{\nabla}_n \mathcal{H} f_{i_k}, \tilde{\nabla}_n \mathcal{H} f \rangle_\nu + \langle \mathcal{H} f_{i_k}, \mathcal{H} f \rangle_\nu \right\} \\ &\leq \lim_{i_k \rightarrow \infty} \left\{ \|\tilde{\nabla}_n \mathcal{H} f_{i_k}\|_{2,n,\nu} \cdot \|\tilde{\nabla}_n \mathcal{H} f\|_{2,n,\nu} + \|\mathcal{H} f_{i_k}\|_{2,\nu} \cdot \|\mathcal{H} f\|_{2,n,\nu} \right\}, \end{aligned} \quad (\text{B.96})$$

where the inequality on the last line is due to Cauchy-Schwarz. In (B.96), set  $a_1 = \|\tilde{\nabla}_n \mathcal{H} f_{i_k}\|_{2,n,\nu}$ ,  $b_1 = \|\tilde{\nabla}_n \mathcal{H} f\|_{2,n,\nu}$ ,  $a_2 = \|\mathcal{H} f_{i_k}\|_{2,\nu}$  and  $b_2 = \|\mathcal{H} f\|_{2,n,\nu}$ , and consider the inequality

$$\begin{aligned} a_1 b_1 + a_2 b_2 &= \sqrt{a_1^2 b_1^2 + a_2^2 b_2^2 + 2a_1 b_2 a_2 b_1} \\ &\leq \sqrt{a_1^2 b_1^2 + a_2^2 b_2^2 + a_1^2 b_2^2 + a_2^2 b_1^2} \quad \text{since } 2cd \leq c^2 + d^2, \forall c, d \in \mathbb{R} \\ &= \sqrt{(a_1^2 + a_2^2)(b_1^2 + b_2^2)}. \end{aligned} \quad (\text{B.97})$$

As a consequence of (B.97), one has

$$\begin{aligned} &\text{RHS of (B.96)} \\ &\leq \lim_{i_k \rightarrow \infty} \left\{ \sqrt{\left( \|\tilde{\nabla}_n \mathcal{H} f_{i_k}\|_{2,n,\nu}^2 + \|\mathcal{H} f_{i_k}\|_{2,n,\nu}^2 \right) \cdot \left( \|\tilde{\nabla}_n \mathcal{H} f\|_{2,n,\nu}^2 + \|\mathcal{H} f\|_{2,n,\nu}^2 \right)} \right\} \\ &= \|f\|_{S'} \times \lim_{i_k \rightarrow \infty} \|f_{i_k}\|_{S'}. \end{aligned}$$

Thus,  $\|f\|_{S'} \leq \lim_{i_k \rightarrow \infty} \|f_{i_k}\|_{S'}$ . Furthermore, the subsequence  $f_{i_k}$  is bounded in  $S'$ , and  $\liminf_{i_k \rightarrow \infty} \|f_{i_k}\|_{S'}$  is the largest number smaller than  $\lim_{i_k \rightarrow \infty} \|f_{i_k}\|_{S'}$ . Thus,

$$\|f\|_{S'} \leq \lim_{i_k \rightarrow \infty} \|f_{i_k}\|_{S'} \implies \|f\|_{S'} \leq \liminf_{i_k \rightarrow \infty} \|f_{i_k}\|_{S'}. \quad (\text{B.98})$$

Similarly, the weak convergence of the bounded subsequence  $f_{i_k}$  in  $S$  gives

$$\|f\|_S \leq \lim_{i_k \rightarrow \infty} \|f_{i_k}\|_S \implies \|f\|_S \leq \liminf_{i_k \rightarrow \infty} \|f_{i_k}\|_S. \quad (\text{B.99})$$

Finally, due to (B.98) and (B.99)

$$\begin{aligned} 2F(f) &= \int_M |\tilde{\nabla}_m f|_m^2 d\mu_r + \int_N |\tilde{\nabla}_n \mathcal{H} f|_n^2 d\nu_r \\ &= \|f\|_S^2 - \|f\|_{2,m,\mu}^2 + \|f\|_{S'}^2 - \|\mathcal{H} f\|_{2,n,\nu}^2 \\ &\leq \liminf_{i_k \rightarrow \infty} \|f_{i_k}\|_S^2 - \|f\|_{2,m,\mu}^2 + \liminf_{i_k \rightarrow \infty} \|f_{i_k}\|_{S'}^2 - \|\mathcal{H} f\|_{2,n,\nu}^2 \\ &= \liminf_{i_k \rightarrow \infty} \{ \|f_{i_k}\|_S^2 - \|f\|_{2,m,\mu}^2 \} + \liminf_{i_k \rightarrow \infty} \{ \|f_{i_k}\|_{S'}^2 - \|\mathcal{H} f\|_{2,n,\nu}^2 \}. \end{aligned} \quad (\text{B.100})$$

By the strong convergence of  $f_{i_k} \rightarrow f$  in  $L^2(M, m, \mu_r)$ , and the strong convergence of  $\mathcal{H}f_{i_k} \rightarrow \mathcal{H}f$  in  $L^2(N, n, \nu_r)$ , one has

$$\begin{aligned}
 \text{RHS of (B.100)} &= \liminf_{i_k \rightarrow \infty} \{ \|f_{i_k}\|_S^2 - \|f_{i_k}\|_{2,m,\mu}^2 \} + \liminf_{i_k \rightarrow \infty} \{ \|f_{i_k}\|_{S'}^2 - \|\mathcal{H}f_{i_k}\|_{2,n,\nu}^2 \} \\
 &\leq \liminf_{i_k \rightarrow \infty} \{ \|f_{i_k}\|_S^2 - \|f_{i_k}\|_{2,m,\mu}^2 + \|f_{i_k}\|_{S'}^2 - \|\mathcal{H}f_{i_k}\|_{2,n,\nu}^2 \} \\
 &= \liminf_{i_k \rightarrow \infty} \{ \|\nabla_m f_{i_k}\|_{2,m,\mu}^2 + \|\nabla_n \mathcal{H}f_{i_k}\|_{2,n,\nu}^2 \} \\
 &= 2 \liminf_{i_k \rightarrow \infty} F(f_{i_k}) = 2I.
 \end{aligned} \tag{B.101}$$

From (B.100) and (B.101), we conclude that  $F(f) \leq I = \inf\{F(g) : g \in \mathcal{C}\}$ ; thus the minimum of  $F$  is attained by  $f$ . To complete the proof the theorem, it remains to show that  $f \in \mathcal{C}$ ; that is  $G(f) = 0$ . One has

$$\begin{aligned}
 G(f) &= 1 - \int_M f^2 d\mu_r \\
 &= 1 - \|f\|_{2,m,\mu}^2 \\
 &= 1 - \lim_{i_k \rightarrow \infty} \|f_{i_k}\|_{2,m,\mu}^2 \\
 &= \lim_{i_k \rightarrow \infty} G(f_{i_k}) = 0,
 \end{aligned}$$

since  $f_{i_k} \in \mathcal{C}$ . □

Due to Lemma B.5.2, the functionals  $F$  and  $G$  are continuously differentiable. In addition, by Lemma B.5.5 there exists a function  $\bar{f} \in S$  which minimises  $F$  over the constraint set  $\mathcal{C}$ . Therefore, using the method of Lagrange multipliers, one has the equation  $F'(\bar{f})g = \lambda G'(\bar{f})g$  for some  $\lambda \in \mathbb{R}$  and all  $g \in S$ . Expanding this equation with (B.94) and (B.95) yields

$$\int_M m(\tilde{\nabla}_m g, \tilde{\nabla}_m \bar{f}) d\mu_r + \int_N n(\tilde{\nabla}_n \mathcal{H}g, \tilde{\nabla}_n \mathcal{H}\bar{f}) d\nu_r = -2\lambda \int_M g \bar{f} d\mu_r, \tag{B.102}$$

for all  $g \in S$ ,  $\bar{f} \in \{f \in S : G(f) = 0\}$  and some  $\lambda \in \mathbb{R}$ . By comparing (B.91) and (B.102), one sees immediately that  $(\bar{f}, \lambda) \in \{f \in S : G(f) = 0\} \times \mathbb{R}$  is a solution pair for the weak formulation (B.91).

If we fix  $g$  to be  $\bar{f}$  in (B.102), then

$$\begin{aligned}
 2F(\bar{f}) &= \int_M |\tilde{\nabla}_m \bar{f}|_m^2 d\mu_r + \int_N |\tilde{\nabla}_n \mathcal{H}\bar{f}|_n^2 d\nu_r \\
 &= -2\lambda \int_M \bar{f}^2 d\mu_r \\
 &= -2\lambda(G(\bar{f}) + 1).
 \end{aligned} \tag{B.103}$$

Moreover, as a consequence of Lemma B.5.5,  $\bar{f}$  is minimising for  $F$ . Thus rearranging (B.103) yields

$$\begin{aligned}\lambda &= - \inf_{f \in S} \frac{F(f)}{G(f) + 1} \\ &= - \inf_{f \in S} \frac{\int_M |\tilde{\nabla}_m f|_m^2 d\mu_r + \int_N |\tilde{\nabla}_n \mathcal{H}f|_n^2 d\nu_r}{2 \int_M f^2 d\mu_r}.\end{aligned}\quad (\text{B.104})$$

Let the solution  $(\bar{f}, \lambda)$  to (B.91) be denoted by  $(\phi_2, \lambda_2)$ . To find other solution pairs to (B.91) of the form  $(\phi_i, \lambda_i)$ , one follows the standard induction arguments presented in [49] and p.212 in [88]: One constructs a sequence of decreasing, closed and  $L^2(M, m, \mu_r)$ -orthogonal subspaces of  $S$ ; that is for  $k \geq 1$ , a sequence of subspaces of  $S$  of the form  $S_k = \{f \in S : \int_M f \phi_i d\mu_r = 0, \text{ for } i = 1, 2, \dots, k\}$ , where  $\phi_1$  is constant. One then uses the fact that the solutions  $\phi_i$  and  $\phi_j$  are  $L^2(M, m, \mu_r)$ -orthogonal for  $\lambda_i \neq \lambda_j$  (this follows immediately from Lemma C.3. in [49]), and the fact that each  $S_k$  is complete (closed subspace of a Hilbert space), to apply the variational method on  $S^{k-1}$  to obtain

$$\lambda_k = - \inf_{f \in S^{k-1}} \frac{\int_M |\tilde{\nabla}_m f|_m^2 d\mu_r + \int_N |\tilde{\nabla}_n \mathcal{H}f|_n^2 d\nu_r}{2 \int_M f^2 d\mu_r}, \quad (\text{B.105})$$

for  $k = 2, 3, \dots$ . Note that  $(\phi_1, 0)$  is a solution pair to (B.91), thus  $\lambda_1 = 0$ . Additionally, the sequence  $\lambda_i$  is monotone decreasing and tends to  $-\infty$ , with the solution space finite for each  $i$  (Lemma C.4. in [49]).

Furthermore, using the identity  $\tilde{\nabla}_n = T_* \tilde{\nabla}_{T^*n} \mathcal{H}^*$  from Lemma B.2.6,

$$\begin{aligned}\int_N |\tilde{\nabla}_n \mathcal{H}f|_n^2 d\nu_r &= \int_N n(T_* \tilde{\nabla}_{T^*n} f, T_* \tilde{\nabla}_{T^*n} f) d\nu_r \\ &= \int_M T^* n(\tilde{\nabla}_{T^*n} f, \tilde{\nabla}_{T^*n} f) d\mu_r = \int_M |\tilde{\nabla}_{T^*n} f|_{T^*n}^2 d\mu_r,\end{aligned}$$

where the second equality is due to (3.7) and (B.28). Hence, one can write (B.105) as an integral of  $M$  as

$$\lambda_k = - \inf_{f \in S^{k-1}} \frac{\int_M \left( |\tilde{\nabla}_m f|_m^2 + |\tilde{\nabla}_{T^*n} f|_{T^*n}^2 \right) d\mu_r}{2 \int_M f^2 d\mu_r}. \quad (\text{B.106})$$

### B.5.3 Ellipticity and global regularity of weak solutions

To complete the proof of Theorem 3.3.3, it remains to verify that the eigenfunctions  $\phi_i$  of  $\Delta^{dyn}$  are smooth and unique for each  $i$ . For then, the smoothness of  $\phi_i$  on  $M$  implies that the weak solution pairs  $(\phi_i, \lambda_i)$  which solves (B.91) are also solution to (B.89). Moreover, the uniqueness of  $(\phi_i, \lambda_i)$  implies that the solutions of the eigenvalue problem (B.76) are given by (B.105) or (B.106) (with the weak

gradients replaced with standard version due to the additional smoothness of  $\phi_i$ ). To determine the regularity and uniqueness of  $\phi_i$  on  $M$ , we utilise the elliptical regularity theorem (see Theorem 8.14 in [61]).

We say that an operator  $L$  of the form

$$L = \sum_{i,j=1}^r a_{ij} \frac{\partial^2}{\partial x^i \partial x^j} + b_i \frac{\partial}{\partial x^i} + c, \quad (\text{B.107})$$

is strictly uniformly elliptic if  $a_{ij}$ ,  $b_i$  and  $c$  are bounded, real-valued functions on  $M$ , and there exists a constant  $\gamma > 0$  such that

$$\sum_{i,j=1}^r a_{ij} \varepsilon_i \varepsilon_j \geq \gamma |\varepsilon|^2, \quad (\text{B.108})$$

where  $\varepsilon \in \mathbb{R}^r$  is non-zero.

As a consequence of the Elliptical Regularity theorem, if  $\partial M$  is smooth, and  $\Delta^{dyn}$  is a strictly uniformly elliptic operator with  $a_{ij}, b_i, c \in C^\infty(M, \mathbb{R})$  and  $c \leq 0$  in  $M$ , then there exist unique solutions in  $C^\infty(M, \mathbb{R})$  for the eigenproblem (B.76).

**Lemma B.5.6.** *Let  $T : M \rightarrow N$  be a  $C^\infty$ -diffeomorphism, and assume  $h_\mu$  is smooth and uniformly bounded away from zero. The weighted Laplacian  $\Delta^{dyn}$  is a strictly uniformly elliptic operator of the form (B.107), with  $a_{ij}, b_i, c$  in  $C^\infty(M, \mathbb{R})$  and  $c \leq 0$  on  $M$ .*

*Proof.* For this proof, we say that an operator has property  $E$ , if it is a strictly uniformly elliptic, with coefficients  $a_{ij}, b_i, c$  in  $C^\infty(M, \mathbb{R})$  and  $c \leq 0$  on  $M$ . By Lemma B.2.7

$$2\Delta^{dyn} f = \Delta_m f + \mathcal{H}^* \Delta_n \mathcal{H} f + \frac{m(\nabla_m h_\mu, \nabla_m f)}{h_\mu} + \frac{n(\nabla_n h_\nu, \nabla_n \mathcal{H} f) \circ T}{h_\nu \circ T}. \quad (\text{B.109})$$

Clearly the sum of operators with property  $E$  is an operator with property  $E$ . Additionally, if the second and fourth terms of (B.109) has property  $E$ , then by setting  $T$  as the identity, one immediately see that the first and third terms of (B.110) also has property  $E$ . Thus, it is sufficient to show that the second and fourth terms of (B.109) has property  $E$ . To show that second term  $\mathcal{H}^* \Delta_n \mathcal{H}$  of (B.109) has property  $E$ , we note by Corollary B.2.6 that  $\mathcal{H}^* \Delta_n \mathcal{H} = \Delta_{T^*n}$ . Therefore in local coordinates at any point in  $M$ ,

$$\mathcal{H}^* \Delta_n \mathcal{H} f = \Delta_{T^*n} f = \frac{1}{\sqrt{\det G_{T^*n}}} \sum_{i,j=1}^r \partial_j \left( \sqrt{\det G_{T^*n}} (T^*n)^{ij} \partial_i f \right), \quad (\text{B.110})$$

for all  $f \in C^\infty(M, \mathbb{R})$ . Using Jacobi's formula for differentiating the determinant of a matrix  $A$ ; that is  $\partial_k(\det A)(x) = (\det A)(x) \sum_{ij} (A^{-1})_{ij}(x) \partial_k A_{ij}(x)$  for all  $x \in M$ ,

one has

$$\begin{aligned}
 \partial_j(\sqrt{\det G_{T^*n}}) &= \frac{1}{2} \frac{1}{\sqrt{\det G_{T^*n}}} \partial_j(\det G_{T^*n}) \\
 &= \frac{1}{2} \frac{\det G_{T^*n}}{\sqrt{\det G_{T^*n}}} \sum_{k,l=1}^r (G_{T^*n}^{-1})_{kl} \partial_j(G_{T^*n})_{kl} \\
 &= \frac{1}{2} \sqrt{\det G_{T^*n}} \sum_{k,l=1}^r (T^*n)^{kl} \partial_j(T^*n)_{kl}. \tag{B.111}
 \end{aligned}$$

Therefore, by using the product rule to expand the partial derivative in the summation on the RHS of (B.110), and then applying (B.111) to the first term one has

$$\begin{aligned}
 &\text{RHS of (B.110)} \\
 &= \frac{1}{\sqrt{\det G_{T^*n}}} \left( \sum_{i,j=1}^r (T^*n)^{ij} \partial_j(\sqrt{\det G_{T^*n}}) \partial_i f + \sqrt{\det G_{T^*n}} \partial_j(T^*n)^{ij} \partial_i f \right. \\
 &\quad \left. + \sqrt{\det G_{T^*n}} (T^*n)^{ij} \partial_j \partial_i f \right) \\
 &= \sum_{i,j=1}^r \frac{1}{2} \left( \sum_{k,l=1}^r (T^*n)^{kl} \partial_j(T^*n)_{kl} \right) (T^*n)^{ij} \partial_i f + \partial_j(T^*n)^{ij} \partial_i f + (T^*n)^{ij} \partial_j \partial_i f \\
 &= \sum_{i,j=1}^r \left[ \frac{1}{2} \left( \sum_{k,l=1}^r (T^*n)^{kl} \partial_j(T^*n)_{kl} \right) (T^*n)^{ij} + \partial_j(T^*n)^{ij} \right] \partial_i f + (T^*n)^{ij} \partial_j \partial_i f. \tag{B.112}
 \end{aligned}$$

Now the Riemannian metric  $n$  is a  $C^\infty$  bilinear symmetric form and positive-definite for every  $y \in N$ . Moreover, the mapping  $T$  is a  $C^\infty$ -diffeomorphism. Hence, the components  $(T^*n)^{ij}$  and  $\partial_i(T^*n)^{ij}$  are both bounded and smooth for each  $1 \leq i, j \leq r$ . Therefore, the coefficients  $b_i = \sum_j \frac{1}{2} (T^*n)^{ij} \partial_j(T^*n)_{kl} + \partial_j(T^*n)^{ij}$  and  $a_{ij} = (T^*n)^{ij}$  in (B.112) are both bounded and smooth. Additionally, due to Lemma B.2.2 we have at the point  $x \in M$ ,

$$\begin{aligned}
 \sum_{i,j=1}^r a_{ij} \varepsilon_i \varepsilon_j &= \sum_{i,j=1}^r (T^*n)^{ij} \varepsilon_i \varepsilon_j \\
 &= \sum_{i,j=1}^r (J_T^\top \cdot G_n \circ T \cdot J_T)^{ij} \varepsilon_i \varepsilon_j \\
 &= \sum_{i,j=1}^r (J_T^{-1} \cdot G_n^{-1} \circ T \cdot (J_T^\top)^{-1})_{ij} \varepsilon_i \varepsilon_j \\
 &= \sum_{i,j,k,l=1}^r (J_T^{-1})_{ik} \cdot (G_n^{-1} \circ T)_{kl} \cdot (J_T^{-1})_{jl} \varepsilon_i \varepsilon_j \\
 &= \sum_{i,j,k,l=1}^r (J_{T^{-1}} \circ T)_{ik} \cdot (G_n^{-1} \circ T)_{kl} \cdot (J_{T^{-1}} \circ T)_{jl} \varepsilon_i \varepsilon_j, \tag{B.113}
 \end{aligned}$$

where the last line is due to the inverse function theorem. Therefore

$$\text{RHS of (B.113)} = \sum_{k,l=1}^r (J_{T^{-1}} \circ T \cdot \varepsilon)_k \cdot (G_n^{-1} \circ T)_{kl} \cdot (J_{T^{-1}} \circ T \cdot \varepsilon)_l > 0,$$

where we have used the fact that the matrix  $G_n^{-1}$  is positive definite at every  $Tx \in N$  to obtain the last inequality. Hence, there is a  $\gamma > 0$  such that  $\sum_{i,j=1}^r a_{ij}(x)\varepsilon_i\varepsilon_j \geq \gamma|\varepsilon|^2$  for all  $x \in M$ . Thus  $a_{ij}$  satisfies the condition (B.108), so by (B.110)-(B.112) the term  $\mathcal{H}^*\Delta_n\mathcal{H}$  has property  $E$ .

To show that the fourth term  $n(\nabla_n h_\nu, \nabla_n \mathcal{H}f)/\mathcal{H}^*h_\nu$  of (B.109) has property  $E$ , we consider the numerator term. One has at each point  $Tx \in N$ ,

$$\begin{aligned} n(\nabla_n h_\nu, \nabla_n \mathcal{H}f) &= n(\nabla_n h_\nu, T_*\nabla_{T^*n}f) \quad \text{by Lemma B.2.6} \\ &= (T_*\nabla_{T^*n}f)h_\nu \\ &= \nabla_{T^*n}f(h_\nu \circ T) \circ T^{-1} \quad \text{by (3.5)} \\ &= m(\nabla_m(h_\nu \circ T), \nabla_{T^*n}f) \circ T^{-1}. \end{aligned} \tag{B.114}$$

Writing the RHS of (B.114) in local coordinates, one has at any point  $x \in M$

$$\begin{aligned} \text{RHS of (B.114)} &= \sum_{i,j=1}^r m_{ij} \left( \sum_{k=1}^r m^{ki} \partial_k (h_\nu \circ T) \right) \left( \sum_{l=1}^r (T^*n)^{jl} \partial_l f \right) \\ &= \sum_{j=1}^r \partial_j (h_\nu \circ T) \left( \sum_{l=1}^r (T^*n)^{jl} \partial_l f \right) \quad \text{on contracting the index } i \\ &= \sum_{j,l=1}^r \partial_j (h_\nu \circ T) (T^*n)^{jl} \partial_l f. \end{aligned}$$

Therefore, at each  $x \in M$

$$\begin{aligned} \frac{n(\nabla_n h_\nu, \nabla_n \mathcal{H}f) \circ T}{h_\nu \circ T} &= \frac{\sum_{j,l=1}^r \partial_j (h_\nu \circ T) (T^*n)^{jl} \partial_l f}{h_\nu \circ T} \\ &= \sum_{j,l=1}^r \partial_j (\ln(h_\nu \circ T)) (T^*n)^{jl} \partial_l f \end{aligned}$$

As before, due to the properties of the metric  $m$ , the smoothness of  $h_\mu$ , and the fact that  $T$  is a diffeomorphism, the coefficient  $b_l = \sum_j \partial_j (\ln(h_\nu \circ T)) (T^*n)^{jl}$  is bounded and smooth, and so the fourth term of (B.109) has property  $E$ .  $\square$

## B.6 The proof of Theorem 3.3.4

This proof is a straightforward modification of Theorem 3.2 in [49]. Let  $g : M \rightarrow \mathbb{R}^+$  be nonnegative and smooth. Since

$$\int_N \mathcal{H}g \, d\nu_r = \int_M g \, d\mu_r,$$

by (B.27), and densities  $h_\mu, h_\nu$  are both positive and smooth, the function  $\mathcal{H}g$  is also nonnegative and smooth. Denote by  $\Gamma^\beta$  the level surfaces generated by  $g$ ; that is  $\{x \in M : g(x) = \beta\}$ . Then the level surfaces of  $T\Gamma^\beta$  are generated by  $\mathcal{H}g$ . Now, due to the co-area formula given by Lemma B.2.1, one has

$$\begin{aligned}
 & \int_M |\nabla_m g|_m \cdot h_\mu \omega_m^r + \int_N |\nabla_n \mathcal{H}g|_n \cdot h_\nu \omega_n^r \\
 &= \int_0^\infty \left( \int_{\Gamma^\beta} h_\mu \omega_m^{r-1} + \int_{T\Gamma^\beta} h_\nu \omega_n^{r-1} \right) d\beta \\
 &= \int_0^\infty (\mu_{r-1}(\Gamma^\beta) + \nu_{r-1}(T\Gamma^\beta)) d\beta \\
 &\geq 2 \inf_{t \in (0, \infty)} \mathbf{H}^{dyn}(\{g = \beta\}) \int_0^\infty \min\{\mu_r(g > \beta), \mu_r(g < \beta)\} d\beta. \tag{B.115}
 \end{aligned}$$

Let  $f : M \rightarrow \mathbb{R}$  be smooth, and  $\sigma$  the median of  $f$  with respect to  $\mu_r$ ; i.e.  $\mu_r(f \geq \sigma) \geq 1/2$  and  $\mu_r(f \leq \sigma) \geq 1/2$ . Set  $f_+ = \max\{f - \sigma, 0\}$  and  $f_- = -\min\{f - \sigma, 0\}$ , so that  $f - \sigma = f_+ - f_-$ . Observe that for each point  $x \in M$ , either  $|f(x) - \sigma| = f_+(x)$ ,  $|f(x) - \sigma| = f_-(x)$  or  $|f(x) - \sigma| = f_+(x) = f_-(x) = 0$ . Therefore

$$\inf_{\beta \in \mathbb{R}} \mathbf{H}^{dyn}(\{f = \beta\}) = \min \left\{ \inf_{t \in (0, \infty)} \mathbf{H}^{dyn}(\{f_-^2 = \beta\}), \inf_{t \in (0, \infty)} \mathbf{H}^{dyn}(\{f_+^2 = \beta\}) \right\}. \tag{B.116}$$

In addition, if  $f_+$  is positive then  $f > \sigma$ , and if  $f_-$  is positive then  $f < \sigma$ . Hence, by using the fact that  $\sigma$  is the median of  $f$ , one has

$$\mu_r(f_+^2 > \beta) \leq \frac{1}{2} \quad \text{and} \quad \mu_r(f_-^2 > \beta) \leq \frac{1}{2}, \tag{B.117}$$

for all  $\beta \geq 0$ . Moreover, if  $f_+ \neq 0$  then  $f_- = 0$ , and if  $f_- \neq 0$  then  $f_+ = 0$ . Hence,

$$(f - \sigma)^2 = f_+^2 + f_-^2, \tag{B.118}$$

and

$$\begin{aligned}
 |\nabla_m(f_+^2 + f_-^2)|_m^2 &= m(\nabla_m(f_+^2 + f_-^2), \nabla_m(f_+^2 + f_-^2)) \\
 &= |\nabla_m(f_+^2)|_m^2 + 2m(\nabla_m(f_+^2), \nabla_m(f_-^2)) + |\nabla_m(f_-^2)|_m^2 \\
 &= |\nabla_m(f_+^2)|_m^2 + |\nabla_m(f_-^2)|_m^2 \\
 &= |\nabla_m(f_+^2)|_m^2 + 2 \cdot |\nabla_m(f_+^2)|_m \cdot |\nabla_m(f_-^2)|_m + |\nabla_m(f_-^2)|_m^2 \\
 &= (|\nabla_m(f_+^2)|_m + |\nabla_m(f_-^2)|_m)^2. \tag{B.119}
 \end{aligned}$$

Finally, by definition  $\mathcal{H}f_+ = \max\{\mathcal{H}f - \sigma, 0\}$  and  $\mathcal{H}f_- = -\min\{\mathcal{H}f - \sigma, 0\}$ . Hence analogous to (B.118)

$$(\mathcal{H}f - \sigma)^2 = \mathcal{H}f_+^2 + \mathcal{H}f_-^2 \tag{B.120}$$

and analogous to (B.119)

$$|\nabla_n(\mathcal{H}f_+^2 + \mathcal{H}f_-^2)|_n^2 = (|\nabla_n(\mathcal{H}f_+^2) + \nabla_n(\mathcal{H}f_-^2)|_n)^2. \quad (\text{B.121})$$

Due to (B.118)-(B.121), one has

$$\begin{aligned} & \int_M |\nabla_m[(f - \sigma)^2]|_m d\mu_r + \int_M |\nabla_n[(\mathcal{H}f - \sigma)^2]|_n d\nu_r \\ &= \int_M |\nabla_m(f_+^2 + f_-^2)|_m d\mu_r + \int_N |\nabla_n(\mathcal{H}f_+^2 + \mathcal{H}f_-^2)|_n d\nu_r \quad \text{by (B.118) and (B.119)} \\ &= \int_M (|\nabla_m(f_+^2)|_m + |\nabla_m(f_-^2)|_m) d\mu_r + \int_N (|\nabla_n(\mathcal{H}f_+^2)|_n + |\nabla_n(\mathcal{H}f_-^2)|_n) d\nu_r, \end{aligned} \quad (\text{B.122})$$

where the last line is due to (B.119) and (B.121).

Now, consider the RHS of (B.122). Since  $f_+^2$  and  $f_-^2$  are nonnegative and smooth almost everywhere, one can set  $g = f_+^2$  and  $g = f_-^2$  independently in (B.115), and then apply (B.116) to the result to obtain

$$\begin{aligned} & \text{RHS of (B.122)} \\ & \geq 2 \inf_{\beta \in \mathbb{R}} \mathbf{H}^{dyn}(\{f = \beta\}) \int_0^\infty \left( \min\{\mu_r(f_+^2 > t), \mu_r(f_+^2 < \beta)\} \right. \\ & \quad \left. + \min\{\mu_r(f_-^2 > t), \mu_r(f_-^2 < \beta)\} \right) d\beta \\ & = 2 \inf_{\beta \in \mathbb{R}} \mathbf{H}^{dyn}(\{f = \beta\}) \int_0^\infty \mu_r(f_+^2 > \beta) + \mu_r(f_-^2 > \beta) d\beta, \end{aligned} \quad (\text{B.123})$$

where the equality on the last line is due to (B.117). Applying the Cavalieri's principle (Proposition I.3.3 in [23]) to the RHS of (B.123) yields

$$\begin{aligned} \text{RHS of (B.123)} &= 2 \inf_{\beta \in \mathbb{R}} \mathbf{H}^{dyn}(\{f = \beta\}) \int_M (f_+^2 + f_-^2) d\mu_r \\ &= 2 \inf_{\beta \in \mathbb{R}} \mathbf{H}^{dyn}(\{f = \beta\}) \int_M (f - \sigma)^2 d\mu_r. \end{aligned} \quad (\text{B.124})$$

Next, we consider the LHS of (B.122). In local coordinates, one has by (B.12)

$$\begin{aligned} \nabla_m[(f - \sigma)^2] &= \sum_{i,j=1}^r m^{ij} \partial_i (f - \sigma)^2 \partial_j \\ &= 2 \sum_{i,j=1}^r m^{ij} (f - \sigma) \partial_i f \partial_j \\ &= 2(f - \sigma) \nabla_m f. \end{aligned}$$

Therefore, by Cauchy-Schwarz

$$\begin{aligned} \int_M |\nabla_m[(f - \sigma)^2]|_m d\mu_r &= 2 \int_M |f - \sigma| \cdot |\nabla_m f|_m d\mu_r \\ &\leq 2 \|f - \sigma\|_{2,m,\mu} \cdot \|\nabla_m f\|_{2,m,\mu}. \end{aligned} \quad (\text{B.125})$$

Also, analogous to (B.125)

$$\begin{aligned}
 \int_N |\nabla_n [(\mathcal{H}f - \sigma)^2]|_n d\nu_r &\leq 2\|\mathcal{H}f - \sigma\|_{2,n,\nu} \cdot \|\nabla_n \mathcal{H}f\|_{2,n,\nu} \\
 &= 2 \left( \int_N (\mathcal{H}f - \sigma)^2 d\nu_r \right) \cdot \|\nabla_n \mathcal{H}f\|_{2,n,\nu} \\
 &= 2 \left( \int_M (f - \sigma)^2 d\mu_r \right) \cdot \|\nabla_n \mathcal{H}f\|_{2,n,\nu} \quad \text{by (B.27)} \\
 &= 2\|f - \sigma\|_{2,m,\mu} \cdot \|\nabla_n \mathcal{H}f\|_{2,n,\nu}. \tag{B.126}
 \end{aligned}$$

Therefore, by (B.122)-(B.126) one has

$$\begin{aligned}
 \inf_{\beta \in \mathbb{R}} \mathbf{H}^{dyn}(\{f = \beta\}) \int_M (f - \sigma)^2 d\mu_r &\leq \|f - \sigma\|_{2,m,\mu} \cdot (\|\nabla_m f\|_{2,m,\mu} + \|\nabla_n \mathcal{H}f\|_{2,n,\nu}) \\
 \implies \inf_{\beta \in \mathbb{R}} \mathbf{H}^{dyn}(\{f = \beta\}) &\leq \frac{\|\nabla_m f\|_{2,m,\mu} + \|\nabla_n \mathcal{H}f\|_{2,n,\nu}}{(\int_M (f - \sigma)^2 d\mu_r)^{1/2}}. \tag{B.127}
 \end{aligned}$$

Let  $\alpha(f)$  be the mean of  $f$  with respect to  $\mu_r$ ; that is  $\alpha(f) = \int_M f d\mu_r$ . Then  $\int_M (f - c)^2 d\mu_r$  as a function of  $c \in \mathbb{R}$  is minimum when  $c = \alpha(f)$ . Hence, by squaring both sides of (B.127), one has

$$\begin{aligned}
 \left( \inf_{\beta \in \mathbb{R}} \mathbf{H}^{dyn}(\{f = \beta\}) \right)^2 &\leq \frac{(\|\nabla_m f\|_{2,m,\mu} + \|\nabla_n \mathcal{H}f\|_{2,n,\nu})^2}{\int_M (f - \alpha(f))^2 d\mu_r} \\
 &\leq 2 \frac{\int_M |\nabla_m f|_m^2 d\mu_r + \int_N |\nabla_n \mathcal{H}f|_n^2 d\nu_r}{\int_M (f - \alpha(f))^2 d\mu_r}, \tag{B.128}
 \end{aligned}$$

for all  $f \in C^\infty(M, \mathbb{R})$ , where we have used the fact that  $(a + b)^2 \leq 2(a^2 + b^2)$  for  $a, b \in \mathbb{R}$  to obtain the inequality on the last line. Furthermore, if  $\lambda_2$  is the smallest magnitude nonzero eigenvalue of  $\Delta^{dyn}$  with corresponding eigenfunction  $\phi_2$ , then by Theorem 3.3.3, one has  $\phi_2 \in C^\infty(M, \mathbb{R})$ ,  $\alpha(\phi_2) = \int_M \phi_2 d\mu_r = 0$ , and for  $k = 2$  the infimum of (3.34) is attained by  $\phi_2$ . Thus, by setting  $f = \phi_2$  in (B.128),

$$\begin{aligned}
 \left( \inf_{\beta \in \mathbb{R}} \mathbf{H}^{dyn}(\{\phi_2 = \beta\}) \right)^2 &\leq 2 \frac{\int_M |\nabla_m \phi_2|_m^2 d\mu_r + \int_N |\nabla_n \mathcal{H}\phi_2|_n^2 d\nu_r}{\int_M |\phi_2 - \alpha(\phi_2)|^2 d\mu_r} \\
 &= -4\lambda_2.
 \end{aligned}$$

This concludes the proof of the theorem.

### B.6.1 Time-discrete and time-continuous case

To generalise Theorem 3.3.4 to the time-continuous dynamic Cheeger inequality, we note that apart from (B.128) all arguments are applied linearly with respect to time. Hence, the results up to (B.128) are immediate via the constructions outlined

in Sections 3.2.1 and 3.3.2. To modify the argument  $(a + b)^2 \leq 2(a^2 + b^2)$  used to obtain (B.128), we apply Cauchy-Schwarz to obtain

$$\left( \int_0^\tau a_t dt \right)^2 = \left( \int_0^\tau a_t \cdot 1 dt \right)^2 \leq \left( \int_0^\tau a_t^2 dt \right) \cdot \left( \int_0^\tau 1^2 dt \right) = \tau \cdot \int_0^\tau a_t^2 dt.$$

For the time-discrete case, one applies Cauchy-Schwarz analogously.

## B.7 The proof of Theorem 3.4.1

Recall the definition of the diffusion operator  $\mathcal{D}_{X,\epsilon}$  given by (3.41). Given  $f \in C^3(M, \mathbb{R})$ , we wish to evaluate the  $\epsilon \rightarrow 0$  limit of the image of  $f$  under the operator  $\mathcal{H}_\epsilon^* \mathcal{H}_\epsilon$ , where by (3.43) and (3.44),

$$\mathcal{H}_\epsilon^* \mathcal{H}_\epsilon f = \mathcal{D}_{X_\epsilon, \epsilon}^* \circ \mathcal{H}^* \circ \mathcal{D}_{Y'_\epsilon, \epsilon}^* \left( \frac{\mathcal{P}_\epsilon(fh_\mu)}{\mathcal{P}_\epsilon h_\mu} \right), \quad (\text{B.129})$$

with  $\mathcal{P}_\epsilon = \mathcal{D}_{Y'_\epsilon, \epsilon} \circ \mathcal{P} \circ \mathcal{D}_{X, \epsilon}$ . Let  $(U, \varphi)$  be a chart on  $M$  containing the point  $x \in M$ . Recall normal coordinates at the point  $x$ , are the local coordinates on  $(U, \varphi)$  such that the metric tensor satisfies  $m_{ij}(x) = \delta_{ij}$  and  $\partial_i m_{jk}(x) = 0$  for all  $1 \leq i, j, k \leq r$ .

Introducing standard multi-index notation for  $\alpha$ ; i.e.  $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_r)$  such that

$$\begin{aligned} |\alpha| &= \alpha_1 + \dots + \alpha_r \\ \alpha! &= \alpha_1! \alpha_2! \dots \alpha_r! \\ D^\alpha &= \partial_1^{\alpha_1} \partial_2^{\alpha_2} \dots \partial_r^{\alpha_r} \\ \mathbf{v}^\alpha &= v_1^{\alpha_1} v_2^{\alpha_2} \dots v_r^{\alpha_r}, \end{aligned} \quad (\text{B.130})$$

for a vector  $\mathbf{v} = (v_1, \dots, v_r)$ . The following lemmas are well known results regarding normal coordinates

**Lemma B.7.1.** *Let  $(U, \varphi)$  be a chart of  $M$  containing the point  $x_0 \in M$ , with corresponding normal coordinates  $\{x_1, x_2, \dots, x_r\}$ . The Laplace-Beltrami operator satisfies*

$$\Delta_m f(x_0) = \sum_{i=1}^r \frac{\partial^2 (f \circ \varphi^{-1})}{\partial x_i^2} (\varphi(x_0)).$$

*Proof.* See p.90 in [103]. □

**Lemma B.7.2.** *Let  $(U, \varphi)$  be a chart of  $M$  containing the point  $x_0 \in M$  with corresponding coordinates  $\{x_1, x_2, \dots, x_r\}$ . The asymptotic expansion of  $\sqrt{\det G_m}$  about  $B_\epsilon(x_0) \subseteq U$ , centered at  $x_0$  is given by*

$$\sqrt{\det G_m}(x_0) = 1 + \sum_{|\alpha|=2}^{\infty} C_{\mathcal{R}, |\alpha|}(x_0) \cdot (\varphi(x_0))^\alpha,$$

where  $C_{\mathcal{R},|\alpha|}(x_0)$  depend only on the Riemannian curvature tensor  $\mathcal{R}$  and covariant derivatives of  $\mathcal{R}$  at the point  $x_0$ . Moreover, if  $\mathcal{R}$  is bounded on  $B_\epsilon(x_0)$ , then

$$\sum_{|\alpha|=2}^{\infty} |C_{\mathcal{R},|\alpha|}(x_0)| < \infty \quad (\text{B.131})$$

*Proof.* See Corollary 2.10 in [65]. □

The following lemma generalises Lemma D.1 [49] for flat manifolds to the case of general Riemannian manifolds.

**Lemma B.7.3.** *Let  $\mathcal{D}_{X,\epsilon}$  be defined as in (3.41). There exist a constant  $c$ , such that*

$$\lim_{\epsilon \rightarrow 0} \sup_{\|f\|_{C^3(M,\mathbb{R})} \leq K} \left\| \frac{\mathcal{D}_{X,\epsilon} f - f}{\epsilon^2} - (c/2)\Delta_m f \right\|_{C^0(M,\mathbb{R})} = 0,$$

for each  $K < \infty$ .

*Proof.* Let  $f \in C^3(M, \mathbb{R})$  with  $\|f\|_{C^3(M,\mathbb{R})} \leq K$ , fix  $x_0 \in X$  and set  $\epsilon > 0$  to be smaller than the injectivity radius of the point  $x_0 \in M$ . It is well known that the exponential map  $\exp_{x_0}$  at the point  $x_0$  is a diffeomorphism of a neighbourhood of  $\mathbf{0} \in \mathbb{R}^r$  onto the metric ball  $B_\epsilon(x_0)$  (see Theorem 5.11, [7]). Moreover, there exist normal coordinates on the chart  $(B_\epsilon(x_0), \exp_{x_0}^{-1})$ ; that is the components of the metric tensor  $m$  satisfy  $m_{ij} = \delta_{ij}$ , and  $\partial_k m_{ij} = 0$  at the point  $x_0$  for all  $1 \leq i, j, k \leq r$  (see Corollary 5.12, [7]).

Recall the definitions of  $q$  and  $Q_{m,\epsilon}$  from Section 3.4. By the Gauss lemma for Riemannian manifolds, the exponential map  $\exp_{x_0}$  is a radial isometry from the Euclidean ball  $E_\epsilon(0)$  to  $B_\epsilon(x_0)$  (see Lemma 3.5, p.69 in [36]). Thus,

$$\begin{aligned} Q_{m,\epsilon}(x_0, z) &= \epsilon^{-r} q \left( \frac{\text{dist}_m(x_0, z)}{\epsilon} \right) \\ &= \epsilon^{-r} q \left( \frac{|\exp_{x_0}^{-1} x_0 - \exp_{x_0}^{-1} z|}{\epsilon} \right) = \epsilon^{-r} q \left( \frac{|\exp_{x_0}^{-1} z|}{\epsilon} \right), \end{aligned} \quad (\text{B.132})$$

for all  $z \in B_\epsilon(x_0) \subset M$ . Moreover, due to the fact that  $\text{supp } q \subset E_1(0)$ , the function  $Q_{m,\epsilon}$  vanishes for all  $z \in M \setminus B_\epsilon(x_0)$ .

Let  $\{x_1, \dots, x_r\}$  denote normal coordinates on  $(B_\epsilon(x_0), \exp_{x_0}^{-1})$ . Recall that the volume form on  $M$  is given by  $\omega_m^r = \sqrt{\det G_m} \cdot dx_1 \wedge dx_2 \wedge \dots \wedge dx_r$ , where  $G_m$  is a  $r \times r$  matrix with entries  $m_{ij}$ . Hence  $(\exp_{x_0})^* \omega_m^r = \sqrt{\det G_m} \circ \exp_{x_0} d\ell$ , where  $\ell$  is the Lebesgue measure on  $\mathbb{R}^r$ . Moreover, since  $\text{supp } Q_{m,\epsilon}(x_0, \cdot) \subset B_\epsilon(x_0)$ , one has

$$\begin{aligned} \mathcal{D}_{X,\epsilon} f(x_0) &= \int_{B_\epsilon(x_0)} Q_{m,\epsilon}(x_0, z) f(z) \cdot \omega_m^r(z) \\ &= \epsilon^{-r} \int_{E_\epsilon(0)} q \left( \frac{|\mathbf{u}|}{\epsilon} \right) \left( f \cdot \sqrt{\det G_m} \right) \circ \exp_{x_0}(\mathbf{u}) \cdot d\ell(\mathbf{u}), \end{aligned} \quad (\text{B.133})$$

where the last line is due to (B.132). An application of the change of variable  $\mathbf{v} = \mathbf{u}/\epsilon$  to the RHS of (B.133) yields

$$\text{RHS of (B.133)} = \int_{E_1(0)} q(\mathbf{v}) \left( f \cdot \sqrt{\det G_m} \right) \circ \exp_{x_0}(\epsilon \mathbf{v}) d\ell(\mathbf{v}). \quad (\text{B.134})$$

To complete the proof of the lemma from (B.133), we follow the proof of Lemma D.1 [49]. We apply Taylor's theorem to the real-valued function  $\bar{f} := f \circ \exp_x$  on  $E_1(0)$ , centered at 0 to obtain

$$\bar{f}(\epsilon \mathbf{v}) = \sum_{|\alpha|=0}^2 (\epsilon \mathbf{v})^\alpha \frac{D^\alpha \bar{f}(0)}{\alpha!} + \sum_{|\alpha|=3} (\epsilon \mathbf{v})^\alpha R_\alpha(\epsilon \mathbf{v})$$

where the remainder term  $R_\alpha(\epsilon \mathbf{v})$  is given by

$$R_\alpha(\epsilon \mathbf{v}) = \frac{3}{\alpha!} \int_0^1 (1-\gamma)^2 D^\alpha \bar{f}(\gamma \epsilon \mathbf{v}) d\gamma. \quad (\text{B.135})$$

Due to the above Taylor expansion of  $\bar{f}$ , the RHS of (B.134) becomes

$$\int_{E_1(0)} q(|\mathbf{v}|) \left[ \sum_{|\alpha|=0}^2 \epsilon^{|\alpha|} \mathbf{v}^\alpha \frac{D^\alpha \bar{f}(0)}{\alpha!} + \sum_{|\alpha|=3} \epsilon^3 \mathbf{v}^\alpha R_3(\epsilon \mathbf{v}) \right] \cdot \overline{\det G_m}(\epsilon \mathbf{v}) d\ell(\mathbf{v}), \quad (\text{B.136})$$

where  $\overline{\det G_m} = (\sqrt{\det G_m}) \circ \exp_{x_0}$ . We evaluate the above integral term by term.

For the  $|\alpha| = 0$  term, one has

$$\begin{aligned} & \int_{E_1(0)} q(|\mathbf{v}|) \cdot \frac{\bar{f}(0)}{0!} \cdot \overline{\det G_m}(\epsilon \mathbf{v}) d\ell(\mathbf{v}) \\ &= \int_{E_1(0)} q(|\mathbf{v}|) \cdot f(x_0) \cdot \overline{\det G_m}(\epsilon \mathbf{v}) d\ell(\mathbf{v}) \\ &= f(x_0) \int_{B_\epsilon(x_0)} Q_{m,\epsilon}(x_0, z) \cdot \omega_m^r(z) = f(x_0). \end{aligned} \quad (\text{B.137})$$

For the  $|\alpha| = 1$  term, we note that the real-valued function  $q(|\mathbf{v}|)$  is symmetric. Hence,  $v_i q(|\mathbf{v}|)$  are odd functions of  $v$  for  $1 \leq i \leq r$ . Therefore,

$$\begin{aligned} & \int_{E_1(0)} q(|\mathbf{v}|) \cdot \epsilon^1 \sum_{i=1}^r v_i^1 \frac{\partial_i \bar{f}(0)}{1!} \cdot \overline{\det G_m}(\epsilon \mathbf{v}) d\ell(\mathbf{v}) \\ &= \sum_{i=1}^r \partial_i \bar{f}(0) \left( \int_{E_1(0)} v_i q(|\mathbf{v}|) \cdot \epsilon \left( 1 + \sum_{|\beta|=2}^{\infty} C_{\mathcal{R},|\beta|}(x_0) \epsilon^{|\beta|} \mathbf{v}^\beta \right) d\ell(\mathbf{v}) \right) \\ &= \sum_{i=1}^r \partial_i f(x_0) \cdot \left( 0 + \int_{E_1(0)} q(|\mathbf{v}|) \cdot \sum_{|\beta|=2}^{\infty} C_{\mathcal{R},|\beta|}(x_0) v_i \mathbf{v}^\beta \epsilon^{|\beta|+1} d\ell(\mathbf{v}) \right), \end{aligned} \quad (\text{B.138})$$

where we have applied Lemma B.7.2 to obtain the second equality, with constants  $C_{\mathcal{R},|\beta|}(x_0) < \infty$  depend only on the Riemannian curvature tensor  $\mathcal{R}$ , and the covariant derivatives of  $\mathcal{R}$  at the point  $x_0$ . We return to this term later in the proof, but for now we proceed to the  $|\alpha| = 2$  term.

For the  $|\alpha| = 2$  term, due to the property (3.40) for  $q$  and the approximation of  $\sqrt{G_m}$  by Lemma B.7.2, one has

$$\begin{aligned}
 & \int_{E_1(0)} q(|\mathbf{v}|) \cdot \left( \epsilon^2 \sum_{i,j=1}^r v_i v_j \frac{\partial_i \partial_j \bar{f}(0)}{2!} \right) \cdot \left( 1 + \sum_{|\beta|=2}^{\infty} C_{\mathcal{R},|\beta|}(x_0) \epsilon^{|\beta|} \mathbf{v}^\beta \right) d\ell(\mathbf{v}) \\
 &= \sum_{i,j=1}^r \frac{\partial_i \partial_j \bar{f}(0)}{2!} \cdot \left( \int_{E_1(0)} q(|\mathbf{v}|) \cdot \left( v_i v_j \epsilon^2 + \sum_{|\beta|=2}^{\infty} C_{\mathcal{R},|\beta|}(x_0) v_i v_j \mathbf{v}^\beta \epsilon^{|\beta|+2} \right) d\ell(\mathbf{v}) \right) \\
 &= \frac{c\epsilon^2}{2} \sum_{i=1}^r \partial_i^2 \bar{f}(0) + \sum_{i,j=1}^r \frac{\partial_i \partial_j \bar{f}(0)}{2} \cdot \left( \int_{E_1(0)} q(|\mathbf{v}|) \cdot \sum_{|\beta|=2}^{\infty} C_{\mathcal{R},|\beta|}(x_0) v_i v_j \mathbf{v}^\beta \epsilon^{|\beta|+2} d\ell(\mathbf{v}) \right) \\
 &= \frac{c\epsilon^2}{2} \Delta_m f(x_0) + \sum_{i,j=1}^r \frac{\partial_i \partial_j f(x_0)}{2} \cdot \left( \int_{E_1(0)} q(|\mathbf{v}|) \cdot \sum_{|\beta|=2}^{\infty} C_{\mathcal{R},|\beta|}(x_0) v_i v_j \mathbf{v}^\beta \epsilon^{|\beta|+2} d\ell(\mathbf{v}) \right), \tag{B.139}
 \end{aligned}$$

where we have applied Lemma B.7.1 to obtain the last line.

Now set  $\epsilon \leq \min\{\rho, 1\}$ , where  $\rho$  is smaller than the injectivity radius for every  $x \in M$ , then the approximations (B.133)-(B.139) are valid for every point  $x \in M$ . Moreover, since  $M$  is compact  $\mathcal{R}$  is bounded on  $M$ . Therefore, by (B.131), if  $\mathbf{v} \in E_1(0)$  then there exists a constant  $C_1$  such that

$$\sum_{|\beta|=2}^{\infty} |\mathcal{C}_{\mathcal{R},|\beta|}(x) v_i \mathbf{v}^\beta \epsilon^{|\beta|+1}| \leq \epsilon^3 \left( \sum_{|\beta|=2}^{\infty} |\mathcal{C}_{\mathcal{R},|\beta|}(x)| \right) = \frac{C_1}{\int_{E_1(0)} q(|\mathbf{v}|) d\ell(\mathbf{v})} \epsilon^3, \tag{B.140}$$

for each  $i \geq 1$  and all  $x \in M$ . Similarly, if  $v \in E_1(0)$  then there exists a constant  $C_2$  such that

$$\sum_{|\beta|=2}^{\infty} |\mathcal{C}_{\mathcal{R},|\beta|}(x) v_i v_j \mathbf{v}^\beta \epsilon^{|\beta|+2}| \leq \epsilon^4 \left( \sum_{|\beta|=2}^{\infty} |\mathcal{C}_{\mathcal{R},|\beta|}(x)| \right) = \frac{C_2}{\int_{E_1(0)} q(|\mathbf{v}|) d\ell(\mathbf{v})} \epsilon^4, \tag{B.141}$$

for each  $i, j \geq 1$  and all  $x \in M$ . Due to (B.133)-(B.141), one has

$$\begin{aligned}
 & \left| \mathcal{D}_{X,\epsilon} f(x) - f(x) - \frac{c\epsilon^2}{2} \Delta_m f(x) \right| \\
 & \leq \left| \left( \sum_{i=1}^r \partial_i f(x) \right) \cdot C_1 \epsilon^3 \right| + \left| \left( \sum_{i,j=1}^r \partial_i \partial_j f(x) \right) \cdot C_2 \epsilon^4 \right| \\
 & \quad + \left| \int_{E_1(0)} q(|\mathbf{v}|) \cdot \sum_{|\alpha|=3} \epsilon^3 \mathbf{v}^\alpha R_\alpha(\epsilon \mathbf{v}) \cdot \overline{\det G_m(\epsilon \mathbf{v})} d\ell(\mathbf{v}) \right|. \tag{B.142}
 \end{aligned}$$

for all  $x \in M$ . Consider the term on the second line of (B.142), one has

$$\begin{aligned}
 & \sup_{\|f\|_{C^3(M,\mathbb{R})} \leq K} \left\| \int_{E_1(0)} q(|\mathbf{v}|) \cdot \sum_{|\alpha|=3} \epsilon^3 \mathbf{v}^\alpha R_\alpha(\epsilon \mathbf{v}) \cdot \overline{\det G_m(\epsilon \mathbf{v})} d\ell(\mathbf{v}) \right\|_{C^0(M,\mathbb{R})} \\
 & \leq \sup_{\substack{\|f\|_{C^3(M,\mathbb{R})} \leq K \\ |\alpha|=3, \mathbf{u} \in E_\epsilon(0)}} \|R_\alpha(\mathbf{u})\|_{C^0(M,\mathbb{R})} \cdot \epsilon^3 \int_{E_1(0)} q(|\mathbf{v}|) \cdot \sum_{|\alpha|=3} \mathbf{v}^\alpha \overline{\det G_m(\epsilon \mathbf{v})} d\ell(\mathbf{v}) \\
 & = \sup_{\substack{\|f\|_{C^3(M,\mathbb{R})} \leq K \\ |\alpha|=3, \mathbf{u} \in E_\epsilon(0)}} \|R_\alpha(\mathbf{u})\|_{C^0(M,\mathbb{R})} \cdot C_3 \epsilon^3,
 \end{aligned}$$

for some constant  $C_3$ . Therefore, rearranging (B.142) yields

$$\begin{aligned}
 & \sup_{\|f\|_{C^3(M,\mathbb{R})} \leq K} \left\| \frac{(\mathcal{D}_{X,\epsilon} - \text{Id})f}{\epsilon^2} - (c/2)\Delta_m f \right\|_{C^0(M,\mathbb{R})} \\
 & \leq \sup_{\|f\|_{C^3(M,\mathbb{R})} \leq K} \left\| \left( \sum_{i=1}^r \partial_i f \right) \right\|_{C^0(M,\mathbb{R})} \cdot C_1 \epsilon^3 + \left\| \left( \sum_{i,j=1}^r \partial_i \partial_j f \right) \right\|_{C^0(M,\mathbb{R})} \cdot C_2 \epsilon^4 \\
 & \quad + \sup_{\substack{\|f\|_{C^3(M,\mathbb{R})} \leq K \\ |\alpha|=3, \mathbf{u} \in E_\epsilon(0)}} \|R_\alpha(\mathbf{u})\|_{C^0(M,\mathbb{R})} \cdot C_3 \epsilon^1. \tag{B.143}
 \end{aligned}$$

Since the first and second order derivatives of  $f$  are bounded for by  $K$ , the first two terms on the RHS of (B.143) converge to 0 as  $\epsilon \rightarrow \infty$ . Hence, to complete the proof of the theorem it suffices to show that

$$R_\alpha(\mathbf{u}) = \frac{3}{\alpha!} \int_0^1 (1-\gamma)^2 D^\alpha(f \circ \exp_x)(\gamma \mathbf{u}) d\gamma,$$

is uniformly bounded on  $E_\epsilon(0)$  for  $|\alpha| = 3$  and every  $f \in C^3(M, \mathbb{R})$  with  $\|f\|_{C^3(M,\mathbb{R})} \leq K$ .

Let  $\mathbf{u} \in E_\epsilon(0)$  and  $|\alpha| = 3$ . Since  $\epsilon$  is less than the injectivity radius of  $x$ , the exponential map  $\exp_x^{-1}$  is a  $C^\infty$ -diffeomorphism from  $B_\epsilon(x)$  onto  $E_\epsilon(0)$ . Thus, if  $\|f\|_{C^3(M,\mathbb{R})} \leq K$ , then all derivatives of  $f \circ \exp_x$  up to order 3 are bounded above by  $K'$  for some  $K' < \infty$  on  $E_\epsilon(0)$ . Now since  $\mathbf{u} \in E_\epsilon(0)$ , one has  $\gamma \mathbf{u} \in E_\epsilon(0)$  for all  $0 \leq \gamma \leq 1$ . Hence, the term  $D^\alpha(f \circ \exp_x)(\gamma \mathbf{u})$  is uniformly bounded in  $\mathbf{u}$  for  $0 \leq \gamma \leq 1$ , and all  $\|f\|_{C^3(M,\mathbb{R})} \leq K$ . It follows that the remainder  $R_\alpha$  is uniformly bounded on  $E_\epsilon(0)$ , for  $|\alpha| = 3$  and every  $\|f\|_{C^3(M,\mathbb{R})} \leq K$ .  $\square$

**Proof of Theorem 3.4.1.** Let  $\|f\|_{C^3(M,\mathbb{R})} \leq 1$ , and set  $\epsilon > 0$  to be smaller than the injectivity radius of each point in  $M$ . We start with the asymptotic expansions of  $\mathcal{P}_\epsilon(fh_\mu)$ . Since  $\|f\|_{C^3(M,\mathbb{R})} \leq 1$  and  $h_\mu$  is bounded in the  $C^3$ -norm, one has  $\|fh_\mu\|_{C^3(M,\mathbb{R})} \leq K$ , for some constant  $K$ . Consider  $fh_\mu$  such that  $\|fh_\mu\|_{C^3(M,\mathbb{R})} \leq K$ .

Lemma B.7.3 yields  $\mathcal{D}_{X,\epsilon}(fh_\mu) = fh_\mu + \frac{c\epsilon^2}{2}\Delta_m(fh_\mu) + \mathcal{O}(\epsilon^3)$ , where  $\mathcal{O}(\epsilon^3)$  denotes the class of polynomials  $a_3\epsilon^3 + a_4\epsilon^4 + \dots$ , with all coefficients  $a_3, a_4, \dots$  bounded on  $M$  and independent of  $f$ . Combining the expansion of  $\mathcal{D}_{X,\epsilon}(fh_\mu)$  with the linearity of  $\mathcal{P}$ , then  $\mathcal{P}\mathcal{D}_{X,\epsilon}(fh_\mu) = \mathcal{P}(fh_\mu) + \frac{c\epsilon^2}{2}\mathcal{P}\Delta_m(fh_\mu) + \mathcal{O}(\epsilon^3)$ . Now, since  $\mathcal{P}$  is given by (B.22) and  $T$  is a  $C^\infty$ -diffeomorphism, one has  $\mathcal{P}\mathcal{D}_{X,\epsilon}(fh_\mu) \in F^3(N, \mathbb{R})$ . Therefore, by a straightforward modification of Lemma B.7.3, we have uniformly on  $N$

$$\begin{aligned} \mathcal{P}_\epsilon(fh_\mu) &= \mathcal{D}_{Y',\epsilon}\mathcal{P}\mathcal{D}_{X,\epsilon}(fh_\mu) \\ &= \mathcal{P}(fh_\mu) + \frac{c\epsilon^2}{2}\mathcal{P}\Delta_m(fh_\mu) + \frac{c\epsilon^2}{2}[\Delta_n\mathcal{P}(fh_\mu) + \mathcal{O}(\epsilon^2)] + \mathcal{O}(\epsilon^3) \\ &= \mathcal{P}(fh_\mu) + \frac{c\epsilon^2}{2}[\mathcal{P}\Delta_m(fh_\mu) + \Delta_n\mathcal{P}(fh_\mu)] + \mathcal{O}(\epsilon^3), \end{aligned} \quad (\text{B.144})$$

where  $c$  is the same constant as in Lemma B.7.3 (since the constant  $c$  comes from the property (3.40) of  $Q$ , independent of  $f$ ). Therefore, using the fact that  $\mathcal{P}h_\mu = h_\nu$

$$\mathcal{H}_\epsilon f = \frac{\mathcal{P}_\epsilon(fh_\mu)}{\mathcal{P}_\epsilon h_\mu} = \frac{\mathcal{P}(fh_\mu) + \frac{c\epsilon^2}{2}[\mathcal{P}\Delta_m(fh_\mu) + \Delta_n\mathcal{P}(fh_\mu)] + \mathcal{O}(\epsilon^3)}{h_\nu + \frac{c\epsilon^2}{2}[\mathcal{P}\Delta_m h_\mu + \Delta_n h_\nu] + \mathcal{O}(\epsilon^3)}, \quad (\text{B.145})$$

uniformly on  $N$ . Next we apply  $\mathcal{H}_\epsilon^*$  to  $\mathcal{H}_\epsilon f$ . According to (B.129), the first step is the application of the dual diffusion operator  $\mathcal{D}_{Y',\epsilon}^*$  to (B.145). In preparation for this, we consider a general polynomial quotient of the form

$$\frac{a + b\epsilon^2 + c\epsilon^3}{d + e\epsilon^2 + f\epsilon^3}$$

where  $a, b, \dots, f$  are a set of known coefficients. By polynomial long division and truncating at  $\epsilon^3$ , one has

$$\frac{a + b\epsilon^2 + c\epsilon^3}{d + e\epsilon^2 + f\epsilon^3} = \frac{a}{d} + \frac{bd - ae}{d^2}\epsilon^2 + \mathcal{O}(\epsilon^3). \quad (\text{B.146})$$

Applying (B.146) to (B.145), and noting that  $\mathcal{H}f = \mathcal{P}(f \cdot h_\mu)/h_\nu$  (see (3.19)) yields

$$\begin{aligned} \mathcal{H}_\epsilon f &= \frac{\mathcal{P}(fh_\mu)}{h_\nu} + \frac{c\epsilon^2}{2} \left[ \frac{\mathcal{P}\Delta_m(fh_\mu)}{h_\nu} + \frac{\Delta_n\mathcal{P}(fh_\mu)}{h_\nu} - \frac{\mathcal{P}(fh_\mu) \cdot \mathcal{P}\Delta_m h_\mu}{h_\nu^2} \right. \\ &\quad \left. - \frac{\mathcal{P}(fh_\mu) \cdot \Delta_n h_\nu}{h_\nu^2} \right] + \mathcal{O}(\epsilon^3) \\ &= \mathcal{H}f + \frac{c\epsilon^2}{2} \left[ \frac{\mathcal{P}\Delta_m(fh_\mu)}{h_\nu} + \frac{\Delta_n\mathcal{P}(fh_\mu)}{h_\nu} - \frac{\mathcal{H}f \cdot \mathcal{P}\Delta_m h_\mu}{h_\nu} - \frac{\mathcal{H}f \cdot \Delta_n h_\nu}{h_\nu} \right] + \mathcal{O}(\epsilon^3) \end{aligned}$$

uniformly on  $N$ . Since  $h_\nu$  is uniformly bounded away from zero, one can check that  $\mathcal{H}_\epsilon f \in F^3(N, \mathbb{R})$ . Hence, it is now straightforward to compute  $\mathcal{D}_{Y',\epsilon}^*\mathcal{H}_\epsilon f$  via Lemma B.7.3 to obtain

$$\begin{aligned} \mathcal{D}_{Y',\epsilon}^*\mathcal{H}_\epsilon f &= \mathcal{H}f + \frac{c\epsilon^2}{2} \left[ \frac{\mathcal{P}\Delta_m(fh_\mu)}{h_\nu} + \frac{\Delta_n\mathcal{P}(fh_\mu)}{h_\nu} - \frac{\mathcal{H}f \cdot \mathcal{P}\Delta_m h_\mu}{h_\nu} - \frac{\mathcal{H}f \cdot \Delta_n h_\nu}{h_\nu} \right] \\ &\quad + \frac{c\epsilon^2}{2}\Delta_n\mathcal{H}f + \mathcal{O}(\epsilon^3), \end{aligned} \quad (\text{B.147})$$

uniformly on  $N$ . We write

$$\begin{aligned}
 \frac{\mathcal{P}(\Delta_m(fh_\mu))}{h_\nu} &= \mathcal{H}\left(\frac{\Delta_m(fh_\mu)}{h_\mu}\right) \\
 &= \mathcal{H}\left(\frac{f \cdot \Delta_m h_\mu + h_\mu \cdot \Delta_m f + 2m(\nabla_m h_\mu, \nabla_m f)}{h_\mu}\right) \\
 &= \mathcal{H}\left(\frac{f \cdot \Delta_m h_\mu}{h_\mu}\right) + \mathcal{H}\left(\Delta_m f + \frac{2m(\nabla_m h_\mu, \nabla_m f)}{h_\mu}\right) \\
 &= \frac{\mathcal{H}f \cdot \mathcal{P}\Delta_m h_\mu}{h_\nu} + \mathcal{H}\left(\Delta_m f + \frac{2m(\nabla_m h_\mu, \nabla_m f)}{h_\mu}\right),
 \end{aligned}$$

where we have used the linearity of  $\mathcal{H}$  to obtain the penultimate line, and the fact that  $\mathcal{H}f = f \circ T^{-1}$  to obtain the last line. Thus, the 2<sup>nd</sup> and 4<sup>th</sup> terms of (B.147) can be combined to form

$$\frac{c\epsilon^2}{2} \left( \frac{\mathcal{P}\Delta_m(fh_\mu)}{h_\nu} - \frac{\mathcal{H}f \cdot \mathcal{P}\Delta_m h_\mu}{h_\nu} \right) = c\epsilon^2 \mathcal{H} \left( \frac{\Delta_m f}{2} + \frac{m(\nabla_m h_\mu, \nabla_m f)}{h_\mu} \right). \quad (\text{B.148})$$

Also,

$$\begin{aligned}
 \frac{\Delta_n \mathcal{P}(fh_\mu)}{h_\nu} &= \frac{\Delta_n(\mathcal{H}f \cdot h_\nu)}{h_\nu} = \frac{h_\nu \cdot \Delta_n \mathcal{H}f + \mathcal{H}f \cdot \Delta_n h_\nu + 2n(\nabla_n h_\nu, \nabla_n \mathcal{H}f)}{h_\nu} \\
 &= \Delta_n \mathcal{H}f + \frac{\mathcal{H}f \cdot \Delta_n h_\nu}{h_\nu} + \frac{2n(\nabla_n h_\nu, \nabla_n \mathcal{H}f)}{h_\nu}.
 \end{aligned}$$

Thus, the 3<sup>rd</sup>, 5<sup>th</sup> and 6<sup>th</sup> terms of (B.147) can be combined to form

$$\begin{aligned}
 &\frac{c\epsilon^2}{2} \left[ \frac{\Delta_n \mathcal{P}(fh_\mu)}{h_\nu} - \frac{\mathcal{H}f \cdot \Delta_n h_\nu}{h_\nu} \right] + \frac{c\epsilon^2}{2} \Delta_n \mathcal{H}f \\
 &= \frac{c\epsilon^2}{2} \Delta_n \mathcal{H}f + c\epsilon^2 \frac{n(\nabla_n h_\nu, \nabla_n \mathcal{H}f)}{h_\nu} + \frac{c\epsilon^2}{2} \Delta_n \mathcal{H}f \\
 &= c\epsilon^2 \left( \Delta_n f + \frac{n(\nabla_n h_\nu, \nabla_n \mathcal{H}f)}{h_\nu} \right) \\
 &= c\epsilon^2 (\Delta_\nu \mathcal{H}f), \quad (\text{B.149})
 \end{aligned}$$

where the last line is due to (3.27). Substituting (B.148) and (B.149) into the RHS of (B.147) yields,

$$\mathcal{D}_{Y_{\epsilon, \epsilon}}^* \mathcal{H}_\epsilon f = \mathcal{H}f + c\epsilon^2 \left( \mathcal{H} \left( \frac{\Delta_m f}{2} + \frac{m(\nabla_m h_\mu, \nabla_m f)}{h_\mu} \right) + \Delta_\nu \mathcal{H}f \right) + \mathcal{O}(\epsilon^3), \quad (\text{B.150})$$

uniformly on  $N$ . It is straightforward to apply  $\mathcal{D}_{X_{\epsilon, \epsilon}}^* \mathcal{H}^*$  to the RHS of (B.150) via Lemma B.7.3, which yields

$$\begin{aligned}
 \mathcal{H}_\epsilon^* \mathcal{H}_\epsilon f &= f + \frac{c\epsilon^2}{2} \Delta_m f + c\epsilon^2 \left( \left( \frac{\Delta_m f}{2} + \frac{m(\nabla_m h_\mu, \nabla_m f)}{h_\mu} \right) + \mathcal{H}^* \Delta_\nu \mathcal{H}f \right) + \mathcal{O}(\epsilon^3) \\
 &= f + \frac{c\epsilon^2}{2} (\Delta_\mu f + \mathcal{H}^* \Delta_\nu \mathcal{H}f) + \mathcal{O}(\epsilon^3), \quad (\text{B.151})
 \end{aligned}$$

uniformly on  $M$ , where we have used (B.31) to obtain the last line. Since the coefficients of the  $\mathcal{O}(\epsilon^3)$  are uniform on  $M$  and independent of  $f$ , rearranging (B.151) gives

$$\lim_{\epsilon \rightarrow 0} \sup_{\|f\|_{C^3(M, \mathbb{R})}} \left\| \frac{(\mathcal{H}_\epsilon^* \mathcal{H} - I)f}{\epsilon^2} - c \cdot \Delta^{dyn} f \right\|_{C^0(M, \mathbb{R})} = 0.$$

□

# Appendix C

## C.1 The proof of Theorem 4.1.2

Let  $\mathbf{L}^{\mu,\epsilon}$  be as in (4.9), where  $w_{ij}^{\mu,\epsilon}$  are given by (4.8) for  $1 \leq i, j \leq k$ . Due to (4.14), we have for all  $\mathbf{f} \in \mathbb{R}^k$

$$(\mathbf{L}^{k,\mu,\epsilon}\mathbf{f})_i = \sum_{j=1}^k w_{ij}^{\mu,\epsilon}(f_j - f_i). \quad (\text{C.1})$$

Let  $\mathbf{g} \in \mathbb{R}^k$ . By (4.10) and (C.1), one has

$$\begin{aligned} \langle \mathbf{g}, \mathbf{L}^{k,\mu,\epsilon}\mathbf{f} \rangle_\mu &= \sum_{i,j=1}^k w_{ij}^{\mu,\epsilon}(f_j - f_i)g_i \frac{h_\mu(x_i)}{p^{k,\epsilon}(x_i)} \\ &= \sum_{i,j=1}^k Q_{\rho,\epsilon}(x_i, x_j) \frac{\sqrt{h_\mu(x_j)}}{p^{k,\epsilon}(x_j)} \frac{\sqrt{h_\mu(x_i)}}{p^{k,\epsilon}(x_i)} (f_j - f_i)g_i \\ &= \sum_{i,j=1}^k U_{ij}(f_j - f_i)g_i, \end{aligned} \quad (\text{C.2})$$

where

$$U_{ij} := Q_{\rho,\epsilon}(x_i, x_j) \frac{\sqrt{h_\mu(x_j)}}{p^{k,\epsilon}(x_j)} \frac{\sqrt{h_\mu(x_i)}}{p^{k,\epsilon}(x_i)}.$$

We note that  $U_{ij}$  is symmetric due to the symmetry of  $Q_{\rho,\epsilon}$ , and positive since each  $h_\mu$ ,  $Q_{\rho,\epsilon}$  and  $p^{k,\epsilon}$  are positive. Therefore, by (C.2), the conclusions of Theorem 4.1.2 follows by a straightforward modification of the arguments in Section 1.4 [25].

## C.2 The proof of Theorem 4.1.4

We adapt the arguments of Section 5.1 in [12]. Let  $S^k = \{x_i\}_{i=1}^k$  be an i.i.d random sample drawn from  $M$  according to the  $C^5$  probability density  $p$ , and  $L^{\mu,k,\epsilon}$  be as in (4.15). Fix  $f \in C^3(M, \mathbb{R})$  and  $x_0 \in M$ . Define

$$L^{\mu,\epsilon}f(x_0) := \int_M Q_{\rho,\epsilon}(x_0, y) \sqrt{\frac{h_\mu(y)}{h_\mu(x)}} (f(x_0) - f(y)) \omega_m^r(y), \quad (\text{C.3})$$

where  $Q_{\rho,\epsilon} : M \rightarrow \mathbb{R}$  is as in (4.4). First, we show that there is a sequence of scalars  $\{\epsilon_k\}_{k \geq 1}$  with  $\lim_{k \rightarrow \infty} \epsilon_k = 0$ , such that

$$\lim_{k \rightarrow \infty} \left| \frac{1}{\epsilon_k} L^{\mu,k,\epsilon_k} f(x_0) - \frac{1}{\epsilon_k} L^{\mu,\epsilon_k} f(x_0) \right| = 0. \quad (\text{C.4})$$

Since  $S^k$  is an i.i.d random sample with probability density  $p$ ,  $h_\mu > 0$  is bounded, and  $p^{k,\epsilon}(x_j) = \frac{1}{k} \sum_{l=1}^k Q_{\rho,\epsilon}(x_j, x_l) > 0$  the random variables

$$X_j = \frac{Q_{\rho,\epsilon}(x_0, x_j)}{p^{k,\epsilon}(x_j)} \sqrt{\frac{h_\mu(x_j)}{h_\mu(x_0)}} (f(x_0) - f(x_j)),$$

are strictly bounded on the interval  $[a_i, b_i]$  for every  $j \in [1, k]$ . Therefore by (4.15) and the Hoeffding inequality [73], one has with high probability

$$\begin{aligned} \lim_{k \rightarrow \infty} L^{\mu,k,\epsilon} f(x_0) &= \lim_{k \rightarrow \infty} \frac{1}{k} \sum_{j=1}^k \frac{Q_{\rho,\epsilon}(x_0, x_j)}{p^{k,\epsilon}(x_j)} \sqrt{\frac{h_\mu(x_j)}{h_\mu(x_0)}} (f(x_j) - f(x_0)) \\ &= \lim_{k \rightarrow \infty} \int_M \frac{Q_{\rho,\epsilon}(x_0, y)}{p^{k,\epsilon}(y)} \sqrt{\frac{h_\mu(y)}{h_\mu(x_0)}} (f(y) - f(x_0)) p(y) \omega_m^r(y). \end{aligned} \quad (\text{C.5})$$

In addition, since  $p \in C^5(M, \mathbb{R})$  is bounded away from zero on  $M$ , by (4.6) and (4.7) there exists a constant  $a_\rho$  such that

$$\lim_{k \rightarrow \infty} \left( \frac{1}{p(y)} - \frac{1}{p^{k,\epsilon}(y)} \right) = \lim_{k \rightarrow \infty} \frac{\epsilon a_\rho \Delta_m p(y) + R_1(y, \epsilon^{3/2})}{p^{k,\epsilon}(y) p(y)}, \quad R_1(y, \epsilon^{3/2}) \in \mathcal{O}(\epsilon^{3/2}), \quad (\text{C.6})$$

Therefore, by combining (C.5) and (C.6)

$$\begin{aligned} &\lim_{k \rightarrow \infty} \frac{1}{\epsilon} \left( \frac{1}{k} L^{k,\mu,\epsilon} f(x_0) - L^{\mu,\epsilon} f(x_0) \right) \\ &= \lim_{k \rightarrow \infty} \frac{1}{\epsilon} \int_M Q_{\rho,\epsilon}(x_0, y) \left( \frac{1}{p^{k,\epsilon}(y)} - \frac{1}{p(y)} \right) p(y) \sqrt{\frac{h_\mu(y)}{h_\mu(x_0)}} (f(y) - f(x_0)) \omega_m^r(y) \\ &= \lim_{k \rightarrow \infty} \frac{1}{\epsilon} \int_M Q_{\rho,\epsilon}(x_0, y) \left( \frac{\epsilon a_\rho \Delta_m p(y) + R_1(y, \epsilon^{3/2})}{p^{k,\epsilon}(y)} \right) \sqrt{\frac{h_\mu(y)}{h_\mu(x_0)}} (f(y) - f(x_0)) \omega_m^r(y). \end{aligned} \quad (\text{C.7})$$

Since  $p \in C^5(M, \mathbb{R})$ ,  $h_\mu > 0$  is bounded above, and  $p^{k,\epsilon}(y) > 0$  for all  $k$ , there exists a sequence of bounded functions  $g_k$  in  $C^3(M, \mathbb{R})$  such that

$$\begin{aligned} \text{RHS of (C.7)} &= \lim_{k \rightarrow \infty} \int_M Q_{\rho,\epsilon}(x_0, y) \left( g_k(y) - \frac{R_1(y, \epsilon^{3/2})}{p^{k,\epsilon}(y)} \right) \sqrt{\frac{h_\mu(y)}{h_\mu(x_0)}} (f(y) - f(x_0)) \omega_m^r(y) \\ &= \lim_{k \rightarrow \infty} \int_M Q_{\rho,\epsilon}(x_0, y) g_k(y) \sqrt{\frac{h_\mu(y)}{h_\mu(x_0)}} (f(y) - f(x_0)) \omega_m^r(y) + R_2(x_0, \epsilon^{1/2}), \end{aligned} \quad (\text{C.8})$$

where  $R_2(x_0, \epsilon^{1/2}) \in \mathcal{O}(\epsilon^{1/2})$ . Moreover, since  $g_k, h_\mu \in C^3(M, \mathbb{R})$  and  $\rho < \mathcal{S}_M$ , by applying Lemma B.7.3 to the functions  $g_k \cdot \sqrt{h_\mu} \cdot f \in C^3(M, \mathbb{R})$  and  $g_k \cdot \sqrt{h_\mu} \in C^3(M, \mathbb{R})$ , one has

$$\begin{aligned} \text{RHS of (C.8)} &= \lim_{k \rightarrow \infty} \frac{\epsilon b_\rho}{\sqrt{h_\mu(x_0)}} \left[ \Delta_m \left( g_k \cdot \sqrt{h_\mu} \cdot f \right) (x_0) - f(x_0) \Delta_m \left( g_k \cdot \sqrt{h_\mu} \right) (x_0) \right] \\ &\quad + R_3(x_0, \epsilon^{3/2}) + R_2(x_0, \epsilon^{1/2}), \end{aligned} \quad (\text{C.9})$$

where  $R_3(x_0, \epsilon^{3/2}) \in \mathcal{O}(\epsilon^{3/2})$ , and  $b_\rho > 0$  depends on the second moment of  $q_\rho$ . Note that by the assumptions on  $h_\mu$  the fact that  $g_k \in C^3(M, \mathbb{R})$  for every  $k$ , the RHS of (C.9) vanishes as  $\epsilon \rightarrow 0$ . Thus, by substituting  $\epsilon = \epsilon_k$  in (C.7)-(C.9) and using the fact that  $\lim_{k \rightarrow \infty} \epsilon_k = 0$ , we arrive at the result (C.4).

Since  $f \in C^3(M, \mathbb{R})$  and  $x_0 \in M$  were choose arbitrarily, one can easily extend (C.4) to

$$\lim_{k \rightarrow \infty} \left( \sup_{\|f\|_{C^3(M, \mathbb{R})} \leq 1} \left\| \frac{L^{\mu, k, \epsilon_k} - L^{\mu, \epsilon_k}}{\epsilon_k} \right\|_{C^0(M, \mathbb{R})} \right) = 0. \quad (\text{C.10})$$

Therefore, to complete the proof of Theorem 4.2.2, it remains to show that there is a constant  $C_\rho > 0$  such that  $\frac{1}{\epsilon C_\rho} \lim_{\epsilon \rightarrow 0} L^{\mu, \epsilon} f(x) = \Delta_\mu f(x)$  for all  $x \in M$  and  $f \in C^\infty(M, \mathbb{R})$ , because then by (C.10) one has

$$\lim_{k \rightarrow \infty} \frac{1}{\epsilon_k C_\rho} L^{\mu, k, \epsilon_k} f(x) = \lim_{k \rightarrow \infty} \frac{1}{\epsilon_k C_\rho} L^{\mu, \epsilon_k} f(x) = \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon C_\rho} L^{\mu, \epsilon} f(x) = \Delta_\mu f(x), \quad (\text{C.11})$$

for every  $f \in C^3(M, \mathbb{R})$  and  $x \in M$ , and the desired conclusion follows. Since  $h_\mu > 0$  is in  $C^3$ , by a similar argument as in (C.9) one has  $C_\rho > 0$  such that

$$\begin{aligned} &\lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon C_\rho} L^{\mu, \epsilon} f(x) \\ &= \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon C_\rho} \int_M Q_{\rho, \epsilon}(x, y) \sqrt{\frac{h_\mu(y)}{h_\mu(x)}} (f(y) - f(x)) \omega_m^r(y) \quad \text{by (C.3)} \\ &= \frac{\Delta_m(\sqrt{h_\mu} \cdot f)(x)}{\sqrt{h_\mu(x)}} - \frac{f(x) \Delta_m(\sqrt{h_\mu})(x)}{\sqrt{h_\mu(x)}} \quad \text{by Lemma B.7.3} \\ &= \frac{f(x) \Delta_m(\sqrt{h_\mu})(x)}{\sqrt{h_\mu(x)}} + \Delta_m f(x) + \frac{2m(\nabla_m(\sqrt{h_\mu}), \nabla_m f)_x}{\sqrt{h_\mu(x)}} - \frac{f(x) \Delta_m(\sqrt{h_\mu})(x)}{\sqrt{h_\mu(x)}} \\ &= \Delta_m f(x) + \frac{m(\nabla_m h_\mu, \nabla_m f)_x}{h_\mu(x)} = \Delta_\mu f(x), \end{aligned} \quad (\text{C.12})$$

where we had applied to the chain rule of differentiation on  $\nabla_m(\sqrt{h_\mu})$  to obtain the penultimate equality.

### C.3 The proof of Theorem 4.2.2

Recall from Section 4.2, since  $M \cup N \subset \mathbb{R}^d$ , by a straightforward modification of (4.4) one has  $Q_{\rho,\epsilon} : (M \cup N) \times (M \cup N) \rightarrow \mathbb{R}^+$ . Also, recall that the random sample  $\hat{S}^k = \{Tx_i\}_{i=1}^k$  is drawn from  $N$  according to the  $C^5$  probability density  $\hat{p}$ . Finally, recall the Definition 3.3.1

$$\Delta^{dyn} := \frac{1}{2}(\Delta_\mu + \mathcal{H}^* \Delta_\nu \mathcal{H}), \quad (\text{C.13})$$

where  $\mathcal{H} : L^2(M, m, \mu_r) \rightarrow L^2(M, n, \nu_r)$  is as in (3.19), with adjoint  $\mathcal{H}^*$ . Define

$$L^{\nu,k,\epsilon} f(x) := \sum_{j=1}^k \frac{Q_{\rho,\epsilon}(Tx, Tx_j)}{\hat{p}^{k,\epsilon}(Tx_j)} \sqrt{\frac{h_\nu(Tx_j)}{h_\nu(Tx)}} (f(x) - f(x_j)), \quad (\text{C.14})$$

where  $\hat{p}^{k,\epsilon}$  is as in (4.23). By (4.30) and (C.14), one has

$$L^{dyn,k,\epsilon} = \frac{1}{2} (L^{\mu,k,\epsilon} + L^{\nu,k,\epsilon}). \quad (\text{C.15})$$

Furthermore, by Theorem 4.1.4, there exists a constant  $C_\rho$  and a sequence of scalars  $\{\hat{\epsilon}_k\}_{k \geq 1}$  with  $\lim_{k \rightarrow \infty} \epsilon_k \rightarrow 0$ , such that

$$\lim_{k \rightarrow 0} \left( \sup_{\|f\|_{C^3(M, \mathbb{R})} \leq 1} \left\| \frac{1}{\epsilon_k C_\rho} L^{\mu,k,\hat{\epsilon}_k} f - \Delta_\mu f \right\|_{C^3(M, \mathbb{R})} \right), \quad (\text{C.16})$$

for all  $x \in M$  and  $f \in C^3(M, \mathbb{R})$ . Therefore, to obtain the conclusion of Theorem 4.2.2, by the linearity of (C.13) and (C.15), and the convergence (C.16), it is sufficient to show that there exists a sequence of scalars  $\{\epsilon_k\}_{k \geq 1}$  containing  $\{\hat{\epsilon}_k\}_{k \geq 1}$  and with  $\lim_{k \rightarrow \infty} \hat{\epsilon}_k = 0$  such that

$$\lim_{k \rightarrow \infty} \left( \sup_{\|f\|_{C^3(M, \mathbb{R})} \leq 1} \left\| \frac{1}{\hat{\epsilon}_k C_\rho} L^{\nu,k,\epsilon_k} f - \mathcal{H}^* \Delta_\nu \mathcal{H} f \right\|_{C^3(M, \mathbb{R})} \right), \quad (\text{C.17})$$

where  $C_\rho$  is as in (C.16).

To show that (C.17) is valid, we follow the proof of Theorem 4.1.4 outlined in Section C.2. Define analogous to (C.3)

$$L^{\nu,\epsilon} f(x) := \int_N Q_{\rho,\epsilon}(Tx, y) \sqrt{\frac{h_\nu(y)}{h_\nu(Tx)}} (f(T^{-1}y) - f(x)) \omega_n^r(y), \quad (\text{C.18})$$

for all  $x \in M$  and  $f \in C^\infty(M, \mathbb{R})$ . Fix  $x_0 \in M$  and  $f \in C^3(M, \mathbb{R})$ . Since  $\hat{S}^k := \{Tx_1, Tx_2, \dots, Tx_k\}$  is an i.i.d random sample drawn from  $(N, n, \nu_r)$  according to the  $C^5$  probability density  $\hat{p} : N \rightarrow \mathbb{R}^+$ . Therefore, analogous to (C.5), one has with high probability

$$\lim_{k \rightarrow \infty} \frac{1}{k} L^{\nu,k,\epsilon} f(x_0) = \lim_{k \rightarrow \infty} \int_N \frac{Q_{\rho,\epsilon}(Tx_0, y)}{\hat{p}^{k,\epsilon}(y)} \sqrt{\frac{h_\nu(y)}{h_\nu(Tx_0)}} (f(T^{-1}y) - f(x_0)) \hat{p}(y) \omega_n^r(y). \quad (\text{C.19})$$

Moreover, analogous to (C.6) one has for all  $y \in N$ ,

$$\lim_{k \rightarrow \infty} \left( \frac{1}{\hat{p}(y)} - \frac{1}{\hat{p}^{k,\epsilon}(y)} \right) = \lim_{k \rightarrow \infty} \frac{\epsilon \hat{a}_\rho \Delta_n \hat{p}(y) + \hat{R}_1(y, \epsilon^{3/2})}{\hat{p}^{k,\epsilon}(y) \hat{p}(y)}, \quad (\text{C.20})$$

where  $\hat{R}_1(y, \epsilon^{3/2}) \in \mathcal{O}(\epsilon^{3/2})$ . Hence, since  $h_\nu \in C^3(N, \mathbb{R})$  is bounded above and uniformly away from zero, by applying similar arguments as in (C.7)-(C.9), one has analogous to (C.11)

$$\lim_{k \rightarrow \infty} \left( \sup_{\|f\|_{C^3(M, \mathbb{R})} \leq 1} \left\| \frac{L^{\nu, k, \epsilon_k} f - L^{\nu, \epsilon_k} f}{\epsilon_k} \right\|_{C^0(M, \mathbb{R})} \right) = 0. \quad (\text{C.21})$$

Now analogous to (C.12), there is a constant  $C_\rho > 0$  such that for all  $x \in M$  and  $f \in C^3(M, \mathbb{R})$

$$\begin{aligned} & \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon C_\rho} L^{\nu, \epsilon} f(x) \\ &= \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon C_\rho} \int_N Q_{\rho, \epsilon}(Tx, y) \sqrt{\frac{h_\nu(y)}{h_\nu(Tx)}} (f(T^{-1}y) - f(x)) \omega_n^r(y) \quad \text{by (C.18)} \\ &= \frac{\Delta_n(\sqrt{h_\nu} \cdot f \circ T^{-1})(Tx)}{\sqrt{h_\nu(Tx)}} - \frac{f(x) \Delta_n(\sqrt{h_\nu})(Tx)}{\sqrt{h_\nu(Tx)}} \quad \text{by Lemma B.7.3} \\ &= \Delta_n(f \circ T^{-1})(Tx) + \frac{2n(\nabla_n(\sqrt{h_\nu}), \nabla_n(f \circ T^{-1}))_{Tx}}{\sqrt{h_\nu(Tx)}} \\ &= \Delta_n(f \circ T^{-1})(Tx) + \frac{n(\nabla_n h_\nu, \nabla_n(f \circ T^{-1}))_{Tx}}{h_\nu(Tx)} \\ &= \Delta_n \mathcal{H}f(Tx) + \frac{n(\nabla_n h_\nu, \nabla_n(\mathcal{H}f))_{Tx}}{h_\nu(Tx)} \\ &= \mathcal{H}^* \Delta_\nu \mathcal{H}f(x), \end{aligned}$$

where we have applied Lemma (3.2.2) to obtain the last two equality. Therefore

$$\begin{aligned} & \lim_{k \rightarrow \infty} \left( \sup_{\|f\|_{C^3(M, \mathbb{R})} \leq 1} \left\| \frac{1}{\epsilon_k C_\rho} L^{\nu, k, \epsilon_k} f - \mathcal{H}^* \Delta_\nu \mathcal{H}f \right\|_{C^0(M, \mathbb{R})} \right) \\ & \leq \lim_{k \rightarrow \infty} \left( \sup_{\|f\|_{C^3(M, \mathbb{R})} \leq 1} \left\| \frac{1}{\epsilon_k C_\rho} L^{\nu, \epsilon_k} f - \mathcal{H}^* \Delta_\nu \mathcal{H}f \right\|_{C^0(M, \mathbb{R})} \right) \quad \text{by (C.21)} \\ & = \lim_{\epsilon \rightarrow 0} \left( \sup_{\|f\|_{C^3(M, \mathbb{R})} \leq 1} \left\| \frac{1}{\epsilon C_\rho} L^{\nu, \epsilon} f - \mathcal{H}^* \Delta_\nu \mathcal{H}f \right\|_{C^0(M, \mathbb{R})} \right) = 0. \end{aligned}$$



# Bibliography

- [1] R. A. Adams and J. J. F. Fournier. *Sobolev spaces*, volume 140. Academic press, 2003.
- [2] A. Agarwal and B. Triggs. Tracking articulated motion with piecewise learned dynamical models. In *European Conference on Computer Vision*, volume 3, pages 54–65, 2004.
- [3] N. Alon. Eigenvalues and expanders. *Combinatorica*, 6:83–96, 1986.
- [4] C. J. Alpert, A. B. Kahng, and S. Z. Yao. Spectral partitioning with multiple eigenvectors. *Discrete Applied Mathematics*, 90(1):3–26, 1999.
- [5] W. N. Anderson Jr and T. D. Morley. Eigenvalues of the Laplacian of a graph. *Linear and Multilinear Algebra*, 18(2):141–145, 1985.
- [6] H. Aref. The development of chaotic advection. *Physics of Fluids (1994-present)*, 14(4):1315–1325, 2002.
- [7] T. Aubin. *A course in differential geometry*, volume 27. American Mathematical Society, 2001.
- [8] A. S. Bandibra, A. singer, and D. A. Spielman. A cheeger inequality for graph connection Laplacian. *SIAM J. Matrix Anal. Appl.*, 34(4):1661 – 1630, 2013.
- [9] R. Banisch and P. Koltai. Understanding the geometry of transport: Diffusion maps for Lagrangian trajectory data unravel coherent sets. *ArXiv:1603.04709v2*, 2016.
- [10] M. Belkin and P. Niyogi. Laplacian eigenmaps and spectral techniques for embedding and clustering. *NIPS*, 14(14):585–591, 2001.
- [11] M. Belkin and P. Niyogi. Laplacian eigenmaps for dimensionality reduction and data representation. *Neural computation*, 15(6):1373–1396, 2003.

- [12] M. Belkin and P. Niyogi. Towards a theoretical foundation for Laplacian-based manifold methods. *Journal of Computer and System Sciences*, 74(8):1289–1308, 2008.
- [13] P. H. Bérard. *Spectral geometry: direct and inverse problems*, volume 1207. Springer, 2006.
- [14] M. Bernstein, V. De Silva, J. C. Langford, and J. B. Tenenbaum. Graph approximations to geodesics on embedded manifolds. Technical report, Technical report, Department of Psychology, Stanford University, 2000.
- [15] T. Biyikoğlu, P. F. Stadler, and J. Leydold. *Laplacian eigenvectors of graphs*. Springer, 2007.
- [16] A. Buluç, H. Meyerhenke, I. Safro, P. Sanders, and C. Schulz. Recent advances in graph partitioning. *arXiv:1311.3144v2*, 2013.
- [17] P. Buser. A note on the isoperimetric constant. *Annales scientifique de l'École Normale Supérieure*, 15(2):213–230, 1982.
- [18] L. Capogna, D. Donatella, S. D. Pals, and J. Tyson. *An introduction to the Heisenberg group and the sub-Riemannian isoperimetric problem*, volume 259. Springer Science & Business Media, 2007.
- [19] C. Carathéodory and E. Study. Zwei beweis des Satzes dass der Kreis unter alle Figuren gleichen Umgangs den gr ossten Inhalt hat. *Math. Ann.*, 68:133–144, 1909.
- [20] L. Cayton. Algorithms for manifold learning. *Univ. of California San Diego Tech. Rep*, pages 1–17, 2005.
- [21] A. K. Chan, M. D. Schlag, and J. Y. Zien. Spectral K-way ratio-cut partitioning and clustering. *IEEE Transaction on Computer-Aided Design of Integrated Circuits and Systems*, 13(9):1088–1096, 1994.
- [22] I. Chavel. *Eigenvalues in Riemannian geometry*, volume 115. Academic press, 1984.
- [23] I. Chavel. *Isoperimetric Inequalities: Differential geometric and analytic perspectives*, volume 145. Cambridge University Press, 2001.
- [24] J. Cheeger. A lower bound for the smallest eigenvalue of the Laplacian. *Proceedings of the Princeton conference in honor of Professor S. Bochner*, 1969.

- 
- [25] F. Chung. *Spectral Graph Theory*. American Mathematical Society, CBMS Regional Conference Series In Mathematics edition, 1996.
- [26] F. Chung. Laplacians and the cheeger inequality for directed graphs. *Annals of Combinatorics*, 9(1):1–19, 2005.
- [27] R. R. Coifman and M. J. Hirn. Diffusion maps fr changing data. *Applied and Computational Harmonic Analysis*, 36(1):79 – 107, 2014.
- [28] R. R. Coifman and S. Lafon. Diffusion maps. *Applied and computational harmonic analysis*, 21(1):5–30, 2006.
- [29] J. A. Costa and A. O. Hero. Geodesic entropic graphs for dimension and entropy estimation in manifold learning. *IEEE Transactions on Signal Processing*, 52(8):2210–2221, 2004.
- [30] D. Delling, A. V. Goldberg, T. Pajor, and R. F. Werneck. Customizable route planning. *International Symposium on Experimental Algorithms*, pages 376–387, 2011.
- [31] M. Dellnitz, G. Froyland, and S. Sertl. On the isolated spectrum of the Perron-Frobenius operator. *Nonlinearity*, 13(4):1171, 2000.
- [32] M. Dellnitz and O. Junge. On the approximation of complicated dynamical behavior. *SIAM Journal on Numerical Analysis*, 36(2):491–515, 1999.
- [33] G. Demirel, F. Vazquez, G. A. Böhme, and T. Gross. Moment-closure approximation for discrete adaptive networks. *Physica D*, 267:68–88, 2014.
- [34] P. Deuffhard, W. Huisinga, A. Fischer, and C. Schutte. Identification of almost invariant aggregates in reversible nearly uncoupled Markov chains. *Linear Algebra and its Applications*, 315(1):39–59, 2000.
- [35] M. P. do Carmo. *Differential geometry of curves and surfaces*, volume 2. Englewood Cliffs:Prentice-hall, 1976.
- [36] M. P. do Carmo. *Riemannian geometry*. 1992.
- [37] R. Doerner, B. Hubinger, W. Grossmann, and S. Thomae. Stable manifolds and predictability of dynamical systems. *Chaos Solitons and Fractals*, 10(11):1759–1782, 1999.
- [38] W. E. Donath and A. J. Hoffman. Lower bounds for the partitioning of graphs. *IBM Journal of Research and Development*, 17(5):420–425, 1973.

- [39] D. Ediger, J. Riedy, D. A. Bader, and H. Meyerhenke. Tracking structure of streaming social networks. *Parallel and Distributed Processing Workshops and PhD Forum (IPDPSW). 2011 IEEE International Symposium on*, pages 1691–1699, 2011.
- [40] B. Erem, P. Stovicek, and D. H. Brooks. Manifold learning for analysis of low-order nonlinear dynamics in high-dimensional electrocardiographic signals. In *Biomedical Imaging (ISBI), 2012 9th IEEE International Symposium on*, pages 844–847. IEEE, 2012.
- [41] H. Federer and W. H. Fleming. Normal and intergral currents. *Annals of Mathematics*, 72(3):458–520, 1960.
- [42] D. J. Fenn, M. A. Porter, M. McDonald, S. Williams, N. F. Johnson, and N. S. Jones. Dynamic communities in multichannel data: An application to the foreign exchange market during the 2007–2008 credit crisis. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 19(3):033119, 2009.
- [43] M. Fiedler. Algebraic connectivity of graphs. *Czechoslovak Mathematics*, 23(2):698–305, 1973.
- [44] M. Fiedler. A property of eigenvectors of nonnegative symmetric matrices and its applications to graph theory. *Czechoslovak Mathematics*, 25:619–633, 1975.
- [45] P.-O. Fjällström. *Algorithms for graph partitioning: A survey*, volume 3. Link oping: Link oping University Electronic Press, 1998.
- [46] J. H. Fowler and N. A. Christakis. Dynamic spread of happiness in a large social network: Longitudinal analysis over 20 years in the Framingham Heart Study. *British Medical Journal*, 337:1–9, 2008.
- [47] G. Froyland. Statistically optimal almost-invariant sets. *Physica D: Nonlinear Phenomena*, 200(3):205 – 219, 2005.
- [48] G. Froyland. An analytic framework for identifying finite-time coherent sets in time-dependent dynamical systems. *Physica D: Nonlinear Phenomena*, 250:1–19, 2013.
- [49] G. Froyland. Dynamic isoperimetry and the geometry of Lagrangian coherent structures. *Nonlinearity*, 28(10):3587–3622, 2015.

- 
- [50] G. Froyland and M. Dellnitz. Detecting and locating near-optimal almost-invariant sets and cycles. *SIAM Journal on Scientific Computing*, 24(6):1893–1863, 2003.
- [51] G. Froyland and O. Junge. On fast computation of finite-time coherent sets using radial basis functions. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 25(8):087409, 2015.
- [52] G. Froyland and O. Junge. Robust fem-based extraction of finite-time coherent sets using scattered, sparse, and incomplete trajectories. *arXiv preprint arXiv:1705.03640*, 2017.
- [53] G. Froyland and E. Kwok. Partitions of networks that are robust to vertex permutation dynamics. *Special Matrices*, 3(1):22–42, 2015.
- [54] G. Froyland and E. Kwok. A dynamic Laplacian for identifying Lagrangian coherent structures on weighted Riemannian manifolds. *Journal of Nonlinear Science*, 2017.
- [55] G. Froyland, S. Lloyd, and A. Quas. Coherent structures and isolated spectrum for Perron-Frobenius cocycles. *Ergodic Theory and Dynamical Systems*, 30(3):729–456, 2010.
- [56] G. Froyland and K. Padberg-Gehle. Almost-invariant sets and invariant manifolds—connecting probabilistic and geometric descriptions of coherent structures in flows. *Physica D: Nonlinear Phenomena*, 238(16):1507–1523, 2009.
- [57] G. Froyland and K. Padberg-Gehle. A rough-and-ready cluster-based approach for extracting finite-time coherent sets from sparse and incomplete trajectory data. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 25(8):087406, 2015.
- [58] G. Froyland, N. Santitissadeekorn, and A. Monahan. Transport in time-dependent dynamical systems: Finite-time coherent sets. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 20(4):043116, 2010.
- [59] M. R. Garey and D. S. Johnson. *Computers and Intractability: A Guide to Theory of NP-completeness*. W.H Freeman and Co, 1970.
- [60] D. Giannakis and A. J. Majda. Nonlinear laplacian spectral analysis for time series with intermittency and low-frequency variability. *Proceedings of the National Academy of Sciences*, 109(7):2222–2227, 2012.

- [61] D. Gilbarg and N. S. Trudinger. *Elliptic partial differential equations of second order*. Springer-Verlag Berlin Heidelberg New York, 1977.
- [62] V. Gol'dshtein and A. Ukhlov. Weighted Sobolev spaces and embedding theorems. arXiv:math/0703725v4[math.FA].
- [63] E. Gouillart, J.-L. Thiffeault, and M. D. Finn. Topological mixing with ghost rods. *Physical Review E*, 73(3):036311, 2006.
- [64] L. Grady and E. Schwartz. Isoperimetric graph partitioning for image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(3):469–475, 2006.
- [65] A. Gray. The volume of a small geodesic ball of a Riemannian manifold. *The Michigan Mathematical Journal*, 20(4):329–344, 1974.
- [66] A. Grigor'yan. Heat kernels on weighted manifolds and applications. *Contemporary Mathematics*, 398:93–191, 2006.
- [67] K. M. Hall. An r-dimensional quadratic placement algorithm. *Management Science*, 17(3):219–229, 1970.
- [68] G. Haller. Lagrangian coherent structures from approximate velocity data. *Physics of Fluids (1994-present)*, 14(6):1851–1861, 2002.
- [69] G. Haller and F. J. Beron-Vera. Geodesic theory of transport barriers in two-dimensional flows. *Physica D: Nonlinear Phenomena*, 241(20):1680–1702, 2012.
- [70] X. He, S. Yan, Y. Hu, P. Niyogi, and H. J. Zhang. Face recognition using laplacianfaces. *IPSN Transactions on Computer Vision and Applications*, 1:83–94, 2009.
- [71] E. Hebey. *Sobolev spaces on Riemannian manifolds*, volume 1635. Springer Science & Business Media, 1996.
- [72] B. Hendrickson and T. G. Kolda. Graph partitioning models for parallel computing. *Parallel Computing*, 26:1519–1534, 2000.
- [73] W. Hoeffding. Probability inequality for sums of bounded random variables. *Journal of the American statistical association*, 58(301):13–30, 1963.
- [74] P. Holme and J. Saramäki. Temporal networks. *Physics report(s)*, 519(3):97–125, 2012.

- 
- [75] D. F. Hong, N. Yokoya, and X. X. Zhu. Local manifold learning with robust neighbors selection for hyperspectral dimensionality reduction. In *Geoscience and Remote Sensing Symposium (IGARSS), 2016 IEEE International*, pages 40–43. IEEE, 2016.
- [76] R. A. Horn and C. A. Johnson. *Matrix Analysis*. Cambridge, 1990.
- [77] W. Jiang, N. Li, H. Yin, and Y. Chai. *An Improved Laplacian Eigenmaps Algorithm for Nonlinear Dimensionality Reduction*, pages 403–413. Springer Berlin Heidelberg, 2016.
- [78] J. Jost. *Riemannian geometry and geometric analysis*. Springer Science & Business Media, 2008.
- [79] K. Judd, M. Small, and T. Stemler. What exactly are the properties of scale-free and other networks? *EPL (Europhysics Letters)*, 103(5):58004, 2013.
- [80] B. Kawohl and F. Vladislav. Isoperimetric estimates for the first eigenvalue of the p-Laplace operator and the Cheeger constant. *Comment. Math. Univ. Carolin*, 44(4):659–667, 2003.
- [81] V. Kwatra, A. Schoödl, Essa, G. Turk, and A. Bobick. Graphcut textures: Image and video synthesis using graph cuts. In *ACM Transactions on Graphics (ToG)*, 22:277–286, 2003.
- [82] R. Lambiotte, J.-C. Delvenne, and M. Barahona. Laplacian dynamics and multiscale modular structure in networks. *arXiv preprint arXiv:0812.1770*, 2008.
- [83] K. Li, M. Small, and X. Fu. Generation of clusters in complex dynamical networks via pinning control. *Journal of Physics A: Mathematical and Theoretical*, 41(50):505101, 2008.
- [84] U. Luxburg. A tutorial on spectral clustering. *Statistics and Computing*, 17(4):395–416, 2007.
- [85] T. Ma and E. M. Bollt. Differential geometry perspective of shape coherence and curvature evolution by finite-time nonhyperbolic splitting. *SIAM Journal on Applied Dynamical Systems*, 13(3):1106–1136, 2014.
- [86] T. Ma and E. M. Bollt. Shape coherence and finite-time curvature evolution. *International Journal of Bifurcation and Chaos*, 25(05):1550076, 2015.

- [87] G. Mathew, I. Mezić, and L. Petzold. A multiscale measure for mixing. *Physica D*, 211(1):23–46, 2005.
- [88] R. C. McOwen. *Partial differential equations: Methods and applications*. Prentice Hall, 1996.
- [89] J. D. Meiss. Symplectic maps, variational principles, and transport. *Reviews of Modern Physics*, 64(3):795–848, 1992.
- [90] V. D. Milman and G. Schechtman. *Asymptotic theory of finite dimensional normed spaces: Isoperimetric inequalities in Riemannian manifolds*. Springer, 2009.
- [91] B. Mohar. Isoperimetric number of graphs. *Journal of Comb. Theory*, 47:274–291, 1989.
- [92] P. J. Mucha, T. Richardson, K. Macon, M. A. Porter, and J. P. Onnela. Community structure in time-dependent, multiscale, multiple networks. *Science*, 328(5980):876–878, 2010.
- [93] B. Nadler, S. Lafon, R. R. Coifman, and I. G. Kevrekidis. Diffusion maps, spectral clustering and reaction coordinates of dynamical systems. *Applied and Computational Harmonic Analysis*, 21(1):113–127, 2006.
- [94] D. Oettinger, D. Blazeviski, and G. Haller. Global variational approach to elliptic transport barriers in three dimensions. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 26(3):033114, 2016.
- [95] R. Osserman. The isoperimetric inequality. *Bulletin of the American Mathematical Society*, 84(6):1182–1238, 1978.
- [96] J. M. Ottino. *The kinematics of mixing: Stretching, chaos, and transport.*, volume 3. Cambridge University Press, 1989.
- [97] L. E. Payne. Isoperimetric inequalities and their applications. *SIAM review*, 9(2):453–488, 1967.
- [98] R. T. Pierrehumbert. Chaotic mixing of tracer and vorticity by modulated travelling Rossby waves. *Geophysical & Astrophysical Fluid Dynamics*, 84(1-4):285–319, 1991.
- [99] R. T. Pierrehumbert and H. Yang. Global chaotic mixing on isentropic surfaces. *Journal of the atmospheric sciences*, 50(15):2462–2480, 1993.

- 
- [100] G. Piyush, S. D. Ross, M. A. Stremler, and P. Kumar. Topological chaos, braiding and bifurcation of almost-cyclic sets. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 22(4):043135, 2012.
- [101] M. Reuter, F. E. Wolter, and N. Peinecke. Laplace-Beltrami spectra as ‘Shape-DNA’ of surfaces and solids. *Computer-Aided Design*, 38(4):342–366, 2006.
- [102] V. Rom-Kedar, A. Leonard, and S. Wiggins. An analytical study of transport, mixing and chaos in an unsteady vortical flow. *Journal of Fluid Mechanics*, 214:347 – 394, 1990.
- [103] S. Rosenberg. *The Laplacian on a Riemannian manifold: An introduction to analysis on manifolds*. Number 31. Cambridge University Press, 1997.
- [104] W. Rudin. *Real and complex analysis*. Tata McGraw-Hill Education, 1987.
- [105] R. M. Rustamov. Laplace-Beltrami eigenfunctions for deformation invariant shape representation. In *Proceedings of the fifth Eurographics symposium on Geometry processing*, pages 225–233. Eurographics Association, 2007.
- [106] T. Sakai. On Riemannian manifolds admitting a function whose gradient is of constant norm. *Kodai Mathematical Journal*, 19(1):39 – 51, 1996.
- [107] N. Santitissadeekorn and E. Bollt. Identifying stochastic basin hopping by partitioning with graph modularity. *Physica D: Nonlinear Phenomena*, 231(2):95–107, 2007.
- [108] J. Scott. *Social Network Analysis: A Handbook*. SAGE, 2000.
- [109] S. C. Shadden, F. Lekien, and J. E. Marsden. Definition and properties of Lagrangian coherent structures from finite-time Lyapunov exponents in two-dimensional aperiodic flows. *Physica D: Nonlinear Phenomena*, 212(3):271–304, 2005.
- [110] J. Shi and J. Malik. Normalized cuts and image segmentation. *IEEE Transaction on pattern analysis and machine intelligence*, 22(8):888–905, 2000.
- [111] M. Small, Y. Li, T. Stemler, and K. Judd. Growing optimal scale-free networks via likelihood. *Physical Review E*, 91(4):042801, 2015.
- [112] D. A. Spielman. Spectral graph theory and its applications. In *Foundations of Computer Science, 2007. FOCS’07. 48th Annual IEEE Symposium on*, pages 29–38. IEEE, 2007.

- [113] M. Spivak. *Calculus on manifolds*, volume 1. New York: WA Benjamin, 1965.
- [114] E. M. Stein and T. S. Murphy. *Harmonic analysis: Real-variable methods, orthogonality, and oscillatory integrals*, volume 3. Princeton University Press, 1993.
- [115] J. Steiner. Einfacher beweis der isoperimetrischen hauptsatze. *reine angew. Math.*, 8:281–296, 1838.
- [116] J. C. Strikwerda. *Finite difference schemes and partial differential equations*. Siam, 2004.
- [117] J.-L. Thiffeault. Advection–diffusion in Lagrangian coordinates. *Physics Letters A*, 309(5):415–422, 2003.
- [118] J. L. Thiffeault. Using multiscale norm to quantify mixing and transport. *Nonlinearity*, 25(2):R1, 2012.
- [119] B. O. Turesson. *Nonlinear potential theory and weighted Sobolev spaces*, volume 1736. Springer Science & Business Media, 2000.
- [120] S. M. Ulam. *A collection of mathematical problems*, volume 8. Interscience Publishers, 1960.
- [121] J. M. Wang, D. J. Fleet, and A. Hertzmann. Gaussian process dynamical models for human motion. *IEEE transactions on pattern analysis and machine intelligence*, 30(2):283–298, 2008.
- [122] K. Q. Weinberger and L. K. Saul. Unsupervised learning of image manifolds by semidefinite programming. *International Journal of Computer Vision*, 70(1):77–90, 2006.
- [123] S. Wiggins. The dynamical systems approach to Lagrangian transport in oceanic flows. *Annu. Rev. Fluid Mech.*, 37:295 – 328, 2005.