Appl. Comput. Harmon. Anal. 56 (2022) 189–249

Contents lists available at ScienceDirect

Applied and Computational Harmonic Analysis

www.elsevier.com/locate/acha

Spectral analysis of weighted Laplacians arising in data clustering

Franca Hoffmann [b,*], Bamdad Hosseini [b], Assad A. Oberai [a], Andrew M. Stuart [b]

[a] *Department of Aerospace and Mechanical Engineering, University of Southern California, Los Angeles, CA 90089, USA*
[b] *Computing and Mathematical Sciences, Caltech, Pasadena, CA, USA*

ARTICLE INFO

Article history:
Received 13 September 2019
Received in revised form 18 April 2021
Accepted 30 July 2021
Available online 1 September 2021
Communicated by Radu Balan

MSC:
47A75
62H30
68T10
35B20
05C50

Keywords:
Spectral clustering
Graph Laplacian
Large data limits
Elliptic differential operators
Perturbation analysis
Spectral gap
Differential geometry

ABSTRACT

Graph Laplacians computed from weighted adjacency matrices are widely used to identify geometric structure in data, and clusters in particular; their spectral properties play a central role in a number of unsupervised and semi-supervised learning algorithms. When suitably scaled, graph Laplacians approach limiting continuum operators in the large data limit. Studying these limiting operators, therefore, sheds light on learning algorithms. This paper is devoted to the study of a parameterized family of divergence form elliptic operators that arise as the large data limit of graph Laplacians. The link between a three-parameter family of graph Laplacians and a three-parameter family of differential operators is explained. The spectral properties of these differential operators are analyzed in the situation where the data comprises of two nearly separated clusters, in a sense which is made precise. In particular, we investigate how the spectral gap depends on the three parameters entering the graph Laplacian, and on a parameter measuring the size of the perturbation from the perfectly clustered case. Numerical results are presented which exemplify the analysis and which extend it in the following ways: the computations study situations in which there are two nearly separated clusters, but which violate the assumptions used in our theory; situations in which more than two clusters are present, also going beyond our theory; and situations which demonstrate the relevance of our studies of differential operators for the understanding of finite data problems via the graph Laplacian. The findings provide insight into parameter choices made in learning algorithms which are based on weighted adjacency matrices; they also provide the basis for analysis of the consistency of various unsupervised and semi-supervised learning algorithms, in the large data limit.

* Corresponding author.
E-mail addresses: fkoh@caltech.edu (F. Hoffmann), bamdadh@caltech.edu (B. Hosseini), aoberai@usc.edu (A.A. Oberai), astuart@caltech.edu (A.M. Stuart).

https://doi.org/10.1016/j.acha.2021.07.004
1063-5203/

1. Introduction

1.1. Overview

This article presents a spectral analysis of differential operators of the form

$$\begin{cases} \mathcal{L}u := -\frac{1}{\varrho^p}\mathrm{div}\left(\varrho^q \nabla\left(\frac{u}{\varrho^r}\right)\right), & \text{in } \mathcal{Z}, \\ \varrho^q \frac{\partial}{\partial n}\left(\frac{u}{\varrho^r}\right) = 0, & \text{on } \partial\mathcal{Z}, \end{cases} \tag{1.1}$$

for parameters $p, q, r \in \mathbb{R}$ fixed. The analysis is focused on the situation where the density ϱ concentrates on two disjoint connected sets (clusters), and numerical results extend our conclusions to multiple clusters and to more general two cluster data densities ϱ not covered by our analysis. Our motivation is to understand a range of algorithms which learn about geometric information in data, and clusters in particular, by means of graph Laplacians constructed from adjacency matrices whose edge weights reflect affinities between data points at each vertex. Operators of the form (1.1) arise as a large data limit of graph Laplacian operators of the form

$$L_N := \begin{cases} D_N^{\frac{1-p}{q-1}} \left(D_N - W_N\right) D_N^{-\frac{r}{q-1}}, & \text{if } q \neq 1, \\ D_N - W_N, & \text{if } q = 1, \end{cases} \tag{1.2}$$

where the symmetric weighted adjacency matrix $W_N = W_N(q)$ is constructed via a suitably reweighted kernel capturing the similarities between discrete data points and $D_N = D_N(q)$ is an associated weighted degree matrix (see Subsection 5.1 for precise definitions of these matrices).

The three primary contributions of this paper are as follows:

1. Under assumptions on ϱ capturing the notion of data approximately clustered into two sets, we study the low lying spectrum of $\mathcal{L}$, the corresponding eigenfunctions and their dependence on (p, q, r); amongst several results concerning this dependence, we elucidate the special properties of the parametric family $q = p + r$ for clustering tasks, and we refer to $\mathcal{L}$ and L_N as *balanced* in this case.
2. We present numerical experiments which exemplify the analysis in both the continuum and discrete regimes, leading to conjectures concerning aspects of our analysis which are not sharp, and extending our understanding to mixture models and to multiple clusters, situations not covered by the analysis.
3. We explain how $\mathcal{L}$ arises from L_N, and provide numerical simulations illustrating that the characteristic behavior identified for the limiting operators $\mathcal{L}$ in point 1 also manifests in the finite data setting when using L_N.

Understanding the spectral behavior of graph Laplacian operators, their continuum counterparts and links between them are crucial steps in the consistency analysis for algorithms such as spectral clustering. Our work addresses primarily the spectral behavior of the continuum formulations but may, in future works, form a useful foundation for analysis of consistency. Our results may also be of independent interest in the spectral theory of elliptic differential operators. Subsection 1.2 is devoted to the background to our work, and a literature review. In Subsection 1.3 we describe the three contributions above in detail; Subsection 1.4 contains illustrative numerical experiments which demonstrate our contributions; and Subsection 1.5 concludes the introduction with an outline of the paper, by section.

1.2. Literature review

Clustering is a fundamental task in data analysis and in unsupervised and semi-supervised learning in particular; algorithms in these areas seek to detect clusters, and more generally coarse structures, geometry and patterns in data. Our focus is on Euclidean data. Our starting point is a dataset $X = \{x_1, ..., x_N\}$ comprising of N points $x_i \in \mathbb{R}^d$, assumed to be drawn i.i.d. from a (typically unknown) probability distribution with (Lebesgue) density ϱ. The goal of clustering algorithms is to split X into meaningful clusters. Many such algorithms proceed as follows: The data points x_i are associated with the vertices of a graph and a weighted adjacency matrix W_N, measuring affinities between data points, is defined on the edges of the graph. From this matrix, and from a weighted diagonal degree matrix D_N found from summing edge weights originating from a given node, various graph Laplacian matrices L_N can be defined. The success of clustering algorithms is closely tied to the spectrum of L_N. At a high level, k clusters will manifest in k small eigenvalues of L_N, and then a spectral gap; and the k associated eigenvectors will have geometry which encodes the clusters. Unsupervised learning leverages this structure to identify clusters [4,33,41,43] and semi-supervised learning uses this structure as prior information which is enhanced by labeled data [8,9,47]. It is thus of considerable interest to study the spectral properties of L_N, and the dependence of the spectral properties on the data and on the design parameters chosen in constructing L_N.

The operator L_N in (1.2) corresponds to different normalizations of the graph Laplacian. A number of special cases within this general class arise frequently in the implementation of unsupervised and supervised learning algorithms. The *unnormalized graph Laplacian* refers to the choice $(p, q, r) = (1, 2, 0)$, giving the symmetric matrix $L_N = D_N - W_N$; another popular choice is the *normalized graph Laplacian* where $(p, q, r) = (3/2, 2, 1/2)$; the choice $(p, q, r) = (2, 2, 0)$ also gives a widely used normalized operator. The graph Laplacian for $(p, q, r) = (3/2, 2, 1/2)$ is symmetric and studied in [23,33,36,38,43,44], whereas the choice $(2, 2, 0)$ gives an operator that is not symmetric, but can be interpreted as a transition probability of a random walk on a graph [14,38]. A number of other choices for (p, q, r) appear in the literature. For example, the spectrum of the graph Laplacian with $(1, 2, 0)$ is related to the ratio cut, whereas $(2, 2, 0)$ is connected to the Ncut problem. The success of the spectral clustering procedure for the graph Laplacian with parameters $(1, 1, 0)$ was investigated in [21] in the setting of non-parametric mixture models; in this case the Dirichlet energy with respect to the natural density weighted L^2 inner-product is linear in ϱ. In [14,45], general choices of $p = q \geq 0$ and $r = 0$ are investigated in the context of diffusion maps with [45] presenting sharp pointwise error bounds on the spectrum as well as norm convergence of L_N to $\mathcal{L}$. In this case, the limiting operator $\mathcal{L}$ is the generator of a reversible diffusion process, a connection first established in the celebrated paper [14] by Coifman and Lafon.

Whilst many different normalizations of the graph Laplacian have been used for a variety of data analysis tasks, a thorough understanding of the advantages and disadvantages of different parameter choices is still lacking. The papers [43,44] contain comparisons between the normalized, unnormalized and random walk Laplacians. But, to the best of our knowledge, there is a gap in the current literature concerning a systematic understanding of the effects of the entire family of weighted graph Laplacian matrices L_N depending on the family of parameters (p, q, r). Of particular interest is the case where N is large, relevant in large data applications, and in [44] the authors showed that the normalized and random walk Laplacians give consistent spectral clustering as opposed to the unnormalized Laplacian operator in this large N limit. This behavior is attributed to different integral operators to which the normalized and unnormalized Laplacians converge. The normalized Laplacian converges to a compact perturbation of the identity with a discrete spectrum while it is demonstrated that the unnormalized Laplacian may not possess a purely discrete spectrum.

The large data limit convergence of graph Laplacians to integral or differential operators has been the subject of many recent studies including [5,6,13,23,20,24,36,39,40,44,45]. The point of departure in these papers is a kernel η defined on $\mathbb{R}^d \times \mathbb{R}^d$, from which the weighted adjacency matrix W_N defined on the edges of a graph is constructed. In [5,6,25,26,36,39,44] the authors fix a kernel and let $N \to \infty$ obtaining an

integral operator as the limit of graph Laplacians. These limiting integral operators are dependent on the kernel η and subsequently the results of these articles also depend on the choice of the kernel. In real-world applications, fixing the kernel η as N grows results in dense graph Laplacian matrices that are impractical to compute with. To alleviate this issue either k-nearest neighbor (k-NN) graphs are employed or the bandwidth of η is tuned to control the sparsity of the graph Laplacian. Motivated by these observations, more recent articles [12,13,22,23,20,24,40,45] consider the joint limit as $N \to \infty$ and the width of the kernel η vanishes sufficiently slowly thereby controlling the local connectivity of the graph and the sparsity of the graph Laplacian. It then follows that in taking this joint limit graph Laplacian matrices L_N converge to differential operators of a similar form to our $\mathcal{L}$ operator; under this type of limiting procedure the resulting differential operator is independent of the weight kernel η, up to scaling.

The aforementioned articles suggest the potential for further analysis of the continuum limits of graph Laplacians as a means to advance our understanding of clustering algorithms on finite but large data sets. Such continuum approaches, often refereed to as *population level analyses*, proceed by studying graph Laplacian operators and subsequently spectral clustering algorithms in the continuum regime [21,36,39]. The continuum analysis may then be extended to the finite data setting using discrete-to-continuum approximation results such as those in [13,20,36,45]. We employ the same perspective in this work, focusing primarily on the analysis of the continuum operators and providing numerical experiments and formal calculations demonstrating the relevance of the continuum analysis to finite data settings. We note that the paper [34] studies consistency of spectral clustering for finite graph problems, and that similar ideas from linear algebra are used to study large data limits in [15], albeit with very restrictive assumptions on the clusters; no limiting operator is employed, or identified, in [15].

We also note that mathematical studies which are conceptually similar to the spectral analysis that we present here have been prevalent in the study of metastability in chemically reacting systems for some time; see [16,17,28,37] and the references therein for applications. This body of work has led to very subtle and deep analyses of the generators of Markov processes [10,11]; this analysis might, in principle, be used to extend some of the work undertaken here to a wider range of sampling densities.

Finally, the tools developed in this paper may be used to study consistency of semi-supervised learning algorithms in the continuum setting. In particular, we provide the spectral perturbation results needed to generalize the work in [27], which studies consistency of graph-based semi-supervised learning algorithms for finite N and using the graph Laplacian L_N, to the large data limit where $N \to \infty$ and L_N is replaced by $\mathcal{L}$.

1.3. Our contributions

We now detail the three contributions outlined in Subsection 1.1. Contribution 1 is summarized in our main theoretical result characterizing the low-lying spectrum of $\mathcal{L}$ and the effect of the (p, q, r) parameters; Contribution 2 extends our theoretical analyses by various numerical experiments (i) in the unbalanced regime where $q \neq p + r$, revealing that some of our bounds on the eigenvalues of $\mathcal{L}$ can be sharpened, and (ii) to the setting of multiple clusters and more general data densities ϱ, suggesting that the theory provided under Contribution 1 reveals fundamental concepts that hold in more generality than the specific setting considered in Contribution 1; Contribution 3 combines formal calculations and numerical experiments to reveal the relationship between the (p, q, r) parameterized family of differential operators $\mathcal{L}$ and various weightings of discrete graph Laplacians L_N.

1.3.1. Contribution 1

Let us define the notion of a *perfectly separated* density. Let $\mathcal{Z} \subset \mathbb{R}^d$ be bounded and ϱ_0 be a (Lebesgue) probability density with support $\mathcal{Z}' \subset \mathcal{Z}$ strictly contained in $\mathcal{Z}$ and concentrated on two disjoint subsets $\mathcal{Z}^+$ and $\mathcal{Z}^-$ of $\mathcal{Z}$; that is, $\mathcal{Z}' = \mathcal{Z}^+ \cup \mathcal{Z}^-$ and $\mathcal{Z}^+ \cap \mathcal{Z}^- = \emptyset$. We refer to $\mathcal{Z}^\pm$ as clusters, and denote the

operator of the form (1.1) based on ϱ_0 by $\mathcal{L}_0$. Consequently, a *nearly separated* density comprises a class of smooth densities ϱ_ϵ that are $\mathcal{O}(\epsilon)$ perturbations of the perfectly separated case ϱ_0, with density supported everywhere on $\mathcal{Z}$ and such that $\varrho_\epsilon = C\epsilon$ away from $\mathcal{Z}'$ with $C > 0$ a constant[1]; we define this concept precisely in Section 2.2.1. We denote the operator of the form (1.1) based on ϱ_ϵ by $\mathcal{L}_\epsilon$. To this end, our main theoretical result characterizes the low-lying spectrum of $\mathcal{L}_\epsilon$ in the nearly separated regime.

Main Result 1.1. *Assume $q > 0$ and $p + r > 0$.*

(i) The first eigenpair of $\mathcal{L}_\epsilon$ is given by

$$\sigma_{1,\epsilon} = 0\,, \qquad \varphi_{1,\epsilon} = \frac{1}{|\mathcal{Z}|^{1/2}_{\varrho_\epsilon^{p+r}}} \varrho_\epsilon^r(x)\mathbf{1}_{\mathcal{Z}}(x), \qquad \forall x \in \mathcal{Z}$$

where $|\mathcal{Z}|_{\varrho_\epsilon^{p+r}} := \int_{\mathcal{Z}} \varrho_\epsilon^{p+r}(x)dx$.

(ii) The second eigenvalue scales as $\sigma_{2,\epsilon} = \mathcal{O}(\epsilon^q)$ and the corresponding eigenvector is given, approximately in a density weighted L^2 space, by the formula

$$\varphi_{2,\epsilon} \approx \frac{1}{|\mathcal{Z}'|^{1/2}_{\varrho_\epsilon^{p+r}}} \varrho_\epsilon^r(x)\big(\mathbf{1}_{\mathcal{Z}^+}(x) - \mathbf{1}_{\mathcal{Z}^-}(x)\big), \qquad \forall x \in \mathcal{Z}\,. \tag{1.3}$$

(iii) The behavior of the third eigenvalue $\sigma_{3,\epsilon}$ varies depending on the relationship between the parameters q and $p + r$:

- *if $p + r < q < 2(p + r)$, then a spectral ratio gap manifests with $\sigma_{2,\epsilon}/\sigma_{3,\epsilon} = \mathcal{O}(\epsilon^{2(p+r)-q})$ as $\epsilon \to 0$;*
- *if $q = p + r$, then $\sigma_{3,\epsilon} \asymp 1$ and a uniform spectral gap manifests, i.e., $\sigma_{3,\epsilon} - \sigma_{2,\epsilon} \asymp 1$ and $\sigma_{2,\epsilon}/\sigma_{3,\epsilon} = \mathcal{O}(\epsilon^q)$ as $\epsilon \to 0$;*
- *if $q < p + r < 2q$, then a spectral ratio gap manifests with $\sigma_{2,\epsilon}/\sigma_{3,\epsilon} = \mathcal{O}(\epsilon^{2q-(p+r)})$ as $\epsilon \to 0$.*

We precisely state this result, with fully detailed assumptions, in Section 3; the statement is comprised of a combination of theorems and corollaries. Part (i) is contained in Theorem 3.2(i) while part (ii) follows by combining Theorem 3.2(ii) with Theorem 3.4. Finally part (iii) is encompassed by Corollary 3.3. A roadmap of the proofs of these results is explained in Section 3 with the detailed proofs postponed to Section 6.

1.3.2. Contribution 2

We present detailed numerical experiments in Section 4 that both support our Main Result 1.1 and make two substantial extensions. These extensions sharpen our results in the unbalanced cases and extend our results to $K > 2$ clusters. In particular, our experiments in case $K = 2$ demonstrate that the rates for $\sigma_{2,\epsilon}/\sigma_{3,\epsilon}$ in Main Result 1.1(iii) are sharp in the balanced setting where $q = p + r$ but show clear evidence that the theoretical rates obtained in the unbalanced settings where $q \neq p + r$ are slower than the observed rates. The results obtained by combining Main Result 1.1 and this empirical improvement in the unbalanced case are then shown numerically to extend naturally to $K > 2$ clusters. For clarity we summarize these numerical results in the conjecture that follows.

Conjecture 1.2. *Suppose that the conditions of Main Result 1.1 are satisfied with the data density ϱ_ϵ concentrating on $K \geq 2$ clusters in the small ϵ limit. Then*

[1] We emphasize that this restrictive assumption on the form of the family of perturbed densities ϱ_ϵ considered here is made for the purposes of analysis; numerical experiments indicate that the conclusions are expected to hold beyond this restrictive class – see Remark 2.7.

$$\sigma_{K,\epsilon} \asymp \epsilon^q, \qquad \frac{\sigma_{K,\epsilon}}{\sigma_{K+1,\epsilon}} \asymp \epsilon^{\min\{q,p+r\}}.$$

Our numerical simulations in Section 4, and in particular Tables 4.1 to 4.3, suggest the above conjecture in the binary cluster setting that sharpens the decay rate of $\sigma_{2,\epsilon}/\sigma_{3,\epsilon}$ as a function of ϵ, in the unbalanced settings when $q \neq p + r$. Put simply, this conjecture states that when $K = 2$ and $q < p + r$ the third eigenvalue $\sigma_{3,\epsilon}$ exhibits similar behavior to the balanced setting where $q = p + r$ and hence a uniform gap in the spectrum manifests as $\epsilon \to 0$. However, when $q > p + r$ the third eigenvalue $\sigma_{3,\epsilon}$ vanishes like ϵ^{q-p-r} and a spectral ratio gap manifests. Moreover, if this conjecture holds then it allows us to sharpen the approximation error of the second eigenfunction $\varphi_{2,\epsilon}$ in Theorem 3.4, as this result heavily depends on a lower bound for $\sigma_{3,\epsilon}$. We attribute this discrepancy to the lower bound on $\sigma_{3,\epsilon}$ obtained in Theorem 3.2(iii) that in turn relies on a generalization of Cheeger's inequality from Appendix D.

1.3.3. Contribution 3

We demonstrate the relationship between the (p, q, r) dependent family of operators $\mathcal{L}$ in (1.1) showing how they arise as the limit of graph Laplacian matrices L_N of the form (1.2). Subsection 5.2 presents an informal limiting argument to identify the operator $\mathcal{L}$ by considering the large data N limit, followed by small kernel bandwidth δ limit of $L_N = L_N(\delta)$. Our informal calculations in Subsection 5.3 extend these arguments from Dirichlet energies to eigenvalue problems, and indicate that the spectrum of the matrix $C\delta^{-2}N^{2r-q}L_N$ converges to that of $\mathcal{L}$, for a suitable constant $C > 0$, as $(N, \delta^{-1}) \to \infty$. Our numerical experiments in Subsection 5.4 support these informal calculations, demonstrating the convergence of the eigenvalues of L_N to numerically computed eigenvalues of $\mathcal{L}$ for different choices of (p, q, r) and for two different types of mixture models. The numerical experiments and informal arguments are developed in the following setting: we assume that the data at the N vertices of the graph, $\{x_1, ..., x_N\}$, are sampled i.i.d. from the probability density ϱ and we suppose that the resulting weight matrix W_N is constructed using a kernel η_δ with the parameter $\delta > 0$ controlling the local connectivity of the vertices; see Subsection 5.1 for details.

To make a precise theory supporting these observations is beyond the scope of this paper, it is not a task which we are addressing in this work; it requires specification of the relationship between N and δ in the limiting process $(N, \delta^{-1}) \to \infty$. The convergence of L_N to $\mathcal{L}$ for specific choices of (p, q, r) has been established in the literature, and this issue was addressed in those papers. In particular, in [23] convergence of the spectrum of L_N Γ-converges to that of $\mathcal{L}$, and that the eigenfunctions of L_N converge to those of $\mathcal{L}$ in the TL^2 topology. More recently, the articles [13,20,45] further extend these results giving rates for the convergence of eigenvalues and eigenfunctions for $(p, q, r) = (1, 2, 0)$ and also for the convergence of L_N on k-NN graphs to $\mathcal{L}$ with $(p, q, r) = (1, 1 - 2/d, 0)$. We postulate that the methods of proof introduced in [13,20], and extensions to spectral convergence properties proved there, can be generalized to the (p, q, r)-dependent family of graph Laplacian operators introduced here; with the analysis for k-NN graphs departing from the proximity graphs considered here in particular in the construction of the discrete operator L_N and its normalization with different choices of (p, q, r). A full analysis and rigorous proof of the discrete-to-continuum connection is left for future work.

1.4. Illustrative numerical experiments

The contributions detailed in the preceding subsection demonstrate that the manner in which clustering is manifest in the spectral properties of the graph Laplacian depend subtly on the choice of the parameters (p, q, r). Making the balanced choice $q = p + r$ one obtains a family of operators whose second eigenvalue decays rapidly, while the gap between the second and third eigenvalues remains of order one as the parameter ϵ, measuring closeness to perfect clustering, decreases to zero; this uniform separation of second and third eigenvalues does not happen when $q > p + r$. Furthermore the form of the *Fiedler vector* (the

second eigenfunction), whilst always exhibiting the clusters present in the data, can have different behavior away from the clusters, depending on (p, q, r). We demonstrate these facts in Example 1.3, exemplifying Contributions 1 and 2. Additionally, Example 1.4 shows that our theory likely applies without the rather specific assumptions used to define clustering as mentioned in Contribution 2; furthermore, Example 1.4 illustrates that the spectral properties of the limiting operator $\mathcal{L}$ reflect the properties of the discrete graph Laplacian arising when $N < +\infty$ as outlined in our Contribution 3.

Example 1.3 *(Comparison of unnormalized and normalized graph Laplacians).* We study the spectral properties of operator $\mathcal{L}_\epsilon$ with parameter choices (p, q, r) given by $(1, 2, 0)$ and $(3/2, 2, 1/2)$ respectively, corresponding to the unnormalized and normalized graph Laplacians respectively. All our numerical experiments are for a data density ϱ_ϵ of the form (4.2) with two distinct clusters; see Fig. 1.1(a) for a plot of ϱ_ϵ with $\epsilon = 0.0125$.

In the unnormalized case $q > p+r$ it follows from our Main Result 1.1 that as $\epsilon \downarrow 0$ the second eigenvalue of $\mathcal{L}_\epsilon$ scales as ϵ^2 and that a spectral gap is present only in ratio form. In Fig. 1.1(b) we plot the second and third eigenvalues σ_2 and σ_3 against ϵ, on a log scale, and calculate best linear fits to the data; this demonstrates that they converge to zero like ϵ^2 and ϵ respectively, in agreement with our Main Result 1.1 (second eigenvalue) and the first component of Conjecture 1.2 (third eigenvalue). We also compute the second eigenfunction (Fielder vector) $\varphi_{2,\epsilon}$ shown in Fig. 1.1(d). Note that in this case the pointwise distance between $\varphi_{2,\epsilon}$ and the right hand side of (1.3) in Main Result 1.1(ii) is only small within the clusters; this reflects the fact that the weighted $L^2(\mathcal{Z}, \varrho_\epsilon^{p-r})$-norm arising in Theorem 3.4 for this choice of (p, q, r) is not sensitive to large pointwise values of functions in areas where ϱ_ϵ is small.

For comparison we now consider the normalized setting. For $q = p+r$ our Main Result 1.1 predicts that, as $\epsilon \downarrow 0$, there exists a uniform spectral gap between the first two eigenvalues of $\mathcal{L}_\epsilon$: for $(p, q, r) = (3/2, 2, 1/2)$, the second eigenvalue scales as ϵ^2 and the third is of order one with respect to ϵ. In Fig. 1.1(c) we plot the second and third eigenvalues of $\mathcal{L}_\epsilon$ against ϵ in that case, on a log-scale, and provide best fits to the data; the results support the theory. The corresponding Fiedler vector $\varphi_{2,\epsilon}$ is shown in Fig. 1.1(e). In this case $\varphi_{2,\epsilon}$ appears to converge pointwise to the right hand side of (1.3), in contrast to the unnormalized case.

It is well-known that the Fiedler vectors encode information on the clusters $\mathcal{Z}^\pm$ that we are trying to detect. They play a significant role in the context of spectral clustering and binary classification [43]. However, it is noteworthy that the Fiedler vectors in the unnormalized and normalized cases differ substantially within $\mathcal{Z} \setminus \mathcal{Z}'$: in the unnormalized case a smooth transition is made between $\mathcal{Z}^+$ and $\mathcal{Z}^-$, whereas in the normalized case abrupt transitions are made to near zero on the boundaries of $\mathcal{Z}^+$ and $\mathcal{Z}^-$. ◇

Since our primary motivation is data clustering, it is relevant to interpret our contributions in that context. In the following example we demonstrate that although our theory is developed under rather strict assumptions on the sampling density of the data and in the limit $N \to \infty$, our results concerning the dependence of spectral ratio gaps on the (p, q, r) parameters appear to generalize to mixture models that violate some of our assumptions. The mixture model assumption is a natural model for population level analysis of clustering algorithms and is considered in the articles [21,36]. It can be argued to be a more realistic data model for the density ϱ than the one for which our theory is developed and it is therefore of interest to demonstrate that our theory is predictive in this setting.

Example 1.4 *(Clustering a mixture model).* Consider the following mixture on the unit square

$$\varrho_\omega(t) := \frac{1}{2\omega}\left(1 - \exp\left(-\frac{1}{\omega}\right)\right)^{-1}\left[\exp\left(-\frac{t_1}{\omega}\right) + \exp\left(\frac{t_1 - 1}{\omega}\right)\right], \qquad t = (t_1, t_2)^T \in [0,1]^2. \tag{1.4}$$

This density is simply the mixture of two exponential distributions restricted to the unit interval $[0, 1]$ in the t_1 direction, with a uniform distribution in the t_2 direction; see Fig. 1.2(a). The parameter ω controls the

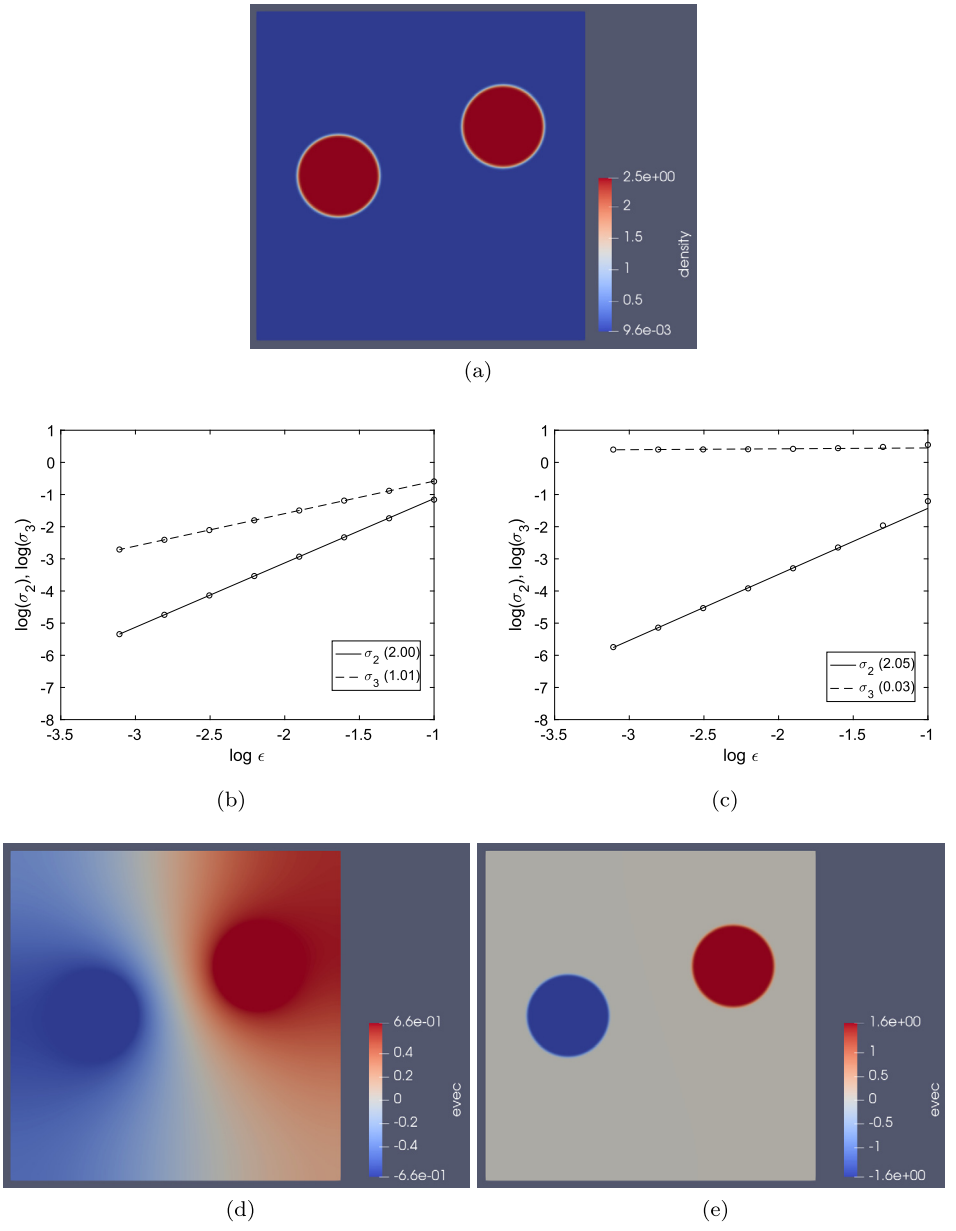


Fig. 1.1. (a) Plot of a density ϱ_ϵ of the form (4.2) with two distinct clusters for $\epsilon = 0.0125$. (b) Showing $\log(\sigma_2)$ and $\log(\sigma_3)$, the second and third eigenvalues of the unnormalized operator $\mathcal{L}_\epsilon$ with $(p, q, r) = (1, 2, 0)$ as functions of ϵ. Values in brackets in the legends indicate numerical slope of the lines. (c) Showing $\log(\sigma_2)$ and $\log(\sigma_3)$ for the normalized operator $\mathcal{L}_\epsilon$ for $(p, q, r) = (3/2, 2, 1/2)$, as functions of ϵ. (d) and (e) The Fiedler vector of $\mathcal{L}_\epsilon$ with $(p, q, r) = (1, 2, 0)$ and $(p, q, r) = (3/2, 2, 1/2)$ respectively for $\epsilon = 0.0125$.

overlap of the mixture components. This model clearly violates our assumptions on the density ϱ outlined in Section 2.2, most notably, (i) letting $\omega \to 0$ the density ϱ_ω concentrates on sets of measure zero as opposed to clusters $\mathcal{Z}^\pm$ of positive measure, and (ii) we cannot ensure that $\varrho_\omega = C\omega$ outside of clusters since the tails of the exponential components decay exponentially as we let $\omega \to 0$.

We generate N samples from ϱ_ω and construct a weighted proximity graph on this dataset using a weight kernel of width $\delta > 0$ as detailed in Subsection 5.4. We then proceed to define a discrete graph Laplacian L_N of the form (1.2) and compute the first four non-trivial eigenvalues $\sigma_{N,\delta}$ of this discrete operator (this notation for the eigenvalues is defined in Subsection 5.2). Fig. 1.2(b,c,d) show the variation of the first few eigenvalues as a function of ω for $N = 2^{13}$ vertices. We consider three choices of the (p, q, r) parameters,

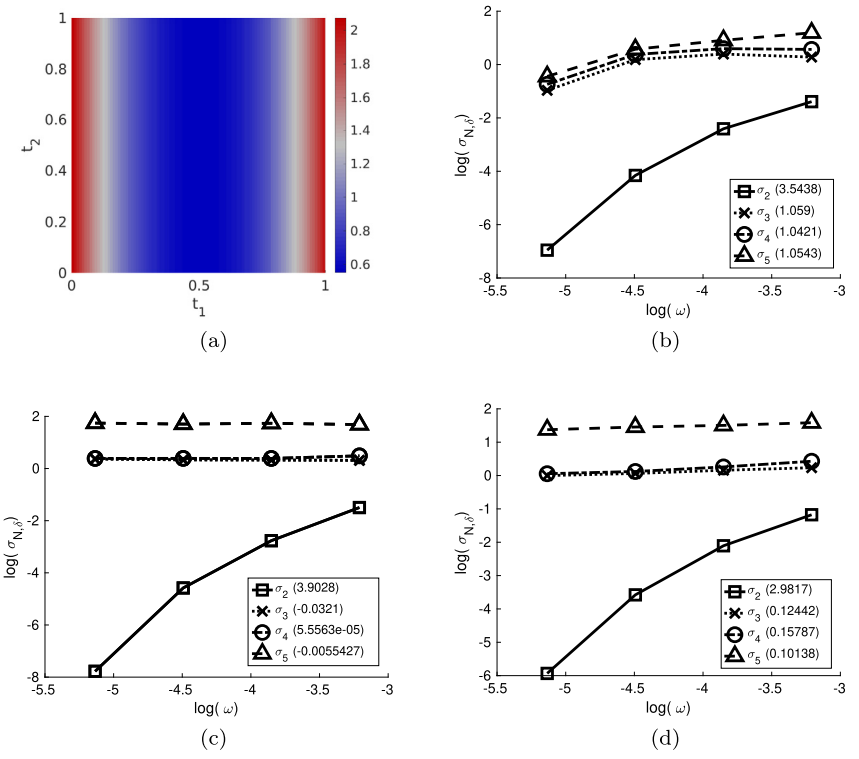


Fig. 1.2. (a) A plot of the mixture density (1.4) for $\omega = 0.25$. (b) The first four non-trivial eigenvalues of the discrete graph Laplacian L_N with parameters $(p, q, r) = (1/2, 2, 1/2)$ as a function of the mean parameter ω. Values reported in brackets in the legends indicate numerical slope of the lines fitted to the data. (c) Showing the first four non-trivial eigenvalues of L_N with $(p, q, r) = (1, 2, 1)$. (d) Showing the same results for parameters $(p, q, r) = (1, 3/2, 1)$.

a balanced case with $(1, 2, 1)$ and two unbalanced cases with $(1/2, 2, 1/2)$ and $(1, 3/2, 1)$. While our theory does not make a prediction regarding the rate at which the second eigenvalue vanishes with ω, we can still use our theoretical insights to postulate uniform or ratio gaps between the second and third eigenvalues.

In the balanced case where $q = p + r$ we observe that the second eigenvalue vanishes with ω while the rest of the spectrum remains bounded away from zero; in contrast, in the unbalanced case $q > p + r$ the third eigenvalue also vanishes and only a spectral ratio gap manifests. The results in the unbalanced case $q < p + r$ are less clear since the higher eigenvalues still vanish, but they do so rather slowly; this may be attributed to numerical error. The results are in agreement with our analysis and numerical results in the continuum limit and suggest that the characteristic behavior we prove for our specific construction of the sampling density ϱ is in fact a more general phenomenon that applies for other type of clustered data and on finite data sets. Further details regarding this experiment are summarized in Subsection 5.4. ◇

1.5. Outline

The remainder of the paper is organized as follows. Section 2 sets up the necessary framework and notation. Section 3 contains the precise statements of the key results Theorems 3.2 and 3.4, relating to Main Result 1.1; proofs of these results are postponed to Section 6. Numerical results illustrating, and extending the Main Result 1.1 and leading to the Conjecture 1.2 are presented in Section 4. Section 5 contains the informal derivation of (1.1) from the parameterized family of graph Laplacians (1.2), and presents the formal calculations and numerical experiments that were summarized under Contribution 3. Our conclusions are given in Section 7. Appendices A, B, C, and D contain, respectively: connections between the diffusion maps and $\mathcal{L}$; discussion of function spaces; the min-max principle; and a weighted Cheeger inequality.

2. The set-up

In this section we set-up the functional analytic framework for our theory and numerics. Subsection 2.1 describes the notation and introduces weighted Laplacian operators in this framework, and Subsection 2.2 is devoted to our precise formulation of binary clustered data in the perfect or nearly separated clustered data setting.

2.1. Preliminaries

For an open subset $\Omega \subseteq \mathcal{Z} \subset \mathbb{R}^d$ with $C^{1,1}$ boundary, consider a probability density function ϱ satisfying

$$\varrho \in C^\infty(\bar{\Omega}), \qquad \int_\Omega \varrho(x)dx = 1, \qquad \varrho^- < \varrho(x) < \varrho^+, \qquad \forall x \in \bar{\Omega}, \tag{2.1}$$

with constants $\varrho^-, \varrho^+ > 0$. We also denote the measure of subsets Ω' of Ω with respect to ϱ with the following notation

$$|\Omega'|_\varrho := \int_{\Omega'} \varrho(x)dx. \tag{2.2}$$

Given a continuous probability density function ϱ as above with full support on $\Omega \subseteq \mathcal{Z}$ we define the weighted space

$$L^2(\Omega, \varrho^s) := \left\{ u : \int_\Omega |u(x)|^2 \varrho(x)^s dx < +\infty \right\}, \tag{2.3}$$

with inner product

$$\langle u, v \rangle_{\varrho^s} := \int_\Omega u(x)v(x)\varrho^s(x)dx, \tag{2.4}$$

for any $s \in \mathbb{R}$. This reduces to the standard $L^2(\Omega)$ space with norm $\|\cdot\|_{L^2(\Omega)}$ and inner product $\langle \cdot, \cdot \rangle$ if $\varrho = 1$ on Ω. Furthermore, for $\varrho > 0$ a.e. on Ω and parameters $(p, q, r) \in \mathbb{R}^3$, we define the weighted Sobolev spaces

$$H^1(\Omega, \varrho) := \left\{ \frac{u}{\varrho^r} \in L^2(\Omega, \varrho^{p+r}) : \|u\|_{H^1(\Omega,\varrho)} := \langle u, u \rangle_V < +\infty \right\},$$

where the $\langle \cdot, \cdot \rangle_V$ inner product is defined as

$$\langle u, v \rangle_V := \left\langle \nabla \left(\frac{u}{\varrho^r} \right), \nabla \left(\frac{v}{\varrho^r} \right) \right\rangle_{\varrho^q} + \left\langle \frac{u}{\varrho^r}, \frac{v}{\varrho^r} \right\rangle_{\varrho^{p+r}}, \tag{2.5}$$

which is the natural inner product induced by the bilinear form $\left\langle (\mathcal{L} + \frac{1}{\varrho^r})u, \frac{v}{\varrho^r} \right\rangle_{\varrho^{p+r}}$. We then introduce the following subspaces of $L^2(\Omega, \varrho^{p+r})$ and $H^1(\Omega, \varrho)$:

$$V^0(\Omega, \varrho) := \left\{ \frac{u}{\varrho^r} \in L^2(\Omega, \varrho^{p+r}) : \left\langle \frac{u}{\varrho^r}, 1 \right\rangle_{\varrho^{p+r}} = \langle u, \varrho^p \rangle = 0 \right\},$$

$$V^1(\Omega, \varrho) := \left\{ u \in H^1(\Omega, \varrho) : \langle u, \varrho^r \rangle_V = 0 \right\} \subset V^0(\Omega, \varrho).$$

We use $H^1(\Omega)$ and $V^1(\Omega)$ to denote the standard H^1 space, and its subspace excluding constants, given by $H^1(\Omega, \mathbf{1}_\Omega)$ and $V^1(\Omega, \mathbf{1}_\Omega)$. The former coincides with the usual Sobolev spaces while the latter coincides with the subspace of $H^1(\Omega)$ consisting of mean zero functions.

In this work, we focus on the class of weighted Laplacian operators defined by equation (1.1), for an appropriate density ϱ and parameters $(p, q, r) \in \mathbb{R}^3$. We generally suppress the dependence of $\mathcal{L}$ on ϱ and the constants p, q, r for convenient notation and make the choice of these parameters explicit in our statements. As we show next, the operator $\mathcal{L}$ is positive semi-definite and since the first eigenpair $(\sigma_1, \varphi_1) = (0, \varrho^r \mathbf{1}_\Omega)$ is known it is convenient to work orthogonal to φ_1 so as to make the operator strictly positive; in other words, we consider the operator $\mathcal{L}$ on the space $V^1(\Omega, \varrho)$.

Lemma 2.1. *If ϱ satisfies (2.1), then the bilinear form*

$$\langle \mathcal{L}u, v \rangle_{\varrho^{p-r}} = \left\langle \varrho^q \nabla \left(\frac{u}{\varrho^r} \right), \nabla \left(\frac{v}{\varrho^r} \right) \right\rangle, \tag{2.6}$$

is symmetric and positive definite on $V^1(\Omega, \varrho) \times V^1(\Omega, \varrho)$. In particular, the operator

$$\mathcal{L} : V^1(\Omega, \varrho) \mapsto V^0(\Omega, \varrho),$$

defined in the weak sense, is self-adjoint and strictly positive definite and the inverse operator

$$\mathcal{L}^{-1} : V^0(\Omega, \varrho) \mapsto V^0(\Omega, \varrho),$$

exists and is compact.

Proof. The fact that $\mathcal{L}$ is self-adjoint and strictly positive on $V^1(\Omega, \varrho)$ can be verified directly. The fact that $\mathcal{L}^{-1}$ is well-defined follows from the Lax-Milgram Lemma [32, Lem. 2.32]. Compactness follows from Proposition B.3. □

Following the spectral theorem [19, Thms. D.6, D.7] we then have:

Proposition 2.2. *Let $(p, q, r) \in \mathbb{R}^3$, and suppose ϱ satisfies (2.1). Then $\mathcal{L} : V^1(\Omega, \varrho) \mapsto V^0(\Omega, \varrho)$ has a discrete spectrum with eigenvalues $0 \leq \sigma_2 \leq \sigma_3 \leq \dots$ and eigenfunctions $\{\varphi_j\}_{j \geq 2} \in V^1(\Omega, \varrho)$ that form an orthogonal basis in both $V^1(\Omega, \varrho)$ and $V^0(\Omega, \varrho)$. Furthermore, we may extend $\mathcal{L}$ to the operator $\mathcal{L} : H^1(\Omega, \varrho) \mapsto L^2(\Omega, \varrho^{p-r})$ and include the eigenpair $(\sigma_1, \varphi_1) = (0, |\Omega|^{1/2}_{\varrho^{p+r}} \varrho^r \mathbf{1}_\Omega)$.*

Remark 2.3. Writing $u = \varrho^r u'$ and $v = \varrho^r v'$ we note that the identity (2.6) may be written as

$$\left\langle \varrho^{p-q} \mathcal{L}(\varrho^r u'), v' \right\rangle_{\varrho^q} = \langle \nabla u', \nabla v' \rangle_{\varrho^q}.$$

From this we see [35] that the operator

$$\mathcal{G} := -\varrho^{p-q} \circ \mathcal{L} \circ \varrho^r$$

is the generator of the reversible diffusion process

$$dX_t = -\nabla \Psi(X_t) dt + \sqrt{2} dB,$$

where $\Psi = -\log(\varrho^q)$, and B denotes a d dimensional Brownian motion. This diffusion process has invariant measure proportional to $\exp(-\Psi) = \varrho^q$. This observation thus establishes a connection between the operator

$\mathcal{L}$ and diffusion processes which, when $q > 0$, concentrate in regions where ϱ is large and sampling density of the data is high. For a more detailed discussion on the connections between diffusion maps and the operators weighted elliptic operators $\mathcal{L}$, see Appendix A.

2.2. Perturbations of densities

We now consider a specific setting of a density ϱ_0 that is supported on a strict subset $\mathcal{Z}' \subset \mathcal{Z}$, consisting of two disjoint sets $\mathcal{Z}^+$ and $\mathcal{Z}^-$. We then consider a sequence of probability densities ϱ_ϵ supported on the whole set $\mathcal{Z}$ that approximate ϱ_0. In the next two subsections we outline our assumptions regarding $\mathcal{Z}'$, ϱ_0 and ϱ_ϵ and introduce weighted Laplacian operators using these densities.

2.2.1. Assumptions on the clusters and densities

We begin by introducing a set of assumptions on the domains $\mathcal{Z}, \mathcal{Z}'$, the density ϱ_0, and the approximating sequence of densities ϱ_ϵ.

Assumption 2.4. The sets $\mathcal{Z}, \mathcal{Z}' = \mathcal{Z}^+ \cup \mathcal{Z}^- \subset \mathbb{R}^d$ satisfy the following:

(a) $\mathcal{Z}$ is open, bounded and connected.
(b) $\mathcal{Z}'$ is a subset of $\mathcal{Z}$ consisting of two open connected subsets $\mathcal{Z}^+$ and $\mathcal{Z}^-$.
(c) $\mathcal{Z}^\pm$ are disjoint from one another and from $\partial\mathcal{Z}$, the boundary of $\mathcal{Z}$: $\exists l, l' > 0$ so that

$$\operatorname{dist}(\mathcal{Z}^+, \mathcal{Z}^-) > l > 0, \qquad \text{and} \qquad \operatorname{dist}(\mathcal{Z}^\pm, \partial\mathcal{Z}) > l' > 0.$$

(d) $\partial\mathcal{Z}$ and $\partial\mathcal{Z}'$ are at least $C^{1,1}$.

The assumption that $\mathcal{Z}^\pm$ are well separated from $\partial\mathcal{Z}$ in Assumption 2.4(c) is not crucial but allows for more convenient presentation of our results. We think of $\mathcal{Z}^\pm$ as "clusters" in the continuum limit.

Assumption 2.5. The density ϱ_0 satisfies the following:

(a) (*Supported on clusters*) $\varrho_0 = 0$ on $\mathcal{Z} \setminus \bar{\mathcal{Z}}'$.
(b) (*Probability density function*) $\int_{\mathcal{Z}'} \varrho_0(x)dx = 1$.
(c) (*Uniformly bounded within clusters*) $\exists \varrho^\pm > 0$ so that $\varrho^- \le \varrho_0(x) \le \varrho^+$, for all $x \in \bar{\mathcal{Z}}'$.
(d) (*Smoothness*) $\varrho_0 \in C^\infty(\bar{\mathcal{Z}}')$.
(e) (*Equal sized clusters*) Given $p, r \ge 0$, the density ϱ_0^{p+r} assigns equal mass to $\mathcal{Z}^+$ and $\mathcal{Z}^-$, i.e.,

$$\int_{\mathcal{Z}^+} \varrho_0^{p+r}(x)dx = \int_{\mathcal{Z}^-} \varrho_0^{p+r}(x)dx\,.$$

We highlight that Assumption 2.5(b) and (e) are not crucial to our analysis. Condition (b) is natural when considering limits of graph Laplacian operators defined from data distributed according to a measure with density ϱ_0, but all of our analysis can be generalized to integrable ϱ_0 simply by observing that the eigenfunctions of $\mathcal{L}$ are invariant under scaling of ϱ_0 by a constant λ, whilst the eigenvalues scale by λ^{q-p-r}. Condition (e) allows for a more convenient presentation with less cumbersome notation but can be removed at the price of a lengthier exposition; see Remark 6.2 below.

Given a density ϱ_0 satisfying Assumption 2.5, we consider a sequence of densities ϱ_ϵ with full support on $\bar{\mathcal{Z}}$ that converge to ϱ_0 as $\epsilon \to 0$ in a suitable sense. We have in mind densities ϱ_ϵ that become more and more concentrated in $\mathcal{Z}'$ as $\epsilon \to 0$. In what follows, we define

$$\Omega_\delta := \{x : \mathrm{dist}(x, \Omega) \leq \delta\}, \tag{2.7}$$

for any set $\Omega \subseteq \bar{\mathcal{Z}}$ and denote the Minkowski (exterior) boundary measure of Ω as

$$|\partial\Omega| := \liminf_{\delta \downarrow 0} \frac{1}{\delta} \left[|\Omega_\delta| - |\Omega| \right].$$

It follows that when ϵ is sufficiently small, $\exists \theta > 0$ so that

$$|\Omega_\epsilon \setminus \Omega| \leq \theta\epsilon |\partial\Omega|. \tag{2.8}$$

Assumption 2.6. Let $0 < L := \min \mathrm{dist}(\mathcal{Z}^{\pm}, \partial\mathcal{Z})$. Then there is $\epsilon_0 \in (0, L/4)$ and constants $K_1, K_2 > 0$ such that, for all $\epsilon \in (0, \epsilon_0)$, the densities ϱ_ϵ satisfy:

(a) (*Full support*) $\mathrm{supp}\varrho_\epsilon = \bar{\mathcal{Z}}$.
(b) (*Probability density function*) $\int_{\mathcal{Z}} \varrho_\epsilon(x)dx = 1$.
(c) (*Approximation within clusters*) $\exists K_1 > 0$ so that $\|\varrho_\epsilon - \varrho_0\|_{C^\infty(\bar{\mathcal{Z}}')} \leq K_1\epsilon$ as $\epsilon \downarrow 0$.
(d) (*Vanishing outside clusters*) $\exists K_2 > 0$ so that $\varrho_\epsilon(x) = K_2\epsilon$ for $x \in \mathcal{Z} \setminus \mathcal{Z}'_\epsilon$.
(e) (*Controlled derivatives*) $\exists K_3 > 0$ so that

$$|\nabla \varrho_\epsilon(x)| \leq K_3 \epsilon^{-1}, \qquad \forall x \in \mathcal{Z}'_\epsilon \setminus \mathcal{Z}'.$$

Once again Assumption 2.6(b) is not crucial to our analysis but is needed to make sure the operator $\mathcal{L}_\epsilon$ defined in (2.13) is the continuum limit of a graph Laplacian. As a consequence of Assumptions 2.5(c) and 2.6(c)-(e), it follows that ϱ_ϵ is uniformly bounded above and below inside $\mathcal{Z}'$: there exist constants $\varrho^{\pm}_{\epsilon_0} > 0$ so that

$$\varrho^-_{\epsilon_0} \leq \varrho_\epsilon(x) \leq \varrho^+_{\epsilon_0}, \qquad \forall x \in \bar{\mathcal{Z}}' \text{ and } \forall \epsilon \in (0, \epsilon_0). \tag{2.9}$$

Note that the upper bound holds on all of $\mathcal{Z}$ as well, whereas the lower bound clearly does not in view of Assumption 2.6(d).

Remark 2.7. The above set of assumptions on ϱ_ϵ, needed for the theoretical results in Subsection 3.2, are rather specific. However, as demonstrated in Example 1.4 (see also the discussion leading to that example) the results appear to generalize to settings where Assumption 2.6 is violated. This suggests that the limitations of Assumption 2.6 reflect our method of proof; relaxing these conditions is an interesting direction for future research.

Example 2.8. Consider the standard mollifier

$$g(x) := \begin{cases} C^{-1} \exp\left(-\dfrac{1}{1-|x|^2}\right) & |x| \leq 1, \\ 0 & |x| > 1. \end{cases}, \qquad g_\epsilon(x) := \frac{1}{\epsilon^d} g\left(\frac{x}{\epsilon}\right), \tag{2.10}$$

where $C = \int_{|x|\leq 1} \exp\left(-\frac{1}{1-|x|^2}\right) dx$ is a normalizing constant. Now, given $\epsilon > 0$ and the density ϱ_0 (extended by zero to all of $\mathcal{Z}$) define

$$\varrho_\epsilon(x) := \frac{1}{K_\epsilon}\Big(\epsilon + g_\epsilon * \varrho_0(x)\Big), \qquad K_\epsilon := \int_{\mathcal{Z}} \Big(\epsilon + g_\epsilon * \varrho_0(x)\Big) dx. \tag{2.11}$$

One can directly verify that the above construction of ϱ_ϵ satisfies Assumption 2.6. ◇

2.2.2. Assumptions on the weighted Laplacian operators

With the densities ϱ_0 and ϱ_ϵ identified we then consider the operators $\mathcal{L}_0$ and $\mathcal{L}_\epsilon$ in the same form as (1.1) as follows:

$$\begin{cases} \mathcal{L}_0 u := -\dfrac{1}{\varrho_0^p}\mathrm{div}\left(\varrho_0^q \nabla\left(\dfrac{u}{\varrho_0^r}\right)\right), & \text{in } \mathcal{Z}' \\ \varrho_0^q \dfrac{\partial}{\partial n}\left(\dfrac{u}{\varrho_0^r}\right) = 0, & \text{on } \partial\mathcal{Z}'. \end{cases} \tag{2.12}$$

Similarly for ϱ_ϵ,

$$\begin{cases} \mathcal{L}_\epsilon u := -\dfrac{1}{\varrho_\epsilon^p}\mathrm{div}\left(\varrho_\epsilon^q \nabla\left(\dfrac{u}{\varrho_\epsilon^r}\right)\right), & \text{in } \mathcal{Z} \\ \varrho_\epsilon^q \dfrac{\partial}{\partial n}\left(\dfrac{u}{\varrho_\epsilon^r}\right) = 0, & \text{on } \partial\mathcal{Z}. \end{cases} \tag{2.13}$$

By Lemma 2.1 and Proposition 2.2, the operators

$$\mathcal{L}_0 : H^1(\mathcal{Z}', \varrho_0) \mapsto L^2(\mathcal{Z}', \varrho_0^{p-r}) \qquad \text{and} \qquad \mathcal{L}_\epsilon : H^1(\mathcal{Z}, \varrho_\epsilon) \mapsto L^2(\mathcal{Z}, \varrho_\epsilon^{p-r})$$

are self-adjoint and positive semi-definite. Furthermore, $\mathcal{L}_0$ has a zero eigenvalue of multiplicity two while $\mathcal{L}_\epsilon$ has a zero eigenvalue of multiplicity one. Both operators have positive, real, discrete eigenvalues after the zero eigenvalues. For $j = 1, 2, 3, ...$ let $\sigma_{j,0}$ and $\sigma_{j,\epsilon}$ denote the eigenvalues of $\mathcal{L}_0$ and $\mathcal{L}_\epsilon$ respectively (in increasing order and accounting for repetitions) and let $\varphi_{j,0}$ and $\varphi_{j,\epsilon}$ denote the corresponding eigenfunctions. Recall that $\varphi_{1,0} = |\mathcal{Z}'|^{-1/2}_{\varrho_0^{p+r}} \varrho_0^r \mathbf{1}_{\mathcal{Z}'}$ and $\varphi_{1,\epsilon} = |\mathcal{Z}|^{-1/2}_{\varrho_\epsilon^{p+r}} \varrho_\epsilon^r \mathbf{1}_{\mathcal{Z}}$, both with corresponding zero eigenvalues. Since we are interested in the eigenpairs for $j \geq 2$ it is more convenient to work orthogonal to the first eigenfunctions from now on, that is, to consider the spaces $V^1(\mathcal{Z}', \varrho_0)$ and $V^1(\mathcal{Z}, \varrho_\epsilon)$ respectively. Thus, we consider the pairs $\{\sigma_{j,0}, \varphi_{j,0}\}$ and $\{\sigma_{j,\epsilon}, \varphi_{j,\epsilon}\}$ for $j \geq 2$ that solve the eigenvalue problems

$$\left\langle \varrho_0^q \nabla\left(\frac{\varphi_{j,0}}{\varrho_0^r}\right), \nabla\left(\frac{v}{\varrho_0^r}\right)\right\rangle = \sigma_{j,0}\left\langle \varrho_0^{p-r}\varphi_{j,0}, v\right\rangle, \qquad \varphi_{j,0}, v \in V^1(\mathcal{Z}', \varrho_0), \tag{2.14}$$

and

$$\left\langle \varrho_\epsilon^q \nabla\left(\frac{\varphi_{j,\epsilon}}{\varrho_\epsilon^r}\right), \nabla\left(\frac{v}{\varrho_\epsilon^r}\right)\right\rangle = \sigma_{j,\epsilon}\left\langle \varrho_\epsilon^{p-r}\varphi_{j,\epsilon}, v\right\rangle, \qquad \varphi_{j,\epsilon}, v \in V^1(\mathcal{Z}, \varrho_\epsilon). \tag{2.15}$$

Throughout the article we take $\varphi_{j,0}$ and $\varphi_{j,\epsilon}$ to be normalized in $L^2(\mathcal{Z}', \varrho_0^{p-r})$ and $L^2(\mathcal{Z}, \varrho_\epsilon^{p-r})$ respectively.

We collect some definitions and notation concerning the spectral gaps of the operators $\mathcal{L}_0$ and $\mathcal{L}_\epsilon$ and Poincaré constants on certain subsets of $\mathcal{Z}$ and $\mathcal{Z}'$; these are used throughout the article.

Definition 2.9 *(Standard spectral gap Λ_Δ).* We say that the standard spectral gap condition holds for a subset Ω of $\mathcal{Z}$ if the Poincaré inequality is satisfied on Ω with an optimal constant $\Lambda_\Delta(\Omega) > 0$, i.e.,

$$\int_\Omega |\nabla u|^2 \, dx \geq \Lambda_\Delta(\Omega) \int_\Omega |u|^2 dx, \qquad \forall u \in V^1(\Omega). \tag{2.16}$$

We also define a certain ϱ_0 weighted version of the above spectral gap definition.

Definition 2.10 *($\mathcal{L}_0$ spectral gap Λ_0).* We say that the $\mathcal{L}_0$ spectral gap condition holds for a subset Ω of $\mathcal{Z}'$ if the following weighted Poincaré inequality is satisfied with an optimal constant $\Lambda_0(\Omega) > 0$

$$\int_\Omega \varrho_0^q \left| \nabla \left(\frac{u}{\varrho_0^r} \right) \right|^2 dx \geq \Lambda_0(\Omega) \int_\Omega \left| \frac{u}{\varrho_0^r} \right|^2 \varrho_0^{p+r} dx, \qquad \forall u \in V^1(\Omega, \varrho_0). \tag{2.17}$$

Observe that condition (2.17) is equivalent to the assumption that the second eigenvalue of the operator $\mathcal{L}_0$ restricted to the set Ω is bounded away from zero. Finally, we define the notion of a uniform spectral gap for $\mathcal{L}_\epsilon$.

Definition 2.11 *($\mathcal{L}_\epsilon$ uniform spectral gap Λ_ϵ).* Given $\epsilon_0 > 0$ we say that the $\mathcal{L}_\epsilon$ uniform spectral gap condition holds for a subset Ω of $\mathcal{Z}$ if $\forall \epsilon \in (0, \epsilon_0)$ there exists an optimal constant $\Lambda_\epsilon(\Omega) > 0$ so that

$$\int_\Omega \left| \nabla \left(\frac{u}{\varrho_\epsilon^r} \right) \right|^2 \varrho_\epsilon^q dx \geq \Lambda_\epsilon(\Omega) \int_\Omega \left| \frac{u}{\varrho_\epsilon^r} \right|^2 \varrho_\epsilon^{p+r} dx, \qquad \forall u \in V^1(\Omega, \varrho_\epsilon). \tag{2.18}$$

Remark 2.12. To connect the spectral gaps of $\mathcal{L}_\epsilon$ restricted to the clusters $\mathcal{Z}^\pm$ with the spectral gaps of the limiting operator $\mathcal{L}_0$ on these clusters, one can make use of the knowledge that ϱ_ϵ converges to ϱ_0 on $\mathcal{Z}^\pm$ by Assumption 2.6(c). More precisely, let us suppose (2.17) holds. We show in Theorem 3.1 that $\sigma_{1,0} = \sigma_{2,0} = 0$ and $\sigma_{3,0} > 0$. Since $\varrho_\epsilon(x)$ converges to $\varrho_0(x)$ pointwise for every $x \in \mathcal{Z}'$, this spectral gap translates to $\mathcal{L}_\epsilon$ for small enough ϵ within the set $\mathcal{Z}'$, and so we can assert (2.18) for $\Omega = \mathcal{Z}^\pm$. The assumption that the restriction of $\mathcal{L}_\epsilon$ to $\mathcal{Z}^\pm$ has a spectral gap is related to the indivisibility parameter in the context of well-separated mixture models of [21].

Remark 2.13. Note that for subsets Ω where ϱ_ϵ is constant, say $\varrho_\epsilon(x) = c_\epsilon$, condition (2.18) reduces to a spectral gap of the standard Laplacian restricted to Ω, with the constant Λ_Δ in (2.16) replaced by $\Lambda_\Delta c_\epsilon^{p+r-q}$. This becomes important when investigating the behavior of $\mathcal{L}_\epsilon$ away from the clusters $\mathcal{Z}^\pm$ and is precisely the reason why we obtain a condition on the sign of $q - p - r$ in our main theorems, see for example Theorem 3.2.

3. Spectral analysis: statement of theorems

In this section we describe the spectral properties of the operators $\mathcal{L}_0$ and $\mathcal{L}_\epsilon$ in relation to certain geometric features in the data summarized in the densities ϱ_0 and ϱ_ϵ. We present precise statements of our key theoretical results, postponing the proofs to Section 6. We define, and then identify, gaps between the second and third eigenvalues of $\mathcal{L}_\epsilon$ together with concentration properties of the second eigenfunction $\varphi_{2,\epsilon}$ as $\epsilon \downarrow 0$. More precisely, we show that the nature and existence of a spectral gap is dependent upon the choice of p, q and r and, under general conditions, concentration properties of $\varphi_{2,\epsilon}$ are directly related to concentration properties of ϱ_ϵ. In Subsection 3.1 we consider the perfectly clustered case pertaining the operator $\mathcal{L}_0$ while Subsection 3.2 perturbs this setting and considers the nearly clustered case corresponding to the operator $\mathcal{L}_\epsilon$.

3.1. Perfectly separated clusters

Recall the concept of perfectly separated clusters from the introduction, the density ϱ_0 and the resulting operator $\mathcal{L}_0$ defined on $\mathcal{Z}'$. The corresponding low-lying spectrum of $\mathcal{L}_0$ can be characterized explicitly:

Theorem 3.1 *(Low-lying spectrum of $\mathcal{L}_0$ and Fiedler vector). Suppose $(p, q, r) \in \mathbb{R}^3$ and Assumptions 2.4 and 2.5 hold. Then $\mathcal{L}_0$ is positive semi-definite and self-adjoint on the weighted Sobolev space $H^1(\mathcal{Z}', \varrho_0)$. Denote its eigenvalues by $\sigma_{1,0} \leq \sigma_{2,0} \leq \cdots$ with corresponding eigenfunctions $\varphi_{j,0}$, $j \geq 1$. Then it holds that:*

(i) The first eigenpair is given by

$$\sigma_{1,0} = 0\,, \qquad \varphi_{1,0} = \frac{1}{|\mathcal{Z}'|^{1/2}_{\varrho_0^{p+r}}} \varrho_0^r(x) \mathbf{1}_{\mathcal{Z}'}(x), \qquad \forall x \in \mathcal{Z}'\,.$$

(ii) The second eigenpair is given by

$$\sigma_{2,0} = 0\,, \qquad \varphi_{2,0} = \frac{1}{|\mathcal{Z}'|^{1/2}_{\varrho_0^{p+r}}} \varrho_0^r(x) \left(\mathbf{1}_{\mathcal{Z}^+}(x) - \mathbf{1}_{\mathcal{Z}^-}(x)\right), \qquad \forall x \in \mathcal{Z}'\,.$$

(iii) $\mathcal{L}_0$ has a uniform spectral gap, i.e., $\sigma_{3,0} > 0$.

Part (i,ii) of Theorem 3.1 can be verified directly by substituting $\varphi_{1,0}$ and $\varphi_{2,0}$ into (2.14). Then it remains to show (iii), the lower bound on the third eigenvalue $\sigma_{3,0}$ which follows from Proposition 6.1, stating that $\mathcal{L}_0$ has a spectral gap on $\mathcal{Z}'$ so long as its restriction to each of the clusters $\mathcal{Z}^\pm$ has a spectral gap. Since ϱ_0 is bounded away from zero on the clusters this condition holds since $\mathcal{Z}^\pm$ are assumed to be connected sets of positive Lebesgue measure.

3.2. Nearly separated clusters

We now turn our attention to the densities ϱ_ϵ that have full support on $\bar{\mathcal{Z}}$, but concentrate around $\mathcal{Z}'$ as ϵ decreases. This represents the practical setting where we do not have perfect clusters $\mathcal{Z}^\pm$ and so the density ϱ_0 is perturbed. A central question here is whether the second eigenpair $\{\sigma_{2,\epsilon}, \varphi_{2,\epsilon}\}$ of $\mathcal{L}_\epsilon$ exhibits behavior similar to the second eigenpair $\{\sigma_{2,0}, \varphi_{2,0}\}$ of $\mathcal{L}_0$ as $\varrho_\epsilon \to \varrho_0$; that is, in the limit as we approach the ideal case of perfect clusters $\mathcal{Z}^\pm$.

In order to establish such a result we first need to approximate the first three eigenvalues of $\mathcal{L}_\epsilon$:

Theorem 3.2 *(Low-lying eigenvalues of $\mathcal{L}_\epsilon$). Let $(p, q, r) \in \mathbb{R}^3$ satisfy $p + r > 0$ and $q > 0$, and suppose Assumptions 2.4, 2.5, and 2.6 hold and that $\Lambda_\Delta(\mathcal{Z} \setminus \mathcal{Z}'_{\epsilon_0}) > 0$ for a sufficiently small $\epsilon_0 > 0$. Then the following holds for all $(\epsilon, \beta) \in (0, \epsilon_0) \times (0, 1)$:*

(i) The first eigenpair is given by

$$\sigma_{1,\epsilon} = 0\,, \qquad \varphi_{1,\epsilon} = \frac{1}{|\mathcal{Z}|^{1/2}_{\varrho_\epsilon^{p+r}}} \varrho_\epsilon^r(x) \mathbf{1}_{\mathcal{Z}}(x) \qquad \forall x \in \mathcal{Z}\,.$$

(ii) The second eigenvalue $\sigma_{2,\epsilon}$ tends to zero as $\epsilon \to 0$,

$$0 \leq \sigma_{2,\epsilon} \leq \Xi_1 \epsilon^{q-\beta},$$

with $\Xi_1 > 0$ a uniform constant independent of ϵ.

(iii) The third eigenvalue behaves differently depending on the (p, q, r) parameters:

- *if $q > p + r$, then $\exists\, \Xi_2, \Xi_3 > 0$ independent of ϵ such that,*

$$\Xi_2 \epsilon^{2(q-p-r)} \leq \sigma_{3,\epsilon} \leq \Xi_3 \epsilon^{q-p-r-2\beta}\,,$$

  and so $\mathcal{L}_\epsilon$ does not have a uniform spectral gap on $\mathcal{Z}$;
- *if $q = p + r$ then there exist constants $\Xi_4, \Xi_5 > 0$, independent of ϵ, so that*

$$\Xi_4 \leq \sigma_{3,\epsilon} \leq \Xi_5,$$

  and so $\mathcal{L}_\epsilon$ has a uniform spectral gap on $\mathcal{Z}$;
- *if $q < p + r$, then there exist constants $\Xi_6, \Xi_7 > 0$, independent of ϵ, so that*

$$\Xi_6 \epsilon^{p+r-q} \leq \sigma_{3,\epsilon} \leq \Xi_7\,.$$

Once again part (i) can be verified directly by substituting $\varphi_{1,\epsilon}$ in (2.15). Part (ii) is a consequence of Proposition 6.4 that obtains an upper bound on $\sigma_{2,\epsilon}$ using a perturbation argument. More precisely, we first construct an explicit approximation $\varphi_{F,\epsilon}$ of $\varphi_{2,\epsilon}$ as a smoothed out version of $\varphi_{2,0}$, normalized in $V^1(\mathcal{Z}, \varrho_\epsilon)$ and supported on a set slightly larger than $\mathcal{Z}'$. We choose a parameter $\beta > 0$ such that $|\nabla \varphi_{F,\epsilon}|$ is controlled by $\epsilon^{-\beta}$ at the boundary of $\mathcal{Z}'$. This is precisely the parameter β appearing in Theorem 3.2. By construction, we then have that $\varphi_{F,\epsilon}$ converges to the normalization of $\varphi_{2,0}$ as $\epsilon \to 0$. Using this approximate eigenfunction as well as $\varphi_{1,\epsilon}$ from part (i) in the min-max principle (see Proposition C.1) yields the desired upper bound on $\sigma_{2,\epsilon}$.

Part (iii) requires more elaborate arguments as outlined in Subsection 6.2.2. The lower bounds on $\sigma_{3,\epsilon}$ follow from Proposition 6.7 that is in turn based on a generalization of Cheeger's inequality (see Proposition D.1). The upper bounds follow from Proposition 6.8 the proof of which uses similar ideas as for the upper bound of $\sigma_{2,\epsilon}$, applying the min-max principle but with a different candidate eigenfunction.

Several interesting conclusions can be drawn from our arguments in Subsection 6.2 aimed at proving Theorem 3.2. The existence of spectral gaps for $\mathcal{L}_\epsilon$ inside the clusters and away from the clusters separately allows us to formally deduce bounds on the low-lying spectrum. Consider the set

$$\mathcal{Z}'_\epsilon := \{x : \operatorname{dist}(x, \mathcal{Z}') \leq \epsilon\},$$

and suppose that for some fixed $\epsilon_0 > 0$, we have $\Lambda_\Delta(\mathcal{Z} \setminus \mathcal{Z}'_{\epsilon_0}) > 0$, that is, the standard Laplacian has a spectral gap away from the clusters according to Definition 2.9. Since $\varrho_\epsilon(x) = K_2\epsilon$ for $x \in \mathcal{Z} \setminus \mathcal{Z}'_{\epsilon_0}$, we have for all $u \perp \mathbf{1}_{\mathcal{Z}\setminus\mathcal{Z}'_{\epsilon_0}}$ in $V^1(\mathcal{Z} \setminus \mathcal{Z}'_{\epsilon_0})$

$$(K_2\epsilon)^{2r-q} \int\limits_{\mathcal{Z}\setminus\mathcal{Z}'_{\epsilon_0}} \left|\nabla\left(\frac{u}{\varrho_\epsilon^r}\right)\right|^2 \varrho_\epsilon^q dx \geq \Lambda_\Delta(\mathcal{Z} \setminus \mathcal{Z}'_{\epsilon_0})(K_2\epsilon)^{r-p} \int\limits_{\mathcal{Z}\setminus\mathcal{Z}'_{\epsilon_0}} \left|\frac{u}{\varrho_\epsilon^r}\right|^2 \varrho_\epsilon^{p+r} dx\,.$$

This simple calculation shows that $\Lambda_\epsilon(\mathcal{Z} \setminus \mathcal{Z}'_{\epsilon_0}) = \mathcal{O}(\epsilon^{q-p-r})$, and so the existence of a uniform $\mathcal{L}_\epsilon$ spectral gap away from the clusters is dependent on the relation between q and $p + r$, in fact we need $q \leq p + r$ to ensure $\Lambda_\epsilon(\mathcal{Z} \setminus \mathcal{Z}'_{\epsilon_0}) > 0$ independent of ϵ which is in line with the conditions in Theorem 3.2(iii).

Combining parts (ii, iii) of Theorem 3.2 yields the following corollary concerning the existence of uniform or ratio gaps in the spectrum of $\mathcal{L}_\epsilon$ depending on (p, q, r). This corollary is a detailed statement of Main Result 1.1(iii).

Corollary 3.3 *(Spectral ratio gap when $q \neq p + r$).*

Suppose that the conditions of Theorem 3.2 are satisfied and that $q \neq p + r$. Then the following holds for all $(\epsilon, \beta) \in (0, \epsilon_0) \times (0, 1)$:

(i) if $q > p + r$ then there exists a constant $\Xi_1 > 0$ independent of ϵ, so that

$$\frac{\sigma_{2,\epsilon}}{\sigma_{3,\epsilon}} \leq \Xi_1 \epsilon^{2(p+r)-q-\beta};$$

(ii) if $q < p + r$ then there exists a constant $\Xi_2 > 0$ independent of ϵ,

$$\frac{\sigma_{2,\epsilon}}{\sigma_{3,\epsilon}} \leq \Xi_2 \epsilon^{2q-p-r-\beta}.$$

We note that while this corollary suggests that there may be no spectral ratio gap when $q > 2(p + r)$ or $2q < p + r$, our numerical experiments in Section 4.2 (and in particular Tables 4.1 to 4.3) suggest that these bounds on the ratio gaps are not sharp due to the fact that our lower bounds on $\sigma_{3,\epsilon}$ from Theorem 3.2(iii) can be improved to match the upper bounds when $q \neq p + r$. We then conjecture that, when $q > p+r$, $\frac{\sigma_{2,\epsilon}}{\sigma_{3,\epsilon}} \leq \Xi_1 \epsilon^{p+r-\beta}$, and when $q < p+r$ we have $\frac{\sigma_{2,\epsilon}}{\sigma_{3,\epsilon}} \leq \Xi_2 \epsilon^{q-\beta}$, as summarized in Conjecture 1.2 in Subsection 1.3.2.

Finally with the spectral gap results established we can characterize the geometry of the second eigenfunction $\varphi_{2,\epsilon}$ and show that as $\epsilon \downarrow 0$ this eigenfunction is nearly aligned with the second eigenfunction $\varphi_{2,0}$ of $\mathcal{L}_0$ for certain choices of (p, q, r).

Theorem 3.4 *(Geometry of the second eigenfunction $\varphi_{2,\epsilon}$). Suppose the conditions of Theorem 3.2 are satisfied. Then there exists $\Xi, \epsilon_0 > 0$ so that $\forall(\epsilon, \beta) \in (0, \epsilon_0) \times (0, 1)$*

$$\left| 1 - \left\langle \frac{\varphi_{2,\epsilon}}{\varrho_\epsilon^r}, \frac{\bar{\varphi}_{2,0}}{\varrho_\epsilon^r} \right\rangle^2_{\varrho_\epsilon^{p+r}} \right| \leq \Xi \epsilon^{\min\{\frac{1}{2}, \frac{p+r}{2}, -|q-(p+r)|+\min\{q,p+r\}-\beta\}},$$

where $\bar{\varphi}_{2,0}$ denotes the normalization of $\varphi_{2,0}$ in $L^2(\mathcal{Z}, \varrho_\epsilon^{p-r})$.

We prove this theorem in Subsection 6.3 by bounding the difference between $\varphi_{2,\epsilon}$ and $\varphi_{F,\epsilon}$ in Proposition 6.9 and then the difference between $\bar{\varphi}_{2,0}$ and $\varphi_{F,\epsilon}$ in Proposition 6.11 and invoking the triangle inequality. Note that the above bound blows up if $2q < p + r$ in the unbalanced case where $q < p + r$ and if $2(p + r) < q$ in the unbalanced case where $q > p + r$. Put simply, if the difference between q and $p + r$ is too large then we may lose convergence of the second eigenfunctions. However, we also expect these conditions are not sharp since they rely on our lower bounds on $\sigma_{3,\epsilon}$ in Theorem 3.2(iii) that we conjectured can be sharpened above. Theorem 3.4 is a detailed statement of Main Result 1.1(ii).

Remark 3.5. Two concrete messages follow from Theorems 3.2 and 3.4: (1) Theorem 3.2(iii) tells us that particular care is needed when looking for a spectral gap characterizing the number of clusters if $q \neq p + r$ as the gap may only be manifest in ratio form, not absolutely, leading to potential overestimation of the number of clusters; (2) Theorem 3.4 tells us the form and geometry of the Fiedler vector which characterizes the two clusters, and its dependence on ϱ_0 and on ϵ; whether or not the problem is balanced determines whether the Fielder vector is approximately piecewise constant, or whether it exhibits smoother transitions across the data. These two observations may be useful to practitioners when interpreting graph Laplacian based analysis of large data sets.

4. Numerical experiments in the continuum

In this section we exemplify, and extend, the main theoretical results stated in the previous section. In Subsections 4.1 and 4.2 we study binary clustered data. The numerical results in these subsections highlight the effects of the parameters (p, q, r) on spectral properties: Subsection 4.1 addresses the balanced case

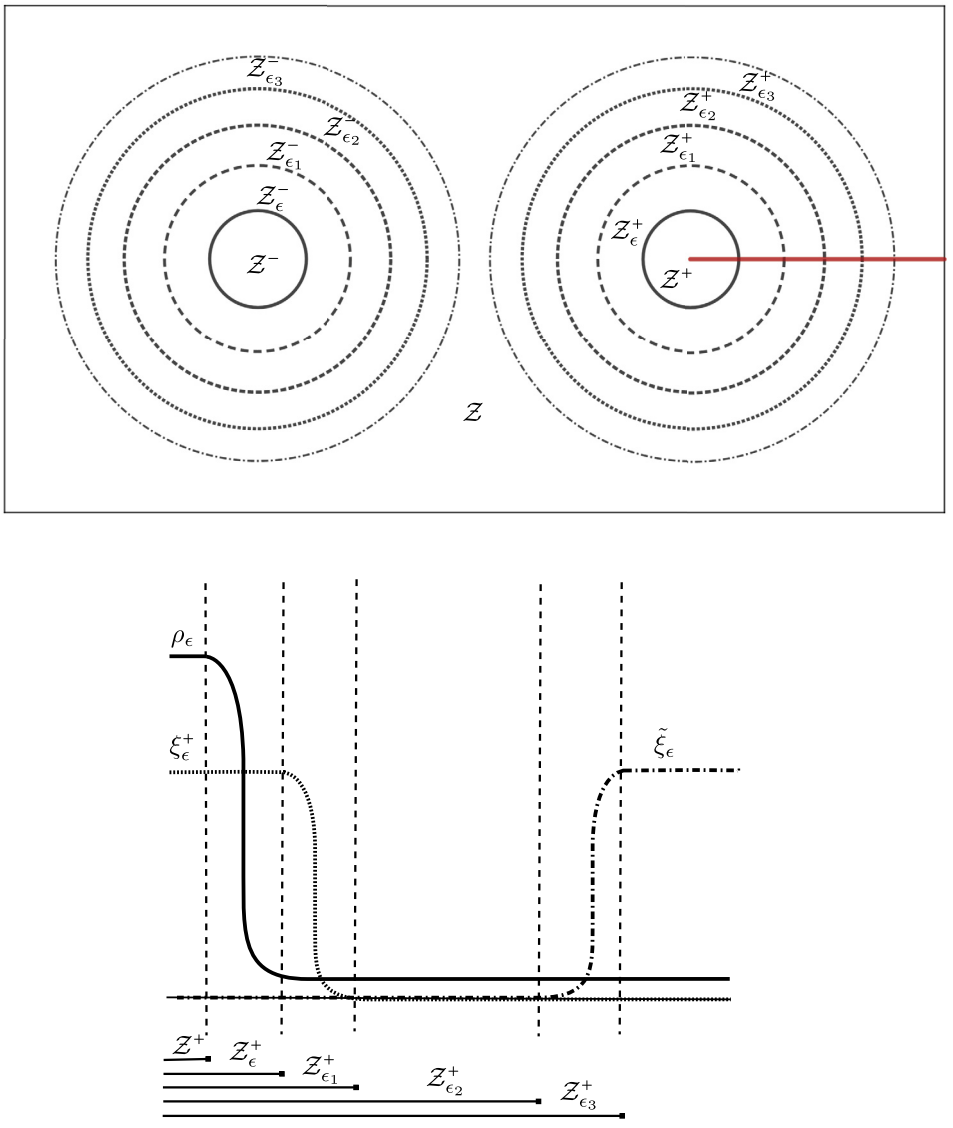


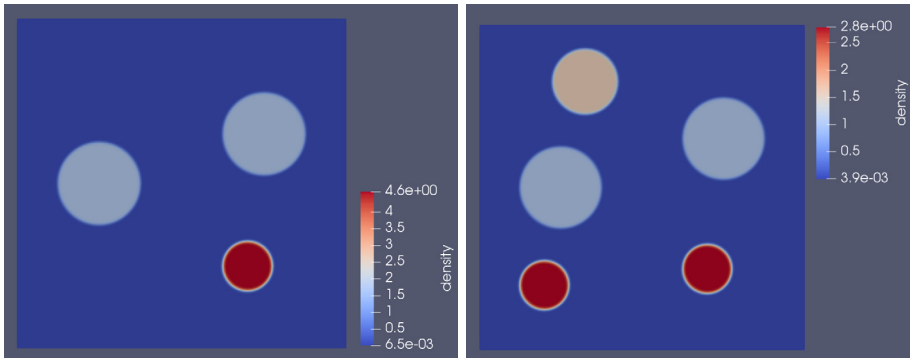
Fig. 4.1. Schematic depiction of the different sets and functions used in construction of $\varphi_{F,\epsilon}$ from (6.3) and $\tilde{\varphi}_{F,\epsilon}$ from (6.15). Top: overhead schematic of the sets $\mathcal{Z}, \mathcal{Z}^\pm, \mathcal{Z}^\pm_\epsilon, \mathcal{Z}^\pm_{\epsilon_1}, \mathcal{Z}^\pm_{\epsilon_2}$, and $\mathcal{Z}^\pm_{\epsilon_3}$. Bottom: cross-section view of ρ_ϵ, ξ^+_ϵ and $\tilde{\xi}_\epsilon$ close to the subset $\mathcal{Z}^+$ along the red line in the top figure. Here, $\epsilon_1 := \epsilon + \epsilon^\beta$, $\epsilon_2 := \epsilon_0 + \epsilon^\beta$, and $\epsilon_3 := \epsilon_0 + 2\epsilon^\beta$ for $\epsilon \in (0, \epsilon_0)$ and $\beta \in (0, 1)$. The function $\varphi_{F,\epsilon}$ is constructed using ϵ_1, concentrates on the clusters, and allows to prove an upper bound on $\sigma_{2,\epsilon}$; the function $\tilde{\varphi}_{F,\epsilon}$ is constructed using ϵ_2 and ϵ_3, concentrates away from the clusters, and allows to prove an upper bound on $\sigma_{3,\epsilon}$. The vertical dashed lines indicate the boundaries of the different sets as indicated below the figure.

where $q = p + r$ and Subsection 4.2 the unbalanced case where $q > p + r$. In Subsection 4.4 we also extend the main theoretical results by considering data comprised of three clusters and five clusters, showing that the intuition from the binary case extends naturally to more than two clusters.

Our numerical simulations in the binary, unbalanced case extend the main theoretical results as they demonstrate the spectral ratio gap of Corollary 3.3, arising when $q > p + r$ is indeed of $\mathcal{O}(\epsilon^{p+r})$ and when $q < p + r$ is of $\mathcal{O}(\epsilon^q)$ suggesting the lower bound on $\sigma_{3,\epsilon}$ can be sharpened.

We proceed by outlining the setting of the numerical experiments. Consider the eigenvalue problem (2.15):

$$\left\langle \varrho_\epsilon^q \nabla \left(\frac{\varphi_{j,\epsilon}}{\varrho_\epsilon^r} \right), \nabla \left(\frac{v}{\varrho_\epsilon^r} \right) \right\rangle = \sigma_{j,\epsilon} \left\langle \varrho_\epsilon^{p+r} \frac{\varphi_{j,\epsilon}}{\varrho_\epsilon^r}, \frac{v}{\varrho_\epsilon^r} \right\rangle, \qquad \varphi_{j,\epsilon}, v \in V^1(\mathcal{Z}, \varrho_\epsilon). \tag{4.1}$$



(a) (b)

Fig. 4.2. Plot of the densities ϱ_ϵ of the form (4.2) with three and five clusters for $\epsilon = 0.0125$.

Our numerics are all performed in dimension $d = 2$. We solve this by the finite element method using the FEniCS software package [31]. We work with the variables $\varphi_{j,\epsilon}/\varrho_\epsilon^r$ and v/ϱ_ϵ^r, rather than directly with $\varphi_{j,\epsilon}$ and v, and discretize these ϱ_ϵ^r scaled variables using the standard linear finite element basis functions in $H^1(\mathcal{Z})$. We approximate ϱ_ϵ using quadratic finite element basis functions. Throughout we take $\mathcal{Z} \equiv (-1,1) \times (-1,1)$. We consider ϵ in the range $(1/1280, 1/10)$. For each value of ϵ, we approximate the eigenvalue problem (4.1) using a mesh of 1.28×10^6 triangular elements defined on a uniform grid of 800×800 nodes. This finite element discretization leads to a generalized matrix eigenvalue problem which is solved using a Krylov-Schur eigenvalue solver in PETSc [3] with a tolerance of 10^{-9}.

Throughout this section we use densities of the form

$$\varrho_\epsilon(s) = C^{-1}\left(\epsilon + \sum_{i=1}^{K} \frac{\operatorname{erf}\left(\epsilon^{-1}(\theta_i - |s - c_i|)\right)}{4\pi\theta_i^2}\right), \qquad \forall s \in \mathcal{Z}, \tag{4.2}$$

where $|\cdot|$ is the two dimensional Euclidean norm, K is the number of circular clusters, c_i denotes the i^{th} cluster center, θ_i the i^{th} cluster radius, and C is a normalizing parameter to make sure that ϱ_ϵ is a probability distribution. In Subsections 4.1 and 4.2 we consider two clusters with parameters $c_1 = (-0.5, 0.0)$, $\theta_1 = 0.25$, $c_2 = (0.5, 0.3)$, and $\theta_2 = 0.25$ as shown in Fig. 1.1(a). In Subsection 4.4 we consider three and five clusters adding the point $c_3 = (0.4, -0.5)$ with radius $\theta_3 = 0.15$, to make three clusters, and then adding $c_4 = (-0.35, 0.65)$ and $c_5 = (-0.6, -0.6)$ with radii $\theta_4 = 0.20$ and $\theta_5 = 0.15$, to generate five clusters. We plot the resulting densities in Fig. 4.2.

4.1. Binary balanced case: $q = p + r$

In Fig. 4.3(a) we plot $\sigma_{2,\epsilon}$ in the balanced case $r = p$, $q = p + r$ and $p \in [0.5, 2]$. For a given value of p each symbol denotes the numerical approximation to $\sigma_{2,\epsilon}$, and the line denotes the best fit determined via linear regression; in the regression we only use data from $\epsilon \leq 0.025$ as consistent asymptotic behavior for $\epsilon \downarrow 0$ is observed in this regime. Theorem 3.2(ii) predicts that $\sigma_{2,\epsilon} = \mathcal{O}(\epsilon^{q-\beta})$ for arbitrarily small $\beta > 0$. Then we expect to observe a slope of approximately $2p$ for each set of simulations. We report the numerical slopes in brackets in the legend of Fig. 4.3(a), and compare the numerical slopes to the analytic prediction in the first four rows of Table 4.1.

In Fig. 4.3(b), we plot the ratio $\sigma_{2,\epsilon}/\sigma_{3,\epsilon}$ for different values of ϵ. By Corollary 3.3 we expect $\sigma_{3,\epsilon}$ to be uniformly bounded away from zero implying that $\sigma_{2,\epsilon}/\sigma_{3,\epsilon} = \mathcal{O}(\epsilon^{q-\beta})$ and so the numerical slopes in Fig. 4.3(b) should be close to $2p$ We compare the numerical slopes to the analytic slopes for the spectral ratio gap in the first four rows of Table 4.1.

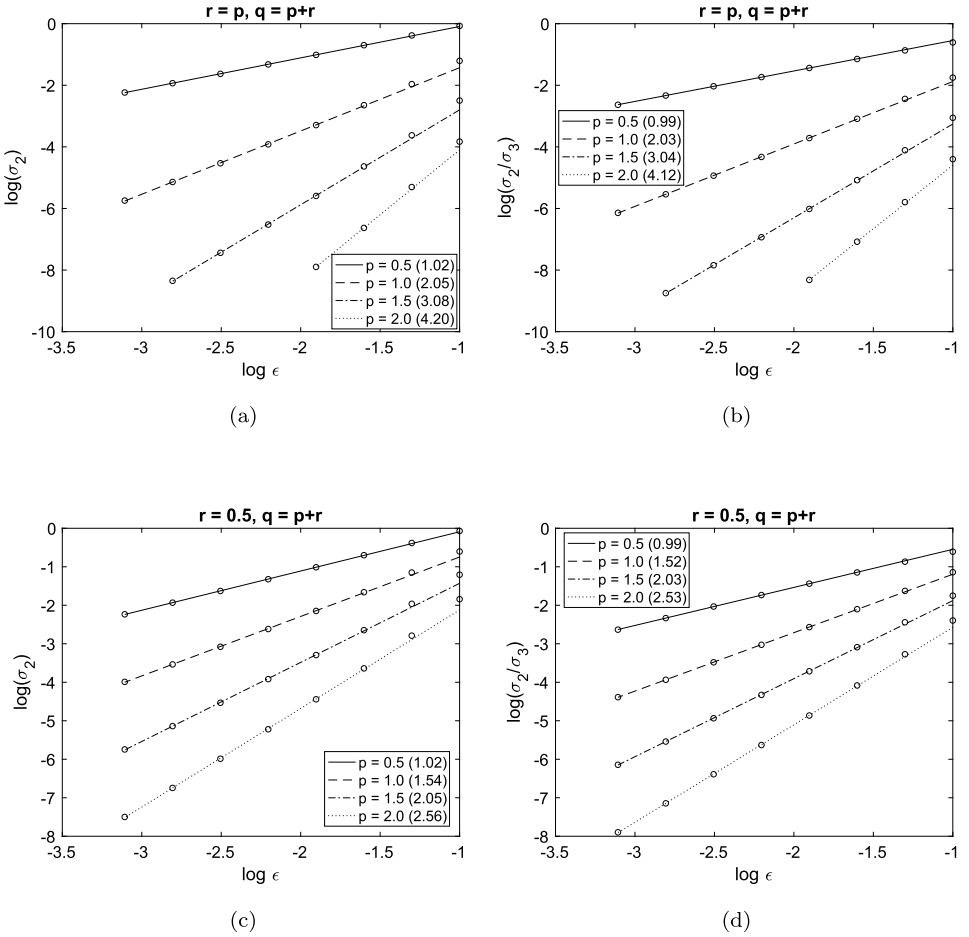


Fig. 4.3. Variation of the second and third eigenvalues of $\mathcal{L}_\epsilon$ in the balanced case with $q = p+r$ and for various values of $p \in [0.5, 2]$. (a, b) consider $r = p$; (c, d) consider fixed $r = 0.5$. (a, c) show $\log(\sigma_{2,\epsilon})$ vs $\log(\epsilon)$ while (b, d) show $\log(\sigma_{2,\epsilon}/\sigma_{3,\epsilon})$ vs $\log(\epsilon)$. The values reported in the brackets in the legends are numerical approximations to the slope of the lines for different values of p.

Table 4.1
Comparison between numerical approximation of the rate of decay of $\log(\sigma_{2,\epsilon})$ and $\log(\sigma_{2,\epsilon}/\sigma_{3,\epsilon})$ as functions of $\log(\epsilon)$ and the analytic predictions in Theorem 3.2 and Corollary 3.3 for the balanced case with $q = p + r$ and different choices of p and r.

		$\frac{\log(\sigma_{2,\epsilon})}{\log \epsilon}$		$\frac{\log(\sigma_{2,\epsilon})-\log(\sigma_{3,\epsilon})}{\log \epsilon}$	
p	r	Analytic	Numerical	Analytic	Numerical
0.5	0.5	1.00	1.02	1.00	0.99
1.0	1.0	2.00	2.05	2.00	2.03
1.5	1.5	3.00	3.08	3.00	3.04
2.0	2.0	4.00	4.20	4.00	4.12
1.0	0.5	1.50	1.54	1.50	1.52
1.5	0.5	2.00	2.05	2.00	2.03
2.0	0.5	2.50	2.56	2.50	2.53

In Fig. 4.3(c,d) we repeat the above study of the second and third eigenvalues for the balanced case $q = p+r$ but this time we fix $r = 0.5$ and vary $p \in (0.5, 2)$. We see similar results to Fig. 4.3(a,b) in that the numerical slopes are in good agreement with the predicted slopes of $q = p + r$. We compare the numerical and analytic slopes for this experiment in the last three rows of Table 4.1.

In summary we note that, in this binary balanced setting the numerical experiments match the theory, quantitatively. The slopes are less accurate for higher values of p. We attribute this to the smaller values of the eigenvalues in these cases, which are evaluated with less numerical precision.

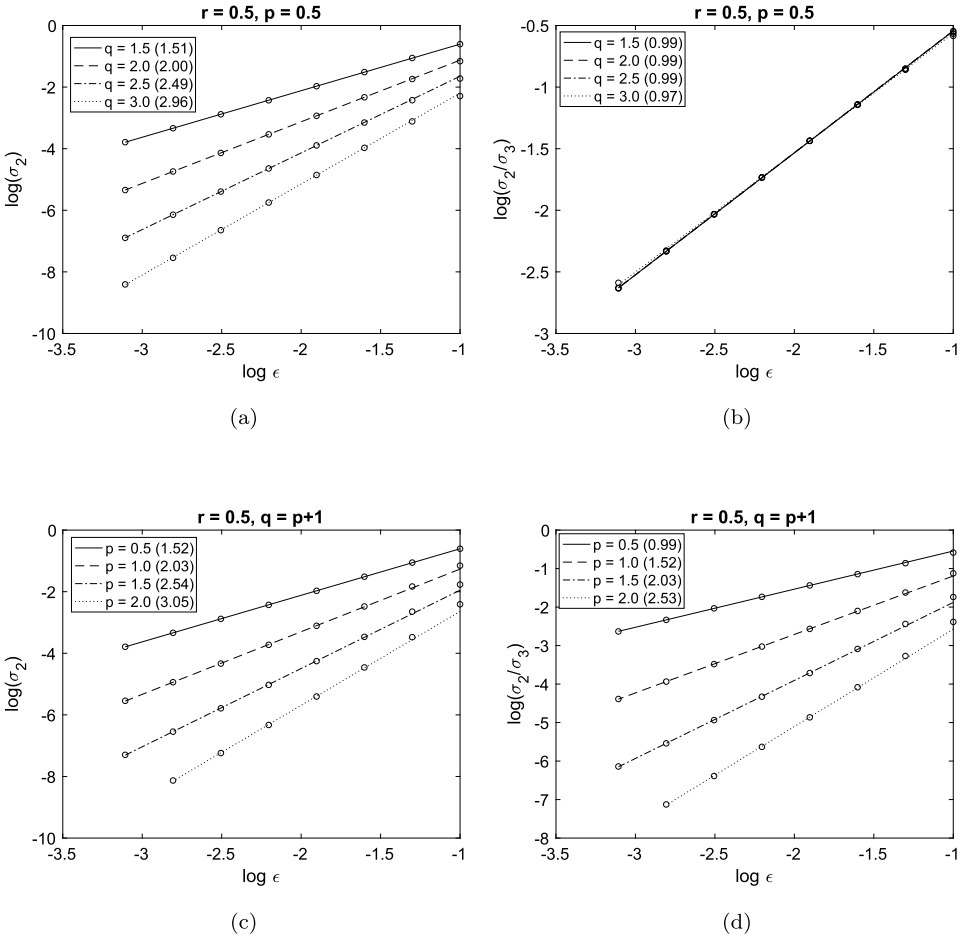


Fig. 4.4. Variation of the second and third eigenvalues of $\mathcal{L}_\epsilon$ in the unbalanced case with $q > p+r$, and for various values of p, q and r. In (a, b) we fix $p = r = 0.5$ and vary $q \in [1.5, 3]$. In (c, d) we fix $r = 0.5$, $q = p+1$ and vary $p \in [0.5, 2]$. (a, c) show $\log(\sigma_{2,\epsilon})$ vs $\log(\epsilon)$ while (b, d) show $\log(\sigma_{2,\epsilon}/\sigma_{3,\epsilon})$ vs $\log(\epsilon)$. The values reported in the brackets in the legends are numerical approximations to the slope of the lines.

4.2. Binary unbalanced case: $q > p + r$

We now turn our attention to the spectrum of $\mathcal{L}_\epsilon$ when $q > p + r$. In Fig. 4.4(a, b) we plot the second eigenvalue $\sigma_{2,\epsilon}$ and the ratio $\sigma_{2,\epsilon}/\sigma_{3,\epsilon}$ for $p = r = 0.5$ and vary q in the range $(1.5, 3)$. As before we fit a line to the computed values of the eigenvalue and the ratio for each value of q and report the numerical slope in brackets in the legend; once again we fit the line to data points with $\epsilon \leq 0.025$ where the $\epsilon \downarrow 0$ regime is manifest. We observe that $\sigma_{2,\epsilon} = \mathcal{O}(\epsilon^q)$ as in the balanced case while the ratio $\sigma_{2,\epsilon}/\sigma_{3,\epsilon} = \mathcal{O}(\epsilon^{p+r})$ which is better than the predicted $\mathcal{O}(\epsilon^{2(p+r)-q})$ rate in Corollary 3.3. As mentioned earlier, these results suggest that the lower bound on $\sigma_{3,\epsilon}$ in Theorem 3.2(iii) can be sharpened to match the upper bound. In Fig. 4.4(c, d), we consider another case with $q > p+r$ but this time we fix $r = 0.5$ vary $p \in (0.5, 2)$ and take $q = p+1$. Once again we observe that $\sigma_{2,\epsilon} \sim \epsilon^q$, which is consistent with Theorem 3.2(ii), and $\sigma_{2,\epsilon}/\sigma_{3,\epsilon} \sim \epsilon^{p+r}$, which is better than the predicted rate in Theorem 3.2(iii); again the results suggest that the lower bound on $\sigma_{3,\epsilon}$ can be sharpened to match the upper bound. We compare the numerical slopes with the analytic upper bounds and with the conjectured $\mathcal{O}(\epsilon^{p+r})$ rate for the spectral ratio gap in Table 4.2.

In summary we note that, in this binary unbalanced setting the numerical experiments are consistent with the theory insight that only a spectral ratio gap will manifest between the second and third eigenvalues. Furthermore, these experiments suggest that the lower and upper bounds on the third eigenvalue should

Table 4.2
Comparison between numerical approximation of the rate of decay of $\log(\sigma_{2,\epsilon})$ and $\log(\sigma_{2,\epsilon}/\sigma_{3,\epsilon})$ as functions of $\log(\epsilon)$ and the analytic predictions in Theorem 3.2 and Corollary 3.3. The last column denotes the conjectured slope of $p+r$ for $\log(\sigma_{2,\epsilon}/\sigma_{3,\epsilon})$ for the unbalanced case $q > p+r$. Negative analytic rates are omitted.

			$\frac{\log(\sigma_{2,\epsilon})}{\log\epsilon}$		$\frac{\log(\sigma_{2,\epsilon})-\log(\sigma_{3,\epsilon})}{\log\epsilon}$		
p	q	r	Analytic	Numerical	Analytic	Numerical	$p+r$
0.5	1.50	0.5	1.50	1.51	0.5	0.99	1.0
0.5	2.0	0.5	2.00	2.00	0.0	0.99	1.0
0.5	2.5	0.5	2.49	2.57	–	0.99	1.0
0.5	3.0	0.5	2.96	3.06	–	0.97	1.0
1.0	2	0.5	2.03	2.11	1.0	1.52	1.5
1.5	2.5	0.5	2.54	2.64	1.5	2.03	2.0
2.0	3.0	0.5	3.05	3.20	2.0	2.53	2.5

Table 4.3
Comparison between numerical approximation of the rate of decay of $\log(\sigma_{2,\epsilon})$ and $\log(\sigma_{2,\epsilon}/\sigma_{3,\epsilon})$ as functions of $\log(\epsilon)$ and the analytic predictions in Theorem 3.2 and Corollary 3.3 for the unbalanced case $q < p+r$. Compare values in the last column with the prescribed values of q. Negative analytic rates are omitted.

			$\frac{\log(\sigma_{2,\epsilon})}{\log\epsilon}$		$\frac{\log(\sigma_{2,\epsilon})-\log(\sigma_{3,\epsilon})}{\log\epsilon}$	
p	q	r	Analytic	Numerical	Analytic	Numerical
1	0.5	1	0.5	0.56	–	0.49
1	1.0	1	1.0	1.07	0.0	1.02
1	1.5	1	1.5	1.56	1.0	1.74
0.5	0.5	1	0.5	0.50	–	0.49
1.5	1.5	1	1.5	1.58	0.5	1.53
2.0	2.0	1	2.0	2.09	1.0	2.03

match, suggesting tighter bounds on the spectral ratio gap could be achievable forming the foundation for the first component of Conjecture 1.2 in Subsection 1.3.2.

4.3. Binary unbalanced case: $q < p+r$

Next we turn our attention to the spectrum of $\mathcal{L}_\epsilon$ when $q < p+r$. Fig. 4.5(a,b) shows the second eigenvalues $\sigma_{2,\epsilon}$ as well as the ratio $\sigma_{2,\epsilon}/\sigma_{3,\epsilon}$ for $p = r = 1$ and $q \in [0.5, 1.5]$. Once again we fit a line to the computed values of the eigenvalues and the ratios and report the slopes within brackets in the legends. We observe that $\sigma_{2,\epsilon} = \mathcal{O}(\epsilon^q)$ as in the $q \geq p+r$ cases; however we also notice that the ratio $\sigma_{2,\epsilon}/\sigma_{3,\epsilon} = \mathcal{O}(\epsilon^q)$, an observation which suggests that Corollary 3.3(ii) can be improved; this in turn would be possible if we could sharpen our lower bound on $\sigma_{3,\epsilon}$ in Theorem 3.2(iii) to match the upper bound, resulting in a uniform spectral gap.

Fig. 4.5(c,d) shows further examples with $q < p+r$ this time with $r = 1$ fixed and taking $q = p \in [0.5, 2.0]$. Once again we observe that $\sigma_{2,\epsilon} = \mathcal{O}(\epsilon^q)$ while $\sigma_{2,\epsilon}/\sigma_{3,\epsilon} = \mathcal{O}(\epsilon^q)$ as well, further reaffirming our conjecture that the lower bound in Theorem 3.2(iii) is too pessimistic. We compare the analytic and numerical slopes for the second eigenvalues as well as the spectral ratio in Table 4.3.

To summarize we derive two conclusions in this unbalanced case: First, our bounds on the second eigenvalue $\sigma_{2,\epsilon}$ are sharp but our bounds on the spectral ratio $\sigma_{2,\epsilon}/\sigma_{3,\epsilon}$ are not sharp similarly to the $q > p+r$ case and due to the fact that our lower bound on $\sigma_{3,\epsilon}$ is too pessimistic. Second, followed by this observation we expect a uniform spectral gap to manifest between the second and third eigenvalues in the unbalanced regime where $q < p+r$, similarly to the balanced regime $q = p+r$. These observations further support the first component of Conjecture 1.2 from Subsection 1.3.2.

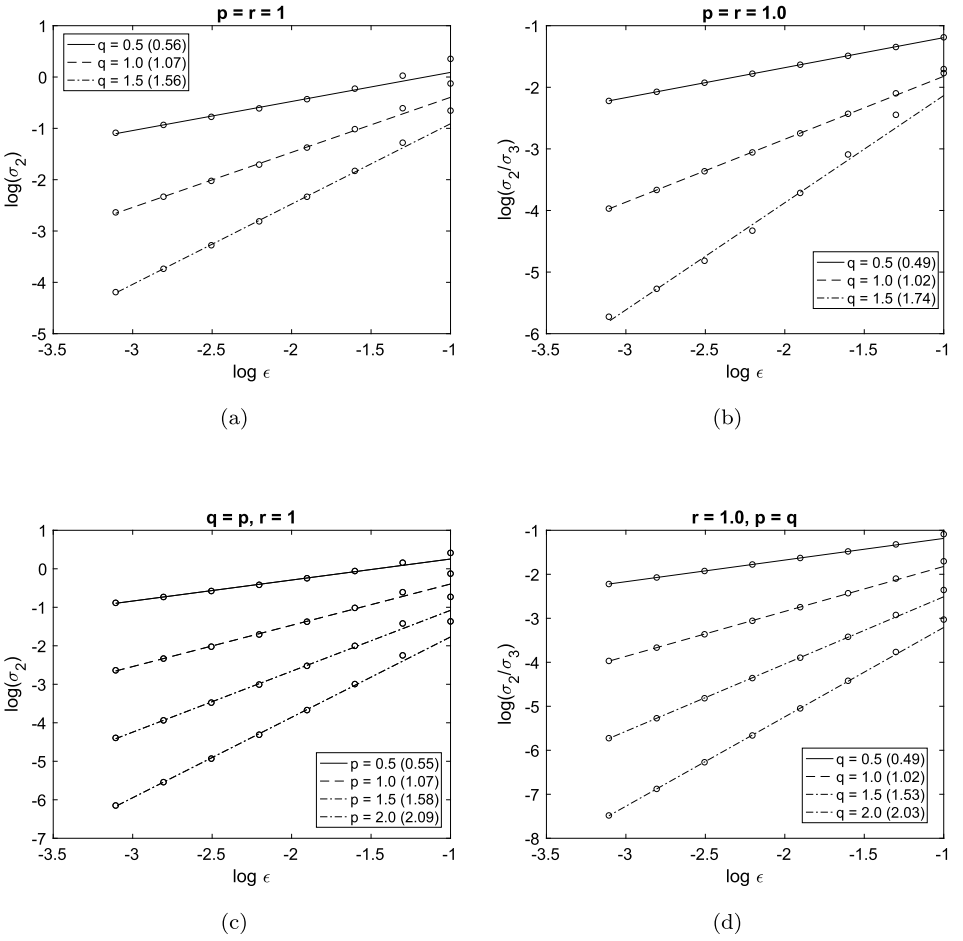


Fig. 4.5. Variation of the second and third eigenvalues of $\mathcal{L}_\epsilon$ in the unbalanced case with $q < p + r$, and for various values of p, q and r. In (a, b) we fix $p = r = 1$ and vary $q \in [0.5, 1.5]$. In (c, d) we fix $r = 1.0$, $q = p$ and vary $p \in [0.5, 2]$. (a, c) show $\log(\sigma_{2,\epsilon})$ vs $\log(\epsilon)$ while (b, d) show $\log(\sigma_{2,\epsilon}/\sigma_{3,\epsilon})$ vs $\log(\epsilon)$. The values reported in the brackets in the legends are numerical approximations to the slope of the lines.

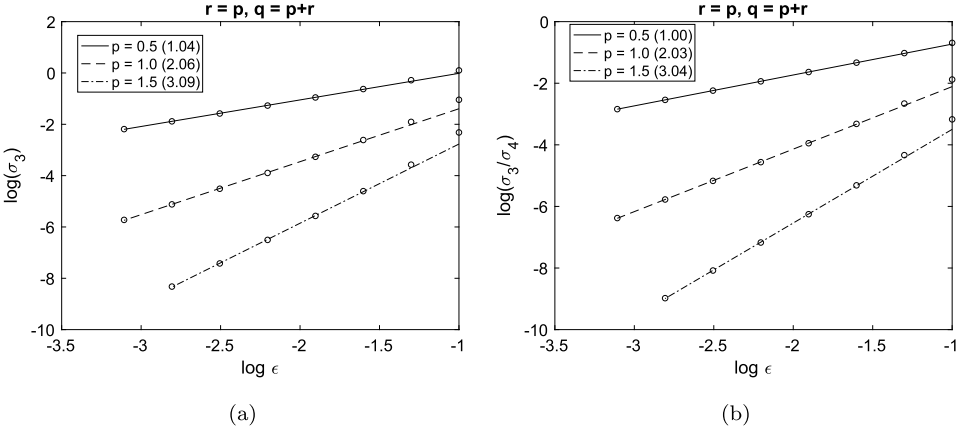
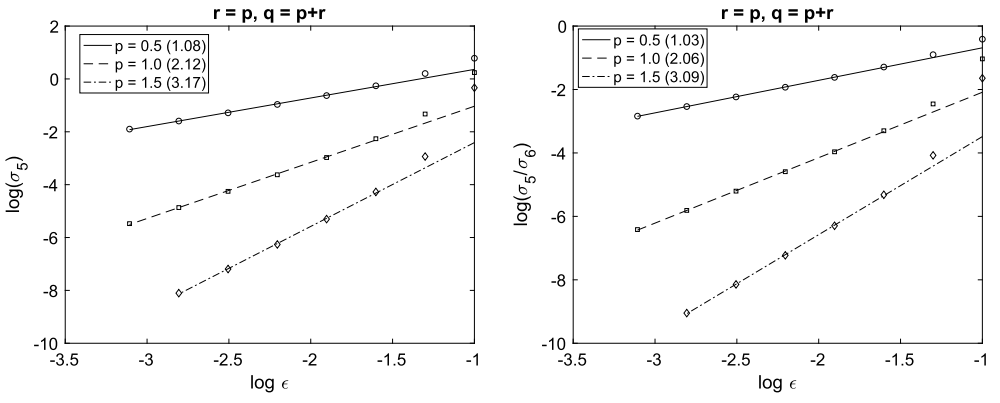
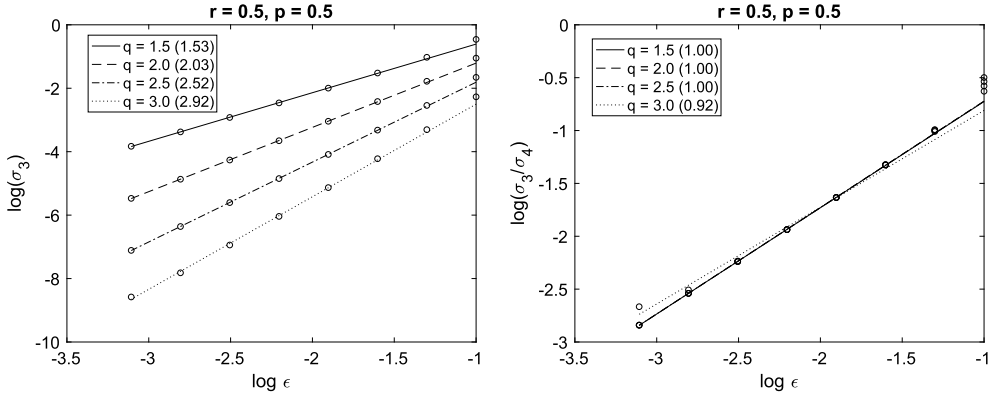


Fig. 4.6. Variation of the third and fourth eigenvalues of $\mathcal{L}_\epsilon$ in the three cluster setting with $q = p + r$, $r = p$ and for $p \in [0.5, 1.5]$. (a) shows $\log(\sigma_{3,\epsilon})$ vs $\log(\epsilon)$ while (b) shows $\log(\sigma_{3,\epsilon}/\sigma_{4,\epsilon})$ vs $\log(\epsilon)$. The values reported in the brackets in the legends are numerical approximations to the slope of the lines for different values of p.



(a) (b)

Fig. 4.7. Variation of the fifth and sixth eigenvalues of $\mathcal{L}_\epsilon$ in the five cluster case with $q = p + r$, $r = p$ and for $p \in [0.5, 1.5]$. (a) shows $\log(\sigma_{5,\epsilon})$ vs $\log(\epsilon)$ while (b) shows $\log(\sigma_{5,\epsilon}/\sigma_{6,\epsilon})$ vs $\log(\epsilon)$. The values reported in the brackets in the legends are numerical approximations to the slope of the lines for different values of p.



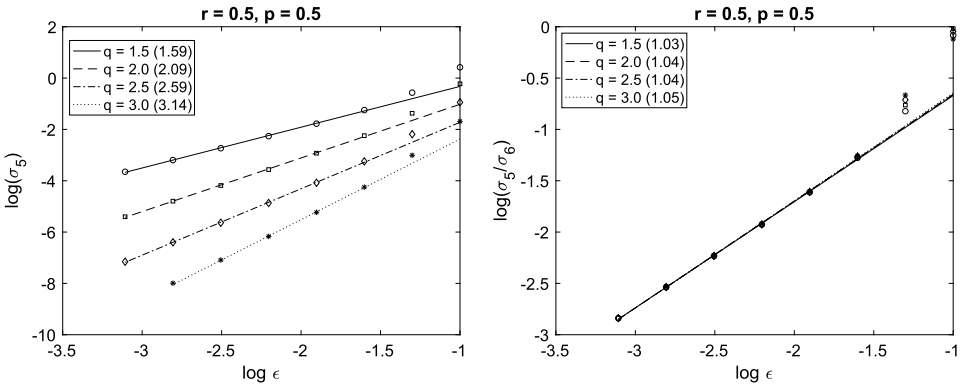
(a) (b)

Fig. 4.8. Variation of the third and fourth eigenvalues of $\mathcal{L}_\epsilon$ in the three cluster setting with $q > p + r$, $r = p = 0.5$ and for $q \in [1.5, 3]$. (a) shows $\log(\sigma_{3,\epsilon})$ vs $\log(\epsilon)$ while (b) shows $\log(\sigma_{3,\epsilon}/\sigma_{4,\epsilon})$ vs $\log(\epsilon)$. The values reported in the brackets in the legends are numerical approximations to the slope of the lines for different values of q.

4.4. Multiple clusters

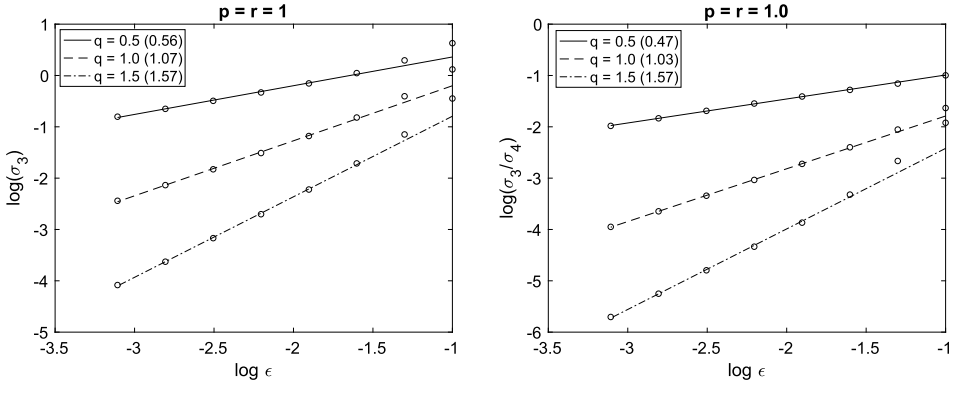
We now consider two densities ϱ_ϵ which concentrate, respectively, on three and five clusters for small ϵ; the quantitative details are given in (4.2) and the text following; see Fig. 4.2. In Figs. 4.6 and 4.7 we display the behavior of the K^{th} eigenvalue and the spectral ratio gap related to it, for $K = 3$ and $K = 5$ respectively. In both cases we let $q = p + r$ and plot $\log(\sigma_{K,\epsilon})$ and $\log(\sigma_{K,\epsilon}/\sigma_{K+1,\epsilon})$ against $\log(\epsilon)$. The numerics are consistent with the hypothesis that $\sigma_{K,\epsilon} \sim \sigma_{K,\epsilon}/\sigma_{K+1,\epsilon} \sim \mathcal{O}(\epsilon^q)$. This suggests a natural extension of Theorem 3.2 and Corollary 3.3 from the binary case to multiple clusters.

In Figs. 4.8–4.11 we collect similar results for the unbalanced regime where $q \neq p + r$. Once again we see strong evidence that the multi-cluster setting behaves similarly to the binary case in that $\sigma_{K,\epsilon} \sim \epsilon^q$ while $\sigma_{K,\epsilon}/\sigma_{K+1,\epsilon} \sim \epsilon^{p+r}$ when $q > p + r$ and $\sigma_{K,\epsilon}/\sigma_{K+1,\epsilon} \sim \epsilon^q$ when $q < p + r$ in both the three and five cluster cases. We provide further evidence for this conjecture in Tables 4.4 and 4.5 where we collect numerical approximations to the above rates for different choices of p, q, r in the balanced and unbalanced regimes. The above results lead to second component of Conjecture 1.2 appearing in Subsection 1.3.2.



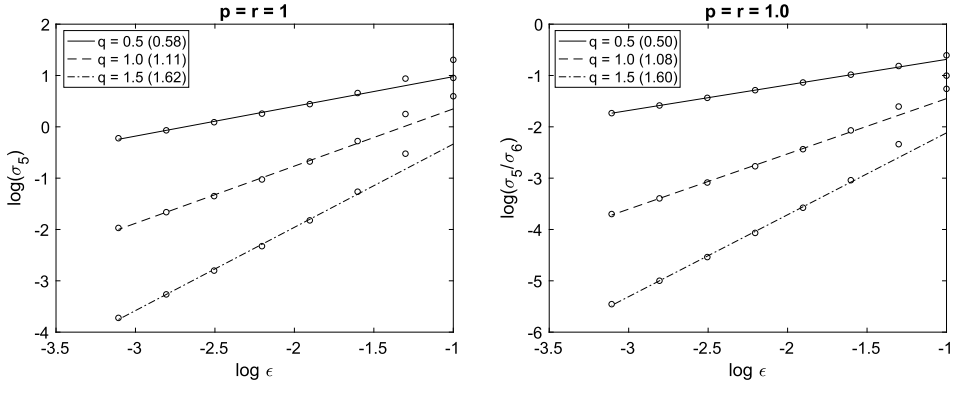
(a) (b)

Fig. 4.9. Variation of the fifth and sixth eigenvalues of $\mathcal{L}_\epsilon$ in the five cluster case with $q > p + r$, $r = p = 0.5$ and for $q \in [1.5, 3]$. (a) shows $\log(\sigma_{5,\epsilon})$ vs $\log(\epsilon)$ while (b) shows $\log(\sigma_{5,\epsilon}/\sigma_{6,\epsilon})$ vs $\log(\epsilon)$. The values reported in the brackets in the legends are numerical approximations to the slope of the lines for different values of q.



(a) (b)

Fig. 4.10. Variation of the third and fourth eigenvalues of $\mathcal{L}_\epsilon$ in the three cluster setting with $q < p + r$, $r = p = 1$ and for $q \in [0.5, 1.5]$. (a) shows $\log(\sigma_{3,\epsilon})$ vs $\log(\epsilon)$ while (b) shows $\log(\sigma_{3,\epsilon}/\sigma_{4,\epsilon})$ vs $\log(\epsilon)$. The values reported in the brackets in the legends are numerical approximations to the slope of the lines for different values of q.



(a) (b)

Fig. 4.11. Variation of the fifth and sixth eigenvalues of $\mathcal{L}_\epsilon$ in the five cluster case with $q < p + r$, $r = p = 1$ and for $q \in [0.5, 1.5]$. (a) shows $\log(\sigma_{5,\epsilon})$ vs $\log(\epsilon)$ while (b) shows $\log(\sigma_{5,\epsilon}/\sigma_{6,\epsilon})$ vs $\log(\epsilon)$. The values reported in the brackets in the legends are numerical approximations to the slope of the lines for different values of q.

Table 4.4
Numerical approximation of the rate of decay of $\log(\sigma_{3,\epsilon})$ and $\log(\sigma_{3,\epsilon}/\sigma_{4,\epsilon})$ as functions of $\log(\epsilon)$ for different choices of p, q, r in the three cluster setting.

	p	q	r	$\frac{\log(\sigma_{3,\epsilon})}{\log\epsilon}$	$\frac{\log(\sigma_{4,\epsilon})-\log(\sigma_{3,\epsilon})}{\log\epsilon}$
$q = p + r$	0.5	1.0	0.5	1.04	1.00
	1.0	2.0	1.0	2.06	2.03
	1.5	3.0	1.5	3.097	3.04
	1.0	1.5	0.5	1.55	1.52
	1.5	2.0	0.5	2.06	2.03
	2.0	2.5	0.5	2.57	2.53
$q > p + r$	0.5	1.5	0.5	1.53	1.00
	0.5	1.0	0.5	1.04	1.00
	0.5	2.0	0.5	2.03	1.00
	0.5	2.5	0.5	2.52	1.00
	0.5	3.0	0.5	2.92	0.92
	1.0	2.0	0.5	2.05	1.52
	1.5	2.5	0.5	2.55	2.03
	2.0	3.0	0.5	3.07	2.53
$q < p + r$	1.0	0.5	1.0	0.56	0.47
	1.0	1.0	1.0	1.07	1.03
	1.0	1.5	1.0	1.57	1.57
	0.5	0.5	1.0	0.54	0.49
	1.5	1.5	1.0	1.58	1.54
	2.0	2.0	1.0	2.09	2.04

Table 4.5
Numerical approximation of the rate of decay of $\log(\sigma_{5,\epsilon})$ and $\log(\sigma_{5,\epsilon}/\sigma_{6,\epsilon})$ as functions of $\log(\epsilon)$ for different choices of p, q, r in the five cluster setting.

	p	q	r	$\frac{\log(\sigma_{4,\epsilon})}{\log\epsilon}$	$\frac{\log(\sigma_{5,\epsilon})-\log(\sigma_{4,\epsilon})}{\log\epsilon}$
$q = p + r$	0.5	1.0	0.5	1.04	1.03
	1.0	2.0	1.0	2.12	2.06
	1.5	3.0	1.5	3.17	3.09
	1.0	1.5	0.5	1.61	1.55
	1.5	2.0	0.5	2.12	2.06
	2.0	2.5	0.5	2.63	2.57
$q > p + r$	0.5	1.5	0.5	1.59	1.03
	0.5	2.0	0.5	2.09	1.04
	0.5	2.5	0.5	2.59	1.04
	0.5	3.0	0.5	3.14	1.05
	1.0	2.0	0.5	2.11	1.56
	1.5	2.5	0.5	2.62	2.07
	2.0	3.0	0.5	3.16	2.59
$q < p + r$	1.0	0.5	1.0	0.58	0.50
	1.0	1.0	1.0	1.11	1.08
	1.0	1.5	1.0	1.62	1.60
	0.5	0.5	1.0	0.57	0.52
	1.5	1.5	1.0	1.63	1.59
	2.0	2.0	1.0	2.14	2.10

5. From discrete to continuum

In this section we present formal calculations, and numerical experiments, demonstrating that the operators of the form $\mathcal{L}$ in (1.1) arise as the large data limit of L_N as in (1.2) for parameters $(p, q, r) \in \mathbb{R}^3$, and for a density ϱ supported on $\mathcal{Z}$ according to which the vertices $\{x_n\}_{n=1}^N$ are i.i.d. Subsection 5.1 discusses the construction of the discrete operators L_N and their properties including self-adjointness and invariance of the spectrum under parameter choices. Subsection 5.2 outlines a roadmap for rigorous proof of convergence of L_N to $\mathcal{L}$ in the framework of [23,40,20] through the study of the convergence of Dirichlet energies, using the law of large numbers and localization of the weights. These arguments reveal the relationship between the discrete and continuum eigenproblems as well as the correct scaling needed in the discrete setting for the spectra to converge, the topic of Subsection 5.3. In Subsection 5.4 we present numerical experiments

demonstrating the convergence of discrete graph Laplacians to continuum limit operators of the form (1.1), as well as manifestations of the theoretical results of Section 3 in the discrete $N < +\infty$ setting.

5.1. The discrete operator L_N

Let $X_N \in \mathbb{R}^{d\times N}$ denote the matrix with columns $\{x_n\}_{n=1}^N$ sampled i.i.d. from a density ϱ on some domain $\mathcal{Z}$. Following [18], we define a similarity graph on X_N by defining a weighted similarity matrix $\tilde{W}_N$ with entries

$$\tilde{W}_{ij} = \begin{cases} \eta_\delta(|x_i - x_j|)\,, & i \neq j, \\ 0 & i = j, \end{cases}$$

where $|\cdot|$ denotes the Euclidean norm, $\eta_\delta(\cdot) = \delta^{-d}\eta(\cdot/\delta)$ for a suitably chosen edge weight profile $\eta : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ that is non-increasing, continuous at zero and has bounded second moment. Furthermore, let $\tilde{D}_N = \mathrm{diag}(\tilde{d}_i)$ where $\tilde{d}_i := \sum_{j=1}^N \tilde{W}_{ij}$ is the degree of node i. Since η_δ is approximately a Dirac distribution for small $\delta > 0$ it follows that $\tilde{d}_i$ is an empirical approximation of $\varrho(x_i)$. Without loss of generality we assume that the resulting similarity graph has no isolated points: $\tilde{d}_i > 0$ for all i. For $q \in \mathbb{R}$, we introduce the matrix $W_N = W_N(q)$, a re-weighting of $\tilde{W}_N$, with entries

$$W_{ij} = \frac{\tilde{W}_{ij}}{\tilde{d}_i^{\,1-q/2}\tilde{d}_j^{\,1-q/2}},$$

with corresponding degree matrix $D_N = \mathrm{diag}(d_i)$ where $d_i := \sum_{j=1}^N W_{ij}$. We now define the graph Laplacian L_N as in (1.2) for $(p, q, r) \in \mathbb{R}^3$,

$$L_N := \begin{cases} D_N^{\frac{1-p}{q-1}}\,(D_N - W_N)\,D_N^{-\frac{r}{q-1}}, & \text{if } q \neq 1\,, \\ D_N - W_N, & \text{if } q = 1. \end{cases}$$

Let $\langle\cdot,\cdot\rangle$ denote the usual Euclidean inner product. Given a symmetric matrix $A \in \mathbb{R}^{N\times N}$ and vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^N$, we define

$$\langle \mathbf{u}, \mathbf{v}\rangle_A := \mathbf{u}^T A \mathbf{v}\,.$$

The matrix L_N is not self-adjoint with respect to the Euclidean inner product for general (p, q, r) but it is self-adjoint with respect to the following (p, q, r)-weighted inner product:

$$\langle\cdot\,,\,\cdot\rangle_{(p,q,r)} := \begin{cases} \langle\cdot\,,\,\cdot\rangle_{D_N^{\frac{p-1-r}{q-1}}} & \text{if } q \neq 1\,, \\ \langle\cdot\,,\,\cdot\rangle & \text{if } q = 1\,. \end{cases}$$

More precisely, in the case $q \neq 1$, writing $\mathbf{v} = D_N^{-\frac{r}{q-1}}\mathbf{u}$ yields

$$\begin{aligned} \langle \mathbf{u}, L_N\mathbf{u}\rangle_{(p,q,r)} &= \langle D_N^{\frac{p-1}{q-1}}\mathbf{v}, D_N^{\frac{1-p}{q-1}}\,(D_N - W_N)\,\mathbf{v}\rangle = \langle \mathbf{v}, (D_N - W_N)\,\mathbf{v}\rangle \\ &= \frac{1}{2}\sum_{i,j} W_{ij}\,|v_i - v_j|^2 = \frac{1}{2}\sum_{i,j} W_{ij}\left|\frac{u_i}{d_i^{r/(q-1)}} - \frac{u_j}{d_j^{r/(q-1)}}\right|^2. \end{aligned} \tag{5.1}$$

If $q = 1$, we have instead

$$\langle \mathbf{u}, L_N \mathbf{u} \rangle_{(p,1,r)} = \langle \mathbf{u}, (D_N - W_N)\, \mathbf{u} \rangle = \frac{1}{2} \sum_{i,j} W_{ij} \, |u_i - u_j|^2 \, .$$

It immediately follows that the first eigenvalue of L_N is zero with corresponding eigenvector $\boldsymbol{\varphi}_1 = D_N^{r/(q-1)} \mathbf{1}$ if $q \neq 1$ and $\boldsymbol{\varphi}_1 = \mathbf{1}$ if $q = 1$, where $\mathbf{1}$ denotes the constant vector of ones. The symmetric expression (5.1) also shows why the graph Laplacian is a useful tool for spectral clustering: If the corresponding similarity graph has more than one disconnected component, then choices of u_i that take different constant multiples of $d_i^{r/(q-1)}$ (if $q \neq 1$; different constants if $q = 1$) on each component of the graph set $\langle \mathbf{u}, L_N \mathbf{u} \rangle_{(p,q,r)}$ to zero. As a consequence, a simple continuity argument (highlighted in [33]) demonstrates that the eigenvectors corresponding to the low lying spectrum of L_N contain information about the clusters in X_N. Note also that for the more common parameter choices $(p, q, r) = (1, 2, 0)$, $(3/2, 2, 1/2)$ and $(1, 1, 0)$ discussed in the introduction (see Subsection 1.2), the weighted inner product $\langle \cdot \, , \, \cdot \rangle_{(p,q,r)}$ reduces to the usual Euclidean inner product. We say $(\sigma, \mathbf{u})$ is an eigenpair of L_N for parameters (p, q, r) if

$$\langle L_N \mathbf{u}, \mathbf{v} \rangle_{(p,q,r)} = \sigma \langle \mathbf{u}, \mathbf{v} \rangle_{(p,q,r)} \qquad \forall \mathbf{v} \in \mathbb{R}^N \, ,$$

and thanks to the assumption that $\tilde{d}_i > 0$ for all i, this statement is equivalent to the matrix equality $L_N \mathbf{u} = \sigma \mathbf{u}$.

Remark 5.1. The spectra of two graph Laplacians with parameters (p_1, q_1, r_1) and (p_2, q_2, r_2) are identical if

$$p_1 + r_1 = p_2 + r_2 \, , \qquad q_1 = q_2 \, . \tag{5.2}$$

This is true both in the discrete setting for the family L_N defined in (1.2), and in the continuum limit for the family of weighted elliptic operators $\mathcal{L}$ defined in (1.1). Here, we focus on the discrete setting; the argument in the continuum limit is analogous.

To see that this result holds, let L_N^i denote the graph Laplacian defined by (1.2) with parameters (p_i, q_i, r_i), for $i = 1, 2$. The second condition in (5.2) ensures that the weights W_N and degrees D_N are the same for both graph Laplacians and the first condition suffices to make their spectra identical.

Indeed, assume that $(\sigma, \mathbf{u})$ is an eigenpair of L_N^1 in the (p_1, q_1, r_1)-inner product,

$$\langle L_N^1 \mathbf{u} \, , \, \mathbf{u} \rangle_{(p_1,q_1,r_1)} = \sigma \langle \mathbf{u} \, , \, \mathbf{u} \rangle_{(p_1,q_1,r_1)} \, .$$

Defining $\tilde{\mathbf{u}} := D_N^{\frac{1}{2}\left(\frac{p_1 - 1 - r_1}{q_1 - 1} - \frac{p_2 - 1 - r_2}{q_2 - 1} \right)} \mathbf{u} = D_N^{\frac{p_1 - p_2}{q_1 - 1}} \mathbf{u}$, we have

$$\langle \mathbf{u} \, , \, \mathbf{u} \rangle_{(p_1,q_1,r_1)} = \langle \tilde{\mathbf{u}} \, , \, \tilde{\mathbf{u}} \rangle_{(p_2,q_2,r_2)} \, .$$

Now writing $\mathbf{v} := D_N^{-\frac{r_1}{q_1 - 1}} \mathbf{u}$ and $\tilde{\mathbf{v}} := D_N^{-\frac{r_2}{q_2 - 1}} \tilde{\mathbf{u}}$ we realize that $\tilde{\mathbf{v}} = \mathbf{v}$ for parameter choices (p_1, q_1, r_1) and (p_2, q_2, r_2) satisfying (5.2). We conclude that

$$\begin{aligned} \langle L_N^2 \tilde{\mathbf{u}} \, , \, \tilde{\mathbf{u}} \rangle_{(p_2,q_2,r_2)} &= \langle (D_N - W_N) \tilde{\mathbf{v}} \, , \, \tilde{\mathbf{v}} \rangle = \langle (D_N - W_N) \mathbf{v} \, , \, \mathbf{v} \rangle \\ &= \langle L_N^1 \mathbf{u} \, , \, \mathbf{u} \rangle_{(p_1,q_1,r_1)} = \sigma \langle \mathbf{u} \, , \, \mathbf{u} \rangle_{(p_1,q_1,r_1)} \\ &= \sigma \langle \tilde{\mathbf{u}} \, , \, \tilde{\mathbf{u}} \rangle_{(p_2,q_2,r_2)} \end{aligned}$$

and so $(\sigma, \tilde{\mathbf{u}})$ is an eigenpair of L_N^2 in the (p_2, q_2, r_2)-inner product.

Remark 5.2. There are a number of graph-based algorithms which proceed by making a preliminary density estimate via a preliminary weight matrix $\tilde{W}$. In the approach described above, and when $q < 2$, the rescaling of the weights from $\tilde{W}$ to W enlarges affinities between points in regions of low sampling density; this adds robustness to graph-based algorithms, minimizing unwanted impact from outliers in the tails of ϱ. This is sometimes also achieved through a rescaling within η_δ defining

$$W_{ij} = \begin{cases} \eta_\delta\big(\tilde{d}_i^{\,1-q/2}\tilde{d}_j^{\,1-q/2}|x_i - x_j|\big)\,, & i \neq j, \\ 0 & i = j. \end{cases}$$

This idea of variable bandwidth originates in the statistical density estimation literature [30,42] and was introduced to the machine learning community, in the context of graph based data analysis, in [46]. It would be of interest to study limiting continuum operators in this context. Analysis that is relevant to this question is undertaken in [7] where aspects of the work of [14] are generalized to the variable bandwidth setting.

5.2. Convergence of Dirichlet energies

In this subsection, we describe why we expect the spectra of discrete operators L_N to converge to the weighted Laplacian operator $\mathcal{L}$. In simple terms, the limit rests on using the law of large numbers to capture the large data limit $N \to \infty$, in tandem with localizing the weight functions η_δ by sending $\delta \to 0$ so that they behave like Dirac measures. To make these ideas rigorous the two limits need to be carefully linked. Here, however, we simply provide intuition about the role of the two limiting processes, considering first large N and then small δ.

For a vector $\mathbf{u} \in \mathbb{R}^N$, we define the *discrete weighted Dirichlet energy* $E_{N,\delta} : \mathbb{R}^N \to [0, \infty)$,

$$E_{N,\delta}(\mathbf{u}) := \frac{N^{2r-q}}{\delta^2}\langle \mathbf{u}, L_N\mathbf{u}\rangle_{(p,q,r)}.$$

This energy can be extended to functions defined on $\mathcal{Z}$. To achieve this, for $u : \mathcal{Z} \to \mathbb{R}$, we write $u_i := u(x_i)$. Our aim is to study the limiting behavior of the functional $E_{N,\delta}$ as $N \to \infty$ and $\delta \to 0$ on a formal level. In the limit, we obtain the *continuous weighted Dirichlet energy* $E : L^2(\mathcal{Z}, \varrho^{p-r}) \to [0, \infty]$ defined as

$$E(u) := \begin{cases} \frac{1}{2}\langle u, \mathcal{L}u\rangle_{\rho^{p-r}} & \text{if } u \in H^1(\mathcal{Z}, \varrho)\,, \\ \infty & \text{if } u \in L^2(\mathcal{Z}, \varrho^{p-r}) \setminus H^1(\mathcal{Z}, \varrho)\,. \end{cases}$$

Once the convergence of the Dirichlet energies has been established, generalizations of the results in [13,20, 23,45] are possible.

The set of feature vectors X_N induces the empirical measure $\mu_N = \frac{1}{N}\sum_{i=1}^N \delta_{x_i}$, which allows one to define the weighted Hilbert space $L^2(\mathcal{Z}, \mu_N)$ with inner product

$$\langle u, v\rangle_{L^2(\mathcal{Z},\mu_N)} = \int_{\mathcal{Z}} u(x)v(x)\, d\mu_N(x) = \frac{1}{N}\sum_{i=1}^N u(x_i)v(x_i)\,.$$

Since the feature vectors x_i are i.i.d. according to the law ϱ, we have $d\mu_N(x) \rightharpoonup \varrho(x)dx$ as $N \to \infty$. Further, we introduce the functions $\tilde{d}^{N,\delta}, d^{N,\delta} : \mathcal{Z} \to \mathbb{R}$ as follows:

$$\tilde{d}^{N,\delta}(x) := \int_{\mathcal{Z}} \eta_\delta(|x - y|)\, d\mu_N(y)\,,$$

$$d^{N,\delta}(x) := \int_{\mathcal{Z}} \frac{\eta_\delta(|x-y|)}{\left(\tilde{d}^{N,\delta}(x)\right)^{1-q/2}\left(\tilde{d}^{N,\delta}(y)\right)^{1-q/2}}\, d\mu_N(y)\,.$$

Note that

$$\tilde{d}_i = N\tilde{d}^{N,\delta}(x_i)\,, \qquad d_i = N^{q-1}d^{N,\delta}(x_i)\,.$$

For a vector $\mathbf{u} \in \mathbb{R}^N$, we can then rewrite the discrete weighted Dirichlet energy $E_{N,\delta}$ using (5.1) (case $q \neq 1$):

$$\begin{aligned} E_{N,\delta}(\mathbf{u}) &:= \frac{N^{2r-q}}{\delta^2}\langle u, L_N u\rangle_{(p,q,r)} = \frac{N^{2r-q}}{2\delta^2}\sum_{i,j} W_{ij}\left|\frac{u_i}{d_i^{r/(q-1)}} - \frac{u_j}{d_j^{r/(q-1)}}\right|^2 \\ &= \frac{N^{2r-q}}{2\delta^2}\sum_{i,j}\left(\frac{\tilde{W}_{ij}}{\tilde{d}_i^{\,1-q/2}\tilde{d}_j^{\,1-q/2}}\right)\left|\frac{u_i}{d_i^{r/(q-1)}} - \frac{u_j}{d_j^{r/(q-1)}}\right|^2 \\ &= \frac{1}{2\delta^2 N^2}\sum_{i,j}\left(\frac{\eta_\delta(|x_i-x_j|)}{\left(\tilde{d}^{N,\delta}(x_i)\right)^{1-q/2}\left(\tilde{d}^{N,\delta}(x_j)\right)^{1-q/2}}\right) \\ &\qquad \times \left|\frac{u_i}{\left(d^{N,\delta}(x_i)\right)^{r/(q-1)}} - \frac{u_j}{\left(d^{N,\delta}(x_j)\right)^{r/(q-1)}}\right|^2. \end{aligned}$$

This formulation allows us to extend $E_{N,\delta}$ from vectors to functions on $\mathcal{Z}$. More precisely, for $u : \mathcal{Z} \to \mathbb{R}$, we have

$$\begin{aligned} E_{N,\delta}(u) = \frac{1}{2\delta^2}\iint_{\mathcal{Z}\times\mathcal{Z}} &\left(\frac{\eta_\delta(|x-y|)}{\left(\tilde{d}^{N,\delta}(x)\right)^{1-q/2}\left(\tilde{d}^{N,\delta}(y)\right)^{1-q/2}}\right) \\ &\times \left|\frac{u(x)}{\left(d^{N,\delta}(x)\right)^{r/(q-1)}} - \frac{u(y)}{\left(d^{N,\delta}(y)\right)^{r/(q-1)}}\right|^2 d\mu_N(x)d\mu_N(y)\,. \end{aligned} \tag{5.3}$$

Now notice that, by the law of large numbers,

$$\tilde{d}^{N,\delta}(x) \to \tilde{d}^{\delta}(x)\,, \qquad d^{N,\delta}(x) \to d^{\delta}(x) \qquad \text{as } N \to \infty \quad \forall x \in \mathcal{Z}\,,$$

where the functions $\tilde{d}^{\delta}, d^{\delta} : \mathcal{Z} \to \mathbb{R}$ are given by

$$\tilde{d}^{\delta}(x) := \int_{\mathcal{Z}} \eta_\delta(|x-y|)\varrho(y)\,dy\,, \qquad d^{\delta}(x) := \int_{\mathcal{Z}} \frac{\eta_\delta(|x-y|)}{\left(\tilde{d}^{\delta}(x)\right)^{1-q/2}\left(\tilde{d}^{\delta}(y)\right)^{1-q/2}}\varrho(y)\,dy\,.$$

Define

$$s_0 := \int_{\mathcal{Z}} \eta(|x|)\,dx\,, \qquad s_2 := \int_{\mathcal{Z}} |e_1\cdot x|^2\eta(|x|)\,dx\,, \tag{5.4}$$

with e_1 denoting the first unit standard normal vector in $\mathbb{R}^d$. Taking $\delta \to 0$ as a second step, we obtain

$$\tilde{d}^{\delta}(x) \to s_0\varrho(x)\,, \qquad d^{\delta}(x) \to s_0^{q-1}\varrho^{q-1}(x) \qquad \forall x \in \mathcal{Z}\,.$$

Therefore, for smooth enough $u : \mathcal{Z} \to \mathbb{R}$, expression (5.3) allows us to estimate

$$
\begin{aligned}
E_{N,\delta}(u) &= \frac{1}{2\delta^2} \iint\limits_{\mathcal{Z}\times\mathcal{Z}} \left(\frac{\eta_\delta(|x-y|)}{\left(\tilde{d}^{N,\delta}(x)\right)^{1-q/2} \left(\tilde{d}^{N,\delta}(y)\right)^{1-q/2}} \right) \\
&\qquad \times \left| \frac{u(x)}{\left(d^{N,\delta}(x)\right)^{r/(q-1)}} - \frac{u(y)}{\left(d^{N,\delta}(y)\right)^{r/(q-1)}} \right|^2 d\mu_N(x) d\mu_N(y) \\
&\overset{N\gg 1}{\approx} \frac{1}{2\delta^2} \iint\limits_{\mathcal{Z}\times\mathcal{Z}} \left(\frac{\eta_\delta(|x-y|)}{\left(\tilde{d}^{\delta}(x)\right)^{1-q/2} \left(\tilde{d}^{\delta}(y)\right)^{1-q/2}} \right) \\
&\qquad \times \left| \frac{u(x)}{\left(d^{\delta}(x)\right)^{r/(q-1)}} - \frac{u(y)}{\left(d^{\delta}(y)\right)^{r/(q-1)}} \right|^2 \varrho(x)\varrho(y)\, dxdy \\
&\overset{\delta\ll 1}{\approx} \frac{1}{2\delta^2} \iint\limits_{\mathcal{Z}\times\mathcal{Z}} \left(\frac{\eta_\delta(|x-y|)}{\left(\tilde{d}^{\delta}(x)\right)^{1-q/2} \left(\tilde{d}^{\delta}(y)\right)^{1-q/2}} \right) \\
&\qquad \times \left| \nabla \left(\frac{u(x)}{\left(d^{\delta}(x)\right)^{r/(q-1)}} \right) \cdot (x-y) \right|^2 \varrho(x)\varrho(y)\, dxdy \\
&\overset{\delta\ll 1}{\approx} \frac{1}{2} \frac{s_2}{s_0^{2r+2-q}} \int\limits_{\mathcal{Z}} \frac{1}{\varrho(x)^{2-q}} \left| \nabla \left(\frac{u(x)}{\varrho(x)^r} \right) \right|^2 \varrho(x)^2\, dx \\
&= \frac{1}{2} \frac{s_2}{s_0^{2r+2-q}} \int\limits_{\mathcal{Z}} \left| \nabla \left(\frac{u(x)}{\varrho(x)^r} \right) \right|^2 \varrho(x)^q\, dx = \frac{s_2}{s_0^{2r+2-q}} E(u)\,.
\end{aligned}
$$

This is the desired result. To develop a theorem based on these calculations requires taking $N \to \infty$ concurrently with $\delta \to 0$, and may be done in the framework of [13,20,45].

Remark 5.3. While the above arguments primarily concern proximity graphs; the method of proof in [13] is more general and can be applied to k-NN graphs as well. However, the resulting limiting process gives a different relationship between the continuum operator $\mathcal{L}$ with a certain choice of (p, q, r) and the correct normalization of the discrete Laplacian L_N.

Remark 5.4. Not all graph Laplacian normalizations lead to differential operators of the type (1.1) in the large data limit, and this is the motivation for introducing the parameters (p, q, r) as graph Laplacian weightings of type (1.2). For example, the operator $D_N^{-s}(D_N - W_N)D_N^{-t}$ with $q = 1$ does *not* correspond to a continuum operator of type (1.1) in the same large data limit, for any choice of $s, t \in \mathbb{R} \setminus \{0\}$.

5.3. Discrete vs continuum eigenproblems

In this subsection, we make explicit the relationship between the discrete and continuum eigenproblems and highlight the correct scaling needed in the discrete setting for the spectra to converge. Let (σ, φ) be an eigenpair of $\mathcal{L}$ and take a test function $\phi \in H^1(\mathcal{Z}, \varrho)$. The arguments in Subsection 5.2 show that for vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^N$ where $u_i = \varphi(x_i), v_i = \phi(x_i)$, we have

$$
\frac{N^{2r-q}}{\delta^2} \langle L_N \mathbf{u}, \mathbf{v} \rangle_{(p,q,r)} \overset{N\gg 1,\delta\ll 1}{\approx} \frac{s_2}{2s_0^{2r+2-q}} \langle \mathcal{L}\varphi, \phi \rangle_{\varrho^{p-r}}\,.
$$

With a similar argument, one can identify the continuum analogue of the weighted inner product $\langle \mathbf{u}, \mathbf{v} \rangle_{(p,q,r)}$ by rewriting it in terms of φ and ϕ:

$$N^{r-p}\langle \mathbf{u},\mathbf{v}\rangle_{(p,q,r)} = N^{r-p}\langle \mathbf{u}, D^{\frac{p-r-1}{q-1}}\mathbf{v}\rangle = N^{r-p}\sum_{i=1}^{N}\mathbf{u}_i\mathbf{v}_i d_i^{\frac{p-r-1}{q-1}}$$
$$= N^{r-p}\sum_{i=1}^{N}\varphi(x_i)\phi(x_i)N^{p-r-1}\left(d^{N,\delta}(x_i)\right)^{\frac{p-r-1}{q-1}}$$
$$= \int_{\mathcal{Z}}\varphi(x)\phi(x)\left(d^{N,\delta}(x)\right)^{\frac{p-r-1}{q-1}}\,d\mu_N(x)\,.$$

Recall from Subsection 5.2 that by the law of large numbers, $d^{N,\delta}(x)\to d^{\delta}(x)$ as $N\to\infty$, and taking $\delta\to 0$ as a next step, we obtain $d^{\delta}(x)\to s_0^{q-1}\varrho^{q-1}(x)$. Therefore,

$$N^{r-p}\langle \mathbf{u},\mathbf{v}\rangle_{(p,q,r)} \overset{N\gg 1}{\approx} \int_{\mathcal{Z}}\varphi(x)\phi(x)\left(d^{\delta}(x)\right)^{\frac{p-r-1}{q-1}}\varrho(x)\,dx$$
$$\overset{\delta\ll 1}{\approx} s_0^{p-r-1}\int_{\mathcal{Z}}\varphi(x)\phi(x)\varrho(x)^{p-r}\,dx\,.$$

In other words, for an eigenpair $(\tilde{\sigma}_{N,\delta},\mathbf{u})$ of the weighted graph Laplacian matrix L_N solving

$$\langle L_N\mathbf{u},\mathbf{v}\rangle_{(p,q,r)} = \tilde{\sigma}_{N,\delta}\langle \mathbf{u},\mathbf{v}\rangle_{(p,q,r)}\,,\qquad \forall \mathbf{v}\in\mathbb{R}^N \tag{5.5}$$

we expect that

$$\frac{2s_0^{p+r-q+1}}{\delta^2 N^{q-p-r}s_2}\tilde{\sigma}_{N,\delta}\to\sigma,\ \text{as } N\to\infty,\delta\to 0,$$

where σ is an eigenvalue of $\mathcal{L}$,

$$\langle \mathcal{L}\varphi,\phi\rangle_{\varrho^{p-r}} = \sigma\langle\varphi,\phi\rangle_{\varrho^{p-r}}\,.$$

These considerations imply that the discrete eigenvalues of L_N need to be scaled appropriately in order to converge to the eigenvalues of $\mathcal{L}$.

Remark 5.5. It is shown in the papers [13,23,40,20] that for the parameter choices $(p,q,r)=(1,2,0)$ and $(3/2,2,1/2)$ and in the limit as $N\to\infty$ and $\delta:=\delta_N\to 0$ at an appropriate rate with N, the discrete operators L_N converge to $\mathcal{L}$ on $\mathcal{Z}$. Those papers analyze the convergence of the Dirichlet forms associated with L_N (defined with respect to real-valued functions on the vertices X_N) to those associated with $\mathcal{L}$ (defined with respect to real-valued functions on $\mathcal{Z}$). In particular, [23,40] use Γ-convergence arguments based on the TL^2 topology to prove convergence. This topology may be used to study Γ-limits of other non-quadratic functionals defined with respect to real-valued functions on the graph – see [18], for example. A similar methodology can be applied to show convergence of L_N to $\mathcal{L}$ for any choice of parameters $(p,q,r)\in\mathbb{R}^3$. However, the Γ-convergence framework does not result in rates of convergence for eigenvalues and eigenvectors of L_N making it difficult to extend continuum analyses, such as our Main Result 1.1, to practical discrete problems. In contrast, the more recent articles [13,20,45] take a more direct approach to proving the convergence of L_N to $\mathcal{L}$ and obtain rates. The rigorous study of this limiting procedure for the general (p,q,r) family of operators is the subject of future research.

Remark 5.6. The fact that the scaling factor in front of $\tilde{\sigma}_{N,\delta}$ has a dependence on N^{p+r-q} once again highlights the special role of the balanced case $q=p+r$.

5.4. Numerical experiments in the discrete setting

In this subsection we present a set of numerical experiments concerning the spectrum of discrete graph Laplacian matrices L_N. Our goal here is twofold: 1) we support the theoretical findings in Subsection 5.3 by showing that as $N \to \infty$ and $\delta \to 0$, the eigenvalues of L_N converge to those of $\mathcal{L}$ after appropriate scaling by N, δ and for different choices of (p, q, r); 2) we show that the continuum spectral analysis of Section 3 manifests for the setting of finitely many samples as well. In particular, we show that a uniform spectral gap for L_N exists when $q = p + r$ but disappears when $q > p + r$.

In what follows, we display two numerical examples: choosing ϱ to be (i) a piecewise constant mixture model, and (ii) a mixture model with exponential components.

5.4.1. A piecewise constant mixture

For the set-up of our numerical experiments, we choose $\mathcal{Z} = (0, 1) \times (0, 1) \subset \mathbb{R}^2$ and define the sequence of densities

$$\varrho_\epsilon(t) = \begin{cases} \epsilon, & t_1 \in (0.2, 0.8), \\ 2.5 - 1.5\epsilon, & t_1 \in [0, 0.2] \cup [0.8, 1], \end{cases} \qquad \forall t = (t_1, t_2)^T \in \mathcal{Z}. \tag{5.6}$$

Thus as $\epsilon \to 0$ the density ϱ_ϵ vanishes inside a strip in the middle of $\mathcal{Z}$ while the rest of the probability mass is split equally between two rectangles to the sides of $\mathcal{Z}$. Note that ϱ_ϵ is discontinuous by definition and so it does not satisfy all of our assumptions from Subsection 2.2. For fixed values of ϵ we sample vertices $\{x_i\}_{i=1}^N$ i.i.d. with respect to ϱ_ϵ and construct a weighted graph $\tilde{W}$ with entries $\tilde{W}_{ij} = \eta_\delta(|x_i - x_j|)$ as in Section 5.1. As for the kernel η_δ we choose

$$\eta_\delta(t) = \frac{1}{\pi\delta^2} \mathbf{1}_{[0,\delta)}(t), \qquad \forall t \in [0, +\infty), \tag{5.7}$$

for which we can easily compute the normalizing constants defined in (5.4) to be $s_0 = 1$ and $s_2 = 1/4$. We can then proceed to define the graph Laplacian matrices L_N as outlined in Subsection 5.1 for different choices of $(p, q, r) \in \mathbb{R}^3$. It remains to choose a relationship between δ, N to ensure convergence of the spectrum of L_N as $N \to \infty$ and $\delta \to 0$. Following [13] we choose

$$\delta = \left(\frac{\log(N)}{N} \right)^{1/3}. \tag{5.8}$$

Although this choice is not justified theoretically at this point we find that it is sufficient numerically to achieve convergence of the eigenvalues.

In Fig. 5.1 we plot the first four non-trivial eigenvalues $\sigma_{N,\delta}$ of L_N as a function of N for $\epsilon = 2^{-3}$ and various choices of (p, q, r) in both balanced and unbalanced cases. Each reported eigenvalue was averaged over twenty redraws of the vertices. We clearly observe that as $N \to \infty$ the eigenvalues converge although the larger eigenvalues appear to converge more slowly. In Fig. 5.2 we plot the relative errors between the discrete eigenvalues $\sigma_{N,\delta}$ and the continuum eigenvalues σ computed using our finite element solver from Section 4 with the density ϱ_ϵ as in (5.6). We observe that in both balanced and unbalanced regimes the discrete eigenvalues converge to their continuum counterparts although the convergence plateau's in the $q > p + r$ case at around $1e-3$ most likely due to numerical errors. We observed that convergence improves for larger values of ϵ.

For our next set of experiments we consider the behavior of the discrete eigenvalues $\sigma_{N,\delta}$ as ϵ vanishes. We fix $N = 2^{13}$ and choose $\epsilon = 2^{-2}, \ldots, 2^{-4}$. Here we redraw the vertices five times and average the computed eigenvalues over these five trials. Fig. 5.3 shows results that are analogous to Fig. 1.1(b,c). We observe that

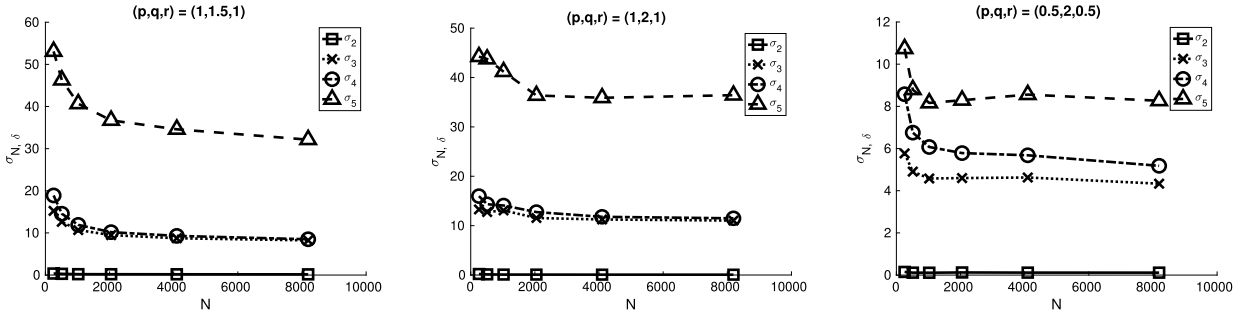


Fig. 5.1. Convergence of the first four non-trivial discrete eigenvalues $\sigma_{N,\delta}$ as a function of N for different values of (p, q, r) and $\epsilon = 2^{-3}$ with vertices distributed according to (5.6).

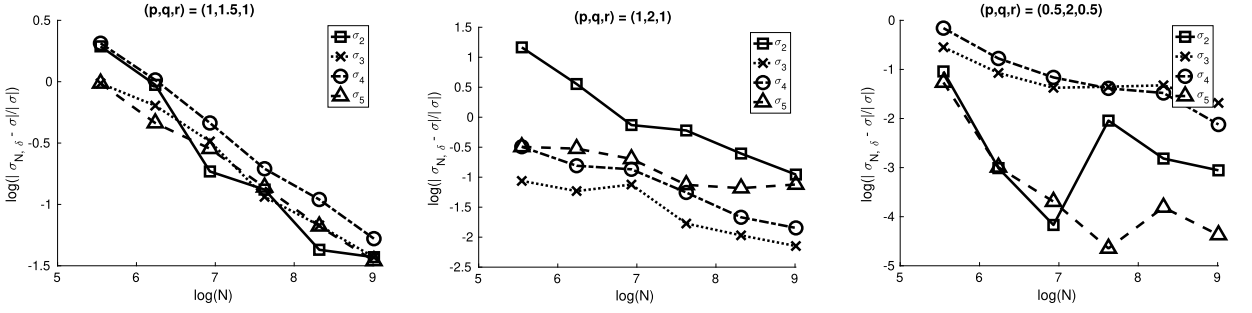


Fig. 5.2. Relative error between the first four non-trivial discrete eigenvalues $\sigma_{N,\delta}$ and the continuum eigenvalues σ as a function of N for different values of (p, q, r) and $\epsilon = 2^{-3}$ with vertices distributed according to (5.6).

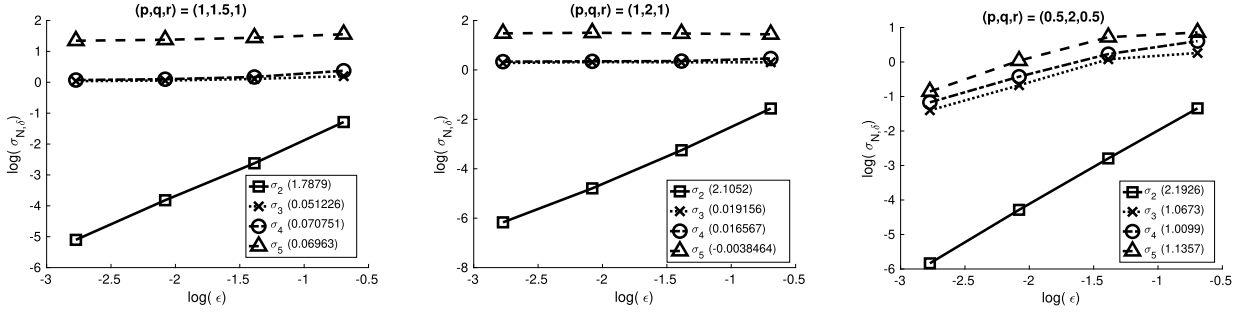


Fig. 5.3. The dependence of the non-trivial discrete eigenvalues $\sigma_{N,\delta}$ as a function of ϵ for different values of (p, q, r) and $N = 2^{13}$ with vertices drawn from (5.6). The reported values within the brackets in the legend are the slopes of a linear fit to the last three data points indicating the rate at which the corresponding eigenvalues vanish with ϵ.

in the balanced case where $q = p + r$ the second eigenvalue vanishes like ϵ^q while the larger eigenvalues remain bounded away from zero as predicted by Theorem 3.2 and confirmed by our numerical experiments in Subsection 4.1. The case where $q > p+r$ also agrees with Theorem 3.2 as well as our continuum numerical experiments in Subsection 4.2 and in turn with the first component of Conjecture 1.2, as we observe that the second eigenvalue vanishes like ϵ^q while the third eigenvalue vanishes like ϵ^{p+r}. Finally, in the $q < p+r$ case we observe a similar behavior to the balanced case where a uniform spectral gap manifests while the second eigenvalue appears to vanish at a rate that is slightly faster than ϵ^q which we attribute to numerical errors. Hence, our discrete experiments are once again in line with continuum experiments from Subsection 4.3 and further support the first component of Conjecture 1.2.

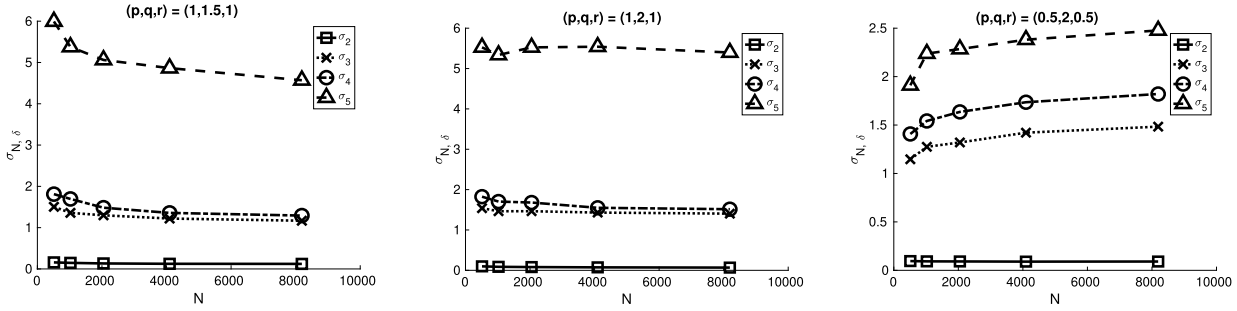


Fig. 5.4. Convergence of the first four scaled discrete eigenvalues $\sigma_{N,\delta}$ as a function of N for different values of (p, q, r) and $\omega = 1.9^{-3}$ with vertices distributed according to (1.4).

5.4.2. An exponential mixture

Here we give full details of the numerical experiments presented in Example 1.4 in Subsection 1.3. We use the same kernel η_δ and parameterization of $\delta(N)$ as in (5.7) and (5.8) respectively. Similarly we choose $\mathcal{Z} = (0,1) \times (0,1) \subset \mathbb{R}^2$ but sample the vertices of the graph from the density ϱ_ω as in (1.4), see Fig. 1.2(a) for a plot of ϱ_ω with $\omega = 1/4$.

In Fig. 1.2(b,c,d) we fix $N = 2^{13}$ and choose $\omega = (1.9)^{-5}, \dots (1.9)^{-8}$. Each data point is obtained by averaging the first four eigenvalues of L_N over five trials where the vertices of the graph are redrawn from ϱ_ω. As we already discussed in Example 1.4 our numerical results indicate that the relationship between p, q and r has a major impact on the gap between the second and third eigenvalues of L_N. In particular, when $q \leq p + r$ a uniform gap is observed while when $q > p + r$ only a ratio gap manifests. We also note that the rate of decay of the second and third eigenvalues as a function of ω in Fig. 1.2(b,c,d) is different from the rates we obtained as a function of the perturbation parameter ϵ since ϱ_ω vanishes exponentially fast in the middle of the domain which violates our assumption that the density satisfies $\varrho = K\epsilon$ away from the clusters. Finally, in Fig. 5.4 we plot the first four non-trivial eigenvalues $\sigma_{N,\delta}$ of L_N for $\epsilon = 1.9^{-6}$ and for different values of N. Analogously to Fig. 5.1 our results show that the first few eigenvalues of L_N converge as $N \to \infty$ for the exponential mixture model as well.

6. Spectral analysis: proofs

In this section we present proofs of the theorems in Section 3. The essential analytical tools in our spectral analysis are the min-max and max-min formulas from Appendix C, together with a new weighted version of Cheeger's inequality given in Appendix D. We adopt the same organizational format as Section 3. In Subsection 6.1 we discuss the perfectly clustered case, and then consider small perturbations of this setting, the nearly clustered case, in Subsection 6.2. Theorem 3.2 is proved in Subsections 6.2.1, 6.2.2 and while the proof of Theorem 3.4 is outlined in Subsection 6.3.

6.1. Proof of Theorem 3.1

As detailed in the discussion following Theorem 3.1 it only remains to characterize the third eigenvalue of $\mathcal{L}_0$.

Proposition 6.1. *Suppose Assumptions 2.4 and 2.5 are satisfied and the $\mathcal{L}_0$ spectral gap condition holds on the clusters $\mathcal{Z}^\pm$ with optimal constants $\Lambda_0^\pm := \Lambda_0(\mathcal{Z}^\pm) > 0$ separately. Then $\sigma_{3,0} \geq \min\{\Lambda_0^+, \Lambda_0^-\} > 0$.*

Proof. Note that Assumption 2.5(e) ensures that $\varphi_{2,0} = |\mathcal{Z}'|^{1/2}_{\varrho_0^{p-r}} \varrho^r (\mathbf{1}_{\mathcal{Z}^+} - \mathbf{1}_{\mathcal{Z}^-})$ belongs to $V^0(\mathcal{Z}', \varrho_0)$. Let $u \in V^1(\mathcal{Z}', \varrho_0)$ so that $u \perp \text{span}\{\varphi_{1,0}, \varphi_{2,0}\}$ in $L^2(\mathcal{Z}', \varrho_0^{p-r})$. A direct calculation shows that this means

the restrictions $u|_{\mathcal{Z}^\pm}$ of u to the clusters $\mathcal{Z}^\pm$ are orthogonal (with respect to the $L^2(\mathcal{Z}^\pm, \varrho_0^{p-r}|_{\mathcal{Z}^\pm})$ inner products) to the restrictions $\varrho_0^r|_{\mathcal{Z}^\pm}$ of ϱ_0^r and belong to $V^0(\mathcal{Z}^\pm, \varrho_0|_{\mathcal{Z}^\pm})$. Thus following the $\mathcal{L}_0$ spectral gap assumption, see Definition 2.10, $u|_{\mathcal{Z}^\pm}$ satisfy Poincaré inequalities of the form (2.17) on $\mathcal{Z}^\pm$ with optimal constants $\Lambda_0^\pm$. Hence

$$\begin{aligned}\int_{\mathcal{Z}'} \left|\nabla\left(\frac{u}{\varrho_0^r}\right)\right|^2 \varrho_0^q dx &= \int_{\mathcal{Z}^+} \left|\nabla\left(\frac{u}{\varrho_0^r}\right)\right|^2 \varrho_0^q dx + \int_{\mathcal{Z}^-} \left|\nabla\left(\frac{u}{\varrho_0^r}\right)\right|^2 \varrho_0^q dx \\ &\geq \min\{\Lambda_0^+, \Lambda_0^-\} \left(\int_{\mathcal{Z}^+} \left|\frac{u}{\varrho_0^r}\right|^2 \varrho_0^{p+r} dx + \int_{\mathcal{Z}^-} \left|\frac{u}{\varrho_0^r}\right|^2 \varrho_0^{p+r} dx \right) \\ &= \min\{\Lambda_0^+, \Lambda_0^-\} \int_{\mathcal{Z}'} \left|\frac{u}{\varrho_0^r}\right|^2 \varrho_0^{p+r} dx.\end{aligned}$$

The result now follows from the max-min formula (C.2) in Theorem C.1. □

Remark 6.2. If Assumption 2.5(e) is dropped then the two terms in the definition of $\varphi_{2,0}$ need to be weighted by appropriate constants to ensure $\int_{\mathcal{Z}'} \varphi_{2,0}(x)\varrho_0^p(x)dx = 0$ so that $\varphi_{2,0} \in V^0(\mathcal{Z}', \varrho_0)$.

6.2. Proof of Theorem 3.2

We now turn our attention to the densities ϱ_ϵ that have full support on $\bar{\mathcal{Z}}$, but concentrate around $\mathcal{Z}'$ as ϵ decreases. Throughout this section, we routinely assume that Assumptions 2.4, 2.5 and 2.6 are satisfied by the domains $\mathcal{Z}, \mathcal{Z}'$ and densities ϱ_0 and ϱ_ϵ. Throughout, the constants Ξ and Ξ_j for any j are arbitrary and can change from one line to the next.

We start by constructing an approximation for $\varphi_{2,\epsilon}$ (the second eigenfunction of $\mathcal{L}_\epsilon$) that is used throughout this section. Fix $\epsilon > 0$ and define the sets $\mathcal{Z}_{\epsilon_1}^\pm$ and $\mathcal{Z}_\epsilon^\pm$ as in (2.7), where $\epsilon_1 = \epsilon + \epsilon^\beta$ with a parameter $0 < \beta < 1$. We choose ϵ small enough so that $\mathcal{Z}_{\epsilon_1}^+$ and $\mathcal{Z}_{\epsilon_1}^-$ are disjoint. Consider functions $\xi_\epsilon^\pm \in C^\infty(\bar{\mathcal{Z}})$ that satisfy

$$\begin{aligned} &\xi_\epsilon^\pm(x) = 1, && x \in \mathcal{Z}_\epsilon^\pm, \\ &0 < \xi_\epsilon^\pm(x) < 1, \quad |\nabla \xi_\epsilon^\pm(x)| \leq \vartheta \epsilon^{-\beta}, && x \in \mathcal{Z}_{\epsilon_1}^\pm \setminus \mathcal{Z}_\epsilon^\pm, \\ &\xi_\epsilon^\pm(x) = 0, && x \in \mathcal{Z} \setminus \mathcal{Z}_{\epsilon_1}^\pm, \end{aligned}$$

for some constant $\vartheta > 0$ independent of β. The $\xi_\epsilon^\pm$ are smooth extensions of the set functions $\mathbf{1}_{\mathcal{Z}_\epsilon^\pm}$. They can be constructed by convolution with the standard mollifier g_ϵ in the same manner in which ϱ_ϵ was constructed in (2.11) (also see [32, Thm. 3.6]). Now define the functions $\chi_\epsilon^\pm \in C^\infty(\bar{\mathcal{Z}})$ by renormalizing $\xi_\epsilon^\pm$ in $L^2(\bar{\mathcal{Z}}, \varrho_\epsilon^{p-r})$,

$$\chi_\epsilon^+ := b_\epsilon^+ \xi_\epsilon^+, \qquad \chi_\epsilon^- := b_\epsilon^- \xi_\epsilon^-, \tag{6.1}$$

where the coefficients $b_\epsilon^\pm \in \mathbb{R}_+$ are chosen to satisfy

$$\begin{aligned} \int_{\mathcal{Z}_{\epsilon_1}^+} \varrho_\epsilon^{p+r} \chi_\epsilon^+ \, dx &= \int_{\mathcal{Z}_{\epsilon_1}^-} \varrho_\epsilon^{p+r} \chi_\epsilon^- \, dx, \\ b_\epsilon^+ + b_\epsilon^- &= 2. \end{aligned} \tag{6.2}$$

The first condition ensures that $\varrho_\epsilon^r(\chi_\epsilon^+ - \chi_\epsilon^-) \in V^0(\mathcal{Z}, \varrho_\epsilon)$, whereas the second condition is not necessary and chosen for closure and convenience in the calculations that follow. For a schematic depiction of these constructions, see Fig. 4.1.

We define the following ansatz as an approximation to $\varphi_{2,\epsilon}$

$$\varphi_{F,\epsilon}(x) = \frac{\varrho_\epsilon^r(x)\,[\chi_\epsilon^+(x) - \chi_\epsilon^-(x)]}{\left\|\varrho_\epsilon^r(x)\,[\chi_\epsilon^+(x) - \chi_\epsilon^-(x)]\right\|_{L^2(\mathcal{Z},\varrho_\epsilon^{p-r})}}. \tag{6.3}$$

Observe that $\varphi_{F,\epsilon}$ is simply a smooth approximation to the zero extension of $\varphi_{2,0}$ to all of $\mathcal{Z}$ by an element of $V^0(\mathcal{Z}, \varrho_\epsilon)$. The dependence on $\beta > 0$ has been omitted in $\varphi_{F,\epsilon}$ for notational convenience. One should choose β large enough in order for the set $\mathcal{Z}'_{\epsilon_1}$ to be close to $\mathcal{Z}'_\epsilon$. However, this has to be balanced with small enough β such that the derivatives $\nabla\chi_\epsilon^\pm$ are allowed to be steep enough for $\varphi_{F,\epsilon}$ to be a good approximation of the Fiedler vector $\varphi_{2,0}$. The following lemma is useful throughout the rest of this section.

Lemma 6.3. *Suppose that $p + r \geq 0$ and that Assumptions 2.4, 2.5 and 2.6 hold and let $b_\epsilon^\pm$ be as in (6.1). Suppose $\epsilon \in (0, \epsilon_0)$ for a sufficiently small $\epsilon_0 > 0$. Then there exists a constant $\Xi > 0$, independent of ϵ so that*

$$|b_\epsilon^\pm - 1| \leq \Xi\epsilon^{\min\{1,p+r\}}.$$

Proof. Consider the ratio

$$\Xi_\epsilon := \frac{\int_{\mathcal{Z}_{\epsilon_1}^+} \varrho_\epsilon^{p+r}\xi_\epsilon^+\,dx}{\int_{\mathcal{Z}_{\epsilon_1}^-} \varrho_\epsilon^{p+r}\xi_\epsilon^-\,dx}.$$

Solving (6.2) for $b_\epsilon^\pm$ we obtain $b_\epsilon^+ = \frac{2}{1+\Xi_\epsilon}$ and $b_\epsilon^- = \frac{2\Xi_\epsilon}{1+\Xi_\epsilon}$. Thus if we can show that

$$|\Xi_\epsilon - 1| \leq \Xi_1\epsilon^{\min\{1,p+r\}}, \tag{6.4}$$

then $|\Xi_\epsilon + 1| = |(-2) - (\Xi_\epsilon - 1)| \geq 2 - |\Xi_\epsilon - 1|$, and so

$$|b_\epsilon^\pm - 1| = \frac{|\Xi_\epsilon - 1|}{|\Xi_\epsilon + 1|} \leq \frac{|\Xi_\epsilon - 1|}{2 - |\Xi_\epsilon - 1|} \leq \frac{\Xi_1\epsilon^{\min\{1,p+r\}}}{2 - \Xi_1\epsilon^{\min\{1,p+r\}}} \leq \Xi\epsilon^{\min\{1,p+r\}},$$

for some $\Xi > 0$, which concludes the proof of the lemma. It remains to show (6.4). Following Assumption 2.6(c, d), for sufficiently small ϵ,

$$\begin{aligned}\Xi_\epsilon &\leq \frac{\int_{\mathcal{Z}_\epsilon^+} \varrho_\epsilon^{p+r}dx + K_2^{p+r}\epsilon^{p+r}|\mathcal{Z}_{\epsilon_1}^+ \setminus \mathcal{Z}^+|}{\int_{\mathcal{Z}^-} \varrho_\epsilon^{p+r}dx} \\ &\leq \frac{\int_{\mathcal{Z}^+}(\varrho_0 + K_1\epsilon)^{p+r}dx + \int_{\mathcal{Z}_\epsilon^+\setminus\mathcal{Z}^+} \varrho_\epsilon^{p+r}dx + K_2^{p+r}\epsilon^{p+r}|\mathcal{Z}^+_{(\epsilon_0+\epsilon_0^\beta)} \setminus \mathcal{Z}^+|}{\int_{\mathcal{Z}^-}(\varrho_0 - K_1\epsilon)^{p+r}dx}.\end{aligned}$$

Note that $\int_{\mathcal{Z}_\epsilon^+\setminus\mathcal{Z}^+} \varrho_\epsilon^{p+r}dx \leq (\varrho_{\epsilon_0}^+)^{p+r}|\mathcal{Z}_\epsilon^+ \setminus \mathcal{Z}^+| \leq (\varrho_{\epsilon_0}^+)^{p+r}\theta\epsilon|\partial\mathcal{Z}^+|$ following the remark after (2.9) and using (2.8). For $0 \leq p + r \leq 1$, we use the inequality $(a+b)^{p+q} \leq (a^{p+r} + b^{p+r})$ for any $a, b \geq 0$, and obtain

$$\Xi_\epsilon \leq \frac{\int_{\mathcal{Z}^+} \varrho_0^{p+r}dx + \Xi_2\epsilon^{p+r}}{\int_{\mathcal{Z}^-} \varrho_0^{p+r}dx - \Xi_3\epsilon^{p+r}}.$$

Thanks to Assumption 2.5(e), $\int_{\mathcal{Z}^+} \varrho_0^{p+r}\,dx = \int_{\mathcal{Z}^-} \varrho_0^{p+r}\,dx$, and so Taylor expanding in $\Xi_3\epsilon^{p+r}$ yields

$$\Xi_\epsilon \leq 1 + \left(\frac{\Xi_2}{\int_{\mathcal{Z}^-} \varrho_0^{p+r} dx} + \Xi_3 \right) \epsilon^{p+r} + \mathcal{O}\left(\epsilon^{2(p+r)}\right) \leq \Xi_1 \epsilon^{p+r}$$

since ϱ_0 is bounded below uniformly on $\mathcal{Z}^-$ by Assumption 2.5(c).

If $p + r > 1$ on the other hand, we simply Taylor expand $(\varrho_0 + K_1\epsilon)^{p+r}$ and $(\varrho_0 - K_1\epsilon)^{p+r}$ directly, and obtain

$$\Xi_\epsilon \leq \frac{\int_{\mathcal{Z}^+} \varrho_0^{p+r} dx + \Xi_2 \epsilon}{\int_{\mathcal{Z}^-} \varrho_0^{p+r} dx - \Xi_3 \epsilon} \leq 1 + \Xi_1 \epsilon \,,$$

again using the uniform upper and lower bounds for ϱ_0 on $\mathcal{Z}^\pm$. The lower bound on $\pm(\Xi_\epsilon - 1)$ follows in a similar manner. □

6.2.1. Proof of Theorem 3.2(ii) (second eigenvalue of $\mathcal{L}_\epsilon$)

Proposition 6.4 *(Second eigenvalue of $\mathcal{L}_\epsilon$). Let $(p, q, r) \in \mathbb{R}^3$ satisfying $p + r > 0$ and $q > 0$, and suppose Assumptions 2.4, 2.5, and 2.6 hold. Then $\exists\, \epsilon_0 > 0$ so that $\forall (\epsilon, \beta) \in (0, \epsilon_0) \times (0, 1)$,*

$$0 \leq \sigma_{2,\epsilon} \leq \Xi \epsilon^{q-\beta},$$

where $\Xi > 0$ is a uniform constant independent of ϵ.

Proof. Fix an $\epsilon_0 > 0$ and let $\epsilon \in (0, \epsilon_0]$. Recall that $\varphi_{F,\epsilon} \in V^0(\mathcal{Z}, \varrho_\epsilon)$ thanks to (6.2) and is normalized with respect to the $L^2(\mathcal{Z}, \varrho_\epsilon^{p-r})$ norm. Now consider the Rayleigh quotient

$$\mathcal{R}_\epsilon(u) := \frac{\int_{\mathcal{Z}} \left| \nabla \left(\frac{u}{\varrho_\epsilon^r} \right) \right|^2 \varrho_\epsilon^q dx}{\int_{\mathcal{Z}} \left| \frac{u}{\varrho_\epsilon^r} \right|^2 \varrho_\epsilon^{p+r} dx},$$

for functions $u \in \operatorname{span}\{\varphi_{1,\epsilon}, \varphi_{F,\epsilon}\}$. Note that $\mathcal{R}_\epsilon(\varphi_{1,\epsilon}) = 0$, and so $\mathcal{R}_\epsilon(u) \leq \mathcal{R}_\epsilon(\varphi_{F,\epsilon})$. Therefore, we can consider $u \in V^1(\mathcal{Z}, \varrho_\epsilon)$. Following the min-max principle (C.1) we simply need to bound $\mathcal{R}_\epsilon(\varphi_{F,\epsilon})$ to find an upper bound for $\sigma_{2,\epsilon}$. Let

$$\Xi_0 = \inf_{\epsilon \in (0,\epsilon_0]} \| \varrho_\epsilon^r [\chi_\epsilon^+ - \chi_\epsilon^-] \|_{L^2(\mathcal{Z}, \varrho_\epsilon^{p-r})}$$

and note that provided ϵ_0 is small enough, $\Xi_0 > 0$ following Lemma 6.3, the fact that $\chi_\epsilon^\pm$ have disjoint supports, $p + r \geq 0$ and using that ϱ_ϵ is bounded above and below on $\mathcal{Z}'$ by (2.9) (see also Lemma 6.10 in Section 6.3 for a more detailed argument). Using $0 < b_\epsilon^\pm < 2$ and Assumption 2.6(d), we have

$$\begin{aligned} \mathcal{R}_\epsilon(\varphi_{F,\epsilon}) &\leq \frac{4}{\Xi_0} \int_{\mathcal{Z}} \left| \nabla \left(\xi_\epsilon^+ - \xi_\epsilon^- \right) \right|^2 \varrho_\epsilon^q dx \\ &= \frac{4}{\Xi_0} \int_{\mathcal{Z}'_{\epsilon_1} \setminus \mathcal{Z}_\epsilon{}'} \left| \nabla \left(\xi_\epsilon^+ - \xi_\epsilon^- \right) \right|^2 \varrho_\epsilon^q dx \\ &\leq \frac{16 K_2^q \vartheta^2}{\Xi_0} |\mathcal{Z}'_{\epsilon_1} \setminus \mathcal{Z}_\epsilon{}'| \epsilon^{q-2\beta} \leq \Xi \epsilon^{q-\beta}, \end{aligned} \tag{6.5}$$

since $|\mathcal{Z}'_{\epsilon_1} \setminus \mathcal{Z}_\epsilon{}'| \leq |\mathcal{Z}'_{\epsilon_1} \setminus \mathcal{Z}'| \leq \theta(\epsilon + \epsilon^\beta)|\partial \mathcal{Z}'| \leq \Xi_1 \epsilon^\beta$ by (2.8) and since $\beta < 1$. It now follows from (C.1) that $\sigma_{2,\epsilon} \leq \Xi \epsilon^{q-\beta}$. □

6.2.2. Proof of Theorem 3.2 (third eigenvalue of $\mathcal{L}_\epsilon$)

We prove the bounds on the third eigenvalue of $\mathcal{L}_\epsilon$ in a series of propositions and corollaries. In particular, part (iii) of Theorem 3.2 follows by combining Propositions 6.7 and 6.8 below. We start with a general result that ties the existence of a $\mathcal{L}_\epsilon$ spectral gap on $\mathcal{Z}$ to spectral gaps on subsets of $\mathcal{Z}$.

Proposition 6.5. *Let $(p,q,r) \in \mathbb{R}^3$ satisfying $p+r>0$ and $q>0$, and suppose Assumptions 2.4, 2.5, and 2.6 hold. Let*

$$\Lambda(\epsilon) := \min\{\Lambda_\epsilon(\mathcal{Z}^+_{\epsilon_0}), \Lambda_\epsilon(\mathcal{Z}\setminus\mathcal{Z}^+_{\epsilon_0})\} \geq 0\,, \tag{6.6}$$

for some $\epsilon_0>0$. Then there exist constants $s,t,\Xi_1,\Xi_2,\Xi_3>0$ independent of ϵ so that $\forall \epsilon\in(0,\epsilon_0)$,

$$\sigma_{3,\epsilon} \geq \min\left\{\frac{\Lambda(\epsilon)(1-\Xi_1\epsilon^t)}{1+\Xi_2\Lambda(\epsilon)\epsilon^s}, \Lambda(\epsilon)\left(1-\Xi_3\epsilon^{\min\{t,s\}}\right)\right\}.$$

Proof. Note that it is possible that $\Lambda(\epsilon)=0$ if the spectral gap condition in Definition 2.11 is not satisfied in $\mathcal{Z}^+_{\epsilon_0}$ or $\mathcal{Z}\setminus\mathcal{Z}^+_{\epsilon_0}$. If this happens for some $\epsilon\in(0,\epsilon_0)$, then the proposition trivially holds. Therefore, we assume from now on that $\Lambda(\epsilon)>0$ for all $\epsilon\in(0,\epsilon_0)$.

Let $u\in V^1(\mathcal{Z},\varrho_\epsilon)$ and $u\perp\varphi_{F,\epsilon}$ with respect to the $\langle\cdot,\cdot\rangle_V$-inner product. Without loss of generality assume $\|u\|_{L^2(\mathcal{Z},\varrho_\epsilon^{p-r})}=1$. We will prove the desired lower bound for $\mathcal{R}_\epsilon(u)$ and use the max-min principle (Theorem C.1) to infer the lower bound of $\sigma_{3,\epsilon}$.

By definition of Λ_ϵ we have

$$\begin{aligned}\int_{\mathcal{Z}}\left|\nabla\left(\frac{u}{\varrho_\epsilon^r}\right)\right|^2\varrho_\epsilon^q dx &= \int_{\mathcal{Z}^+_{\epsilon_0}}\left|\nabla\left(\frac{u}{\varrho_\epsilon^r}\right)\right|^2\varrho_\epsilon^q dx + \int_{\mathcal{Z}\setminus\mathcal{Z}^+_{\epsilon_0}}\left|\nabla\left(\frac{u}{\varrho_\epsilon^r}\right)\right|^2\varrho_\epsilon^q dx\\ &\geq \Lambda_\epsilon(\mathcal{Z}^+_{\epsilon_0})\int_{\mathcal{Z}^+_{\epsilon_0}}\left|\frac{u}{\varrho_\epsilon^r}-\bar{u}_{\mathcal{Z}^+_{\epsilon_0}}\right|^2\varrho_\epsilon^{p+r}dx\\ &\quad+\Lambda_\epsilon(\mathcal{Z}\setminus\mathcal{Z}^+_{\epsilon_0})\int_{\mathcal{Z}\setminus\mathcal{Z}^+_{\epsilon_0}}\left|\frac{u}{\varrho_\epsilon^r}-\bar{u}_{\mathcal{Z}\setminus\mathcal{Z}^+_{\epsilon_0}}\right|^2\varrho_\epsilon^{p+r}dx,\end{aligned}$$

where for subsets $\Omega\subseteq\mathcal{Z}$ we used the notation (recall (2.2))

$$\bar{u}_\Omega := \frac{1}{|\Omega|_{\varrho_\epsilon^{p+r}}}\int_\Omega\left(\frac{u}{\varrho_\epsilon^r}\right)\varrho_\epsilon^{p+r}dx. \tag{6.7}$$

After expanding the squared absolute values and rearrangement we get

$$\begin{aligned}\frac{1}{\Lambda(\epsilon)}\int_{\mathcal{Z}}\left|\nabla\left(\frac{u}{\varrho_\epsilon^r}\right)\right|^2\varrho_\epsilon^q dx &\geq \int_{\mathcal{Z}}\left|\frac{u}{\varrho_\epsilon^r}\right|^2\varrho_\epsilon^{p+r}dx\\ &\quad+\bar{u}^2_{\mathcal{Z}^+_{\epsilon_0}}|\mathcal{Z}^+_{\epsilon_0}|_{\varrho_\epsilon^{p+r}}+\bar{u}^2_{\mathcal{Z}\setminus\mathcal{Z}^+_{\epsilon_0}}|\mathcal{Z}\setminus\mathcal{Z}^+_{\epsilon_0}|_{\varrho_\epsilon^{p+r}}\\ &\quad-2\bar{u}_{\mathcal{Z}^+_{\epsilon_0}}\int_{\mathcal{Z}^+_{\epsilon_0}}u\varrho_\epsilon^p dx-2\bar{u}_{\mathcal{Z}\setminus\mathcal{Z}^+_{\epsilon_0}}\int_{\mathcal{Z}\setminus\mathcal{Z}^+_{\epsilon_0}}u\varrho_\epsilon^p dx.\end{aligned} \tag{6.8}$$

We further discard the terms in the second line as they are positive, which leaves us with the lower bound:

$$\frac{1}{\Lambda(\epsilon)}\int\limits_{\mathcal{Z}}\left|\nabla\left(\frac{u}{\varrho_\epsilon^r}\right)\right|^2\varrho_\epsilon^q dx \geq \int\limits_{\mathcal{Z}}\left|\frac{u}{\varrho_\epsilon^r}\right|^2\varrho_\epsilon^{p+r}dx - 2\left(\bar{u}_{\mathcal{Z}_{\epsilon_0}^+}\int\limits_{\mathcal{Z}_{\epsilon_0}^+}u\varrho_\epsilon^p dx + \bar{u}_{\mathcal{Z}\setminus\mathcal{Z}_{\epsilon_0}^+}\int\limits_{\mathcal{Z}\setminus\mathcal{Z}_\epsilon^+}u\varrho_{\epsilon_0}^p dx\right).$$

Using Hölder's inequality and the normalization $\|u\|_{L^2(\mathcal{Z},\varrho_\epsilon^{p-r})}=1$ we obtain

$$\begin{aligned}&\frac{1}{\Lambda(\epsilon)}\int\limits_{\mathcal{Z}}\left|\nabla\left(\frac{u}{\varrho_\epsilon^r}\right)\right|^2\varrho_\epsilon^q dx\\ &\quad\geq\left[1-2\left(\bar{u}_{\mathcal{Z}_{\epsilon_0}^+}|\mathcal{Z}_{\epsilon_0}^+|^{1/2}_{\varrho_\epsilon^{p+r}}+\bar{u}_{\mathcal{Z}\setminus\mathcal{Z}_{\epsilon_0}^+}|\mathcal{Z}\setminus\mathcal{Z}_{\epsilon_0}^+|^{1/2}_{\varrho_\epsilon^{p+r}}\right)\right]\int\limits_{\mathcal{Z}}\left|\frac{u}{\varrho_\epsilon^r}\right|^2\varrho_\epsilon^{p+r}dx\\ &\quad=:\left[1-2(T_1+T_2)\right].\end{aligned}$$

It remains to bound the T_1 and T_2 terms.

Recall that $\langle u,\varrho_\epsilon^r\rangle_V=0$, implying that $\int_{\mathcal{Z}}u\varrho_\epsilon^p dx=0$ and so

$$\int\limits_{\mathcal{Z}_{\epsilon_0}^+}u\varrho_\epsilon^p dx+\int\limits_{\mathcal{Z}_{\epsilon_0}^-}u\varrho_\epsilon^p dx=-\int\limits_{\mathcal{Z}\setminus\mathcal{Z}'_{\epsilon_0}}u\varrho_\epsilon^p dx. \tag{6.9}$$

On the other hand since $\langle u,\varphi_{F,\epsilon}\rangle_V=0$ as well we have that

$$\begin{aligned}0=&b_\epsilon^+\int\limits_{\mathcal{Z}}u\xi_\epsilon^+\varrho_\epsilon^p dx+b_\epsilon^+\int\limits_{\mathcal{Z}}\varrho_\epsilon^q\nabla\left(\frac{u}{\varrho_\epsilon^r}\right)\cdot\nabla\xi_\epsilon^+dx\\ &-b_\epsilon^-\int\limits_{\mathcal{Z}}u\xi_\epsilon^-\varrho_\epsilon^p dx-b_\epsilon^-\int\limits_{\mathcal{Z}}\varrho_\epsilon^q\nabla\left(\frac{u}{\varrho_\epsilon^r}\right)\cdot\nabla\xi_\epsilon^-dx.\end{aligned}$$

Using the definition of $\xi_\epsilon^\pm$ we can write

$$\begin{aligned}\int\limits_{\mathcal{Z}_{\epsilon_0}^+}u\varrho_\epsilon^p dx-\int\limits_{\mathcal{Z}_{\epsilon_0}^-}u\varrho_\epsilon^p dx=&\int\limits_{\mathcal{Z}_{\epsilon_0}^+\setminus\mathcal{Z}_\epsilon^+}u\varrho_\epsilon^p dx-\int\limits_{\mathcal{Z}_{\epsilon_0}^-\setminus\mathcal{Z}_\epsilon^-}u\varrho_\epsilon^p dx\\ &-b_\epsilon^+\int\limits_{\mathcal{Z}_{\epsilon_1}^+\setminus\mathcal{Z}_\epsilon^+}u\varrho_\epsilon^p\xi_\epsilon^+dx+b_\epsilon^-\int\limits_{\mathcal{Z}_{\epsilon_1}^-\setminus\mathcal{Z}_\epsilon^-}u\varrho_\epsilon^p\xi_\epsilon^-dx\\ &+(1-b_\epsilon^+)\int\limits_{\mathcal{Z}_\epsilon^+}u\varrho_\epsilon^p dx-(1-b_\epsilon^-)\int\limits_{\mathcal{Z}_\epsilon^-}u\varrho_\epsilon^p dx\\ &-b_\epsilon^+\int\limits_{\mathcal{Z}_{\epsilon_1}^+\setminus\mathcal{Z}_\epsilon^+}\varrho_\epsilon^q\nabla\left(\frac{u}{\varrho_\epsilon^r}\right)\cdot\nabla\xi_\epsilon^+dx+b_\epsilon^-\int\limits_{\mathcal{Z}_{\epsilon_1}^-\setminus\mathcal{Z}_\epsilon^-}\varrho_\epsilon^q\nabla\left(\frac{u}{\varrho_\epsilon^r}\right)\cdot\nabla\xi_\epsilon^-dx.\end{aligned} \tag{6.10}$$

Furthermore, by the Cauchy-Schwartz inequality, and the bound on the derivative of $\xi_\epsilon^\pm$ we obtain

$$\left|\int\limits_{\mathcal{Z}_{\epsilon_1}^\pm\setminus\mathcal{Z}_\epsilon^\pm}\varrho_\epsilon^q\nabla\left(\frac{u}{\varrho_\epsilon^r}\right)\cdot\nabla\xi_\epsilon^\pm dx\right|^2\leq\mathcal{R}_\epsilon(u)\int\limits_{\mathcal{Z}_{\epsilon_1}^\pm\setminus\mathcal{Z}_\epsilon^\pm}|\nabla\xi_\epsilon^\pm|^2\varrho_\epsilon^q dx\leq\Xi_1\mathcal{R}_\epsilon(u)\epsilon^{q-2\beta} \tag{6.11}$$

for a constant $\Xi_1 > 0$ independent of ϵ. Combine (6.9), (6.10), (6.11), and the fact that $0 < b_\epsilon^\pm < 2$ to get

$$2\left|\int_{\mathcal{Z}_{\epsilon_0}^+} u\varrho_\epsilon^p dx\right| \leq \int_{\mathcal{Z}\setminus\mathcal{Z}'_\epsilon} |u\varrho_\epsilon^p| dx + 2\int_{\mathcal{Z}'_{\epsilon_1}\setminus\mathcal{Z}'_\epsilon} |u\varrho_\epsilon^p| dx + \max\{|1-b_\epsilon^+|, |1-b_\epsilon^-|\}\int_{\mathcal{Z}'_\epsilon} |u\varrho_\epsilon^p| dx + 4\sqrt{\Xi_1}\mathcal{R}_\epsilon(u)^{1/2}\epsilon^{\frac{q}{2}-\beta}.$$

Multiple applications of Hölder's inequality along with Lemma 6.3 then give

$$2\left|\int_{\mathcal{Z}_{\epsilon_0}^+} u\varrho_\epsilon^p dx\right| \leq |\mathcal{Z}\setminus\mathcal{Z}'_\epsilon|_{\varrho_\epsilon^{p+r}}^{1/2} + 2\left|\mathcal{Z}'_{\epsilon_1}\setminus\mathcal{Z}'_\epsilon\right|_{\varrho_\epsilon^{p+r}}^{1/2} + \Xi\epsilon^{\min\{1,p+r\}}|\mathcal{Z}'_\epsilon|_{\varrho_\epsilon^{p+r}}^{1/2} + 4\sqrt{\Xi_1}\mathcal{R}_\epsilon(u)^{1/2}\epsilon^{\frac{q}{2}-\beta}.$$

Furthermore, by Assumption 2.6(d) and (2.8),

$$|\mathcal{Z}\setminus\mathcal{Z}'_\epsilon|_{\varrho_\epsilon^{p+r}} = K_2^{p+r}\epsilon^{p+r}|\mathcal{Z}\setminus\mathcal{Z}'_\epsilon| \leq \Xi_2\epsilon^{p+r},$$
$$\left|\mathcal{Z}'_{\epsilon_1}\setminus\mathcal{Z}'_\epsilon\right|_{\varrho_\epsilon^{p+r}} = K_2^{p+r}\epsilon^{p+r}|\mathcal{Z}'_{\epsilon_1}\setminus\mathcal{Z}'_\epsilon| \leq \Xi_3\epsilon^{p+r+\beta}.$$

We can repeat the above calculation by replacing $\mathcal{Z}^+$ with $\mathcal{Z}^-$ and vice versa to get the bound

$$\left|\int_{\mathcal{Z}_{\epsilon_0}^\pm} u\varrho_\epsilon^p dx\right| \leq \Xi_4\epsilon^{\frac{1}{2}\min\{2,p+r\}} + 4\sqrt{\Xi_1}\mathcal{R}_\epsilon(u)^{1/2}\epsilon^{\frac{q}{2}-\beta},$$

for some constant $\Xi_4 > 0$. Note that by (6.9), we also have

$$\left|\int_{\mathcal{Z}\setminus\mathcal{Z}'_{\epsilon_0}} u\varrho_\epsilon^p dx\right| \leq 2\Xi_4\epsilon^{\frac{1}{2}\min\{2,p+r\}} + 8\sqrt{\Xi_1}\mathcal{R}_\epsilon(u)^{1/2}\epsilon^{\frac{q}{2}-\beta}.$$

We conclude that

$$\begin{aligned}|T_1| + |T_2| &= \frac{|\mathcal{Z}_{\epsilon_0}^+|_{\varrho_\epsilon^{p+r}}^{1/2}}{|\mathcal{Z}_{\epsilon_0}^+|_{\varrho_\epsilon^{p+r}}}\left|\int_{\mathcal{Z}_{\epsilon_0}^+} u\varrho_\epsilon^p dx\right| + \frac{|\mathcal{Z}\setminus\mathcal{Z}_{\epsilon_0}^+|_{\varrho_\epsilon^{p+r}}^{1/2}}{|\mathcal{Z}\setminus\mathcal{Z}_{\epsilon_0}^+|_{\varrho_\epsilon^{p+r}}}\left|\int_{\mathcal{Z}_{\epsilon_0}^-} u\varrho_\epsilon^p dx + \int_{\mathcal{Z}\setminus\mathcal{Z}'_{\epsilon_0}} u\varrho_\epsilon^p dx\right| \\ &\leq \frac{1}{|\mathcal{Z}_{\epsilon_0}^+|_{\varrho_\epsilon^{p+r}}^{1/2}}\left|\int_{\mathcal{Z}_{\epsilon_0}^+} u\varrho_\epsilon^p dx\right| + \frac{1}{|\mathcal{Z}\setminus\mathcal{Z}_{\epsilon_0}^+|_{\varrho_\epsilon^{p+r}}^{1/2}}\left(\left|\int_{\mathcal{Z}_{\epsilon_0}^-} u\varrho_\epsilon^p dx\right| + \left|\int_{\mathcal{Z}\setminus\mathcal{Z}'_{\epsilon_0}} u\varrho_\epsilon^p dx\right|\right) \\ &\leq \Xi_5\epsilon^{\frac{1}{2}\min\{2,p+r\}} + \Xi_6\mathcal{R}_\epsilon^{1/2}\epsilon^{\frac{q}{2}-\beta}.\end{aligned}$$

Thus, we obtain

$$\mathcal{R}_\epsilon(u) + \Lambda(\epsilon)\Xi_6\mathcal{R}_\epsilon^{1/2}(u)\epsilon^{\frac{q}{2}-\beta} \geq \Lambda(\epsilon)\left[1 - 2\Xi_5\epsilon^{\frac{1}{2}\min\{2,p+r\}}\right].$$

Now if $\mathcal{R}_\epsilon(u) \geq 1$ then $\mathcal{R}_\epsilon(u) \geq \mathcal{R}_\epsilon^{1/2}(u)$ and we have

$$\mathcal{R}_\epsilon(u) \geq \frac{\Lambda(\epsilon)\left[1 - 2\Xi_5 \epsilon^{\frac{1}{2}\min\{2,p+r\}}\right]}{1 + 2\Lambda(\epsilon)\Xi_6 \epsilon^{\frac{q}{2}-\beta}}.$$

Alternatively, if $\mathcal{R}_\epsilon(u) < 1$ then $\mathcal{R}_\epsilon^{1/2}(u) < 1$ and we instead obtain

$$\mathcal{R}_\epsilon(u) \geq \Lambda(\epsilon)\left[1 - 2\Xi_8 \epsilon^{\frac{1}{2}\min\{2,p+r,q-2\beta\}}\right].$$

Combining these two bounds we get the desired result so long as $p + r > 0$, $q > 0$, and β and ϵ_0 are small enough. □

Next, we investigate the consequences of Proposition 6.5 for different parameter choices p, q and r. The main point of interest here is to analyze how the parameter $\Lambda(\epsilon)$ in (6.6) is controlled by ϵ. We will show in Proposition 6.7 that the choice of q in relation to p and r plays a major role in whether $\Lambda(\epsilon)$ is uniformly bounded away from zero and hence, whether a uniform spectral gap exists between $\sigma_{2,\epsilon}$ and $\sigma_{3,\epsilon}$.

Our method of proof relies on isoperimetric-type inequalities for general Dirichlet forms as in [2, Sec. 8.5.1] viewed as a generalized form of Cheeger's inequality. For open $\Omega \subset \mathcal{Z}$ define the ϱ_ϵ^q weighted Minkowski boundary measure of Ω as follows

$$|\partial\Omega|_{\varrho_\epsilon^q} := \liminf_{\delta \downarrow 0} \frac{1}{\delta}\left[|\Omega_\delta|_{\varrho_\epsilon^q} - |\Omega|_{\varrho_\epsilon^q}\right]. \tag{6.12}$$

Furthermore, given p, q, r we fix a subset $\Omega' \subseteq \mathcal{Z}$ and consider any $\Omega \subset \Omega' \subseteq \mathcal{Z}$. Define the isoperimetric function

$$\mathcal{J}(\Omega, \varrho_\epsilon) := \frac{|\partial\Omega|_{\varrho_\epsilon^q}}{\min\{|\Omega|_{\varrho_\epsilon^{p+r}}, |\Omega' \setminus \Omega|_{\varrho_\epsilon^{p+r}}\}}. \tag{6.13}$$

The following lemma is proven in Appendix D similarly to [2, Prop. 8.5.2].

Lemma 6.6. *Let $(p, q, r) \in \mathbb{R}^3$, and suppose Assumptions 2.4, 2.5 and 2.6 hold. Let $\Omega' \subseteq \mathcal{Z}$. Fix $\epsilon \in (0, \epsilon_0)$. Assume there exist $h(\epsilon) > 0$ so that*

$$h(\epsilon) \leq \inf_{\Omega} \mathcal{J}(\Omega, \varrho_\epsilon), \tag{6.14}$$

where the infimum is over open subsets $\Omega \subset \Omega' \subseteq \mathcal{Z}$ such that $|\Omega|_{\varrho_\epsilon^{p+r}} \leq \frac{1}{2}|\Omega'|_{\varrho_\epsilon^{p+r}}$. Then $\mathcal{L}_\epsilon$ has a spectral gap on Ω' according to Definition 2.11 and (2.18) holds with

$$\Lambda_\epsilon(\Omega') \geq \frac{h(\epsilon)^2}{4}\left(\inf_{\Omega'} \varrho_\epsilon^{p+r-q}\right).$$

Proposition 6.7. *Let $r \in \mathbb{R}$, $q > 0$, $p + r > 0$, and suppose Assumptions 2.4, 2.5 and 2.6 hold. Then there exists $\Xi > 0$ independent of $\epsilon \in (0, \epsilon_0]$ so that*

$$\sigma_{3,\epsilon} \geq \Xi \epsilon^{\max\{p+r-q, 2(q-p-r)\}}.$$

Proof. By Proposition 6.5 we only need to find a lower bound on $\Lambda(\epsilon)$ which in turn requires us to find a lower bound on $\Lambda_\epsilon(\mathcal{Z}_{\epsilon_0}^+)$ and $\Lambda_\epsilon(\mathcal{Z} \setminus \mathcal{Z}_{\epsilon_0}^+)$ separately. We only consider $\Lambda_\epsilon(\mathcal{Z}_{\epsilon_0}^+)$ and note that the same argument can be repeated for $\Lambda_\epsilon(\mathcal{Z} \setminus \mathcal{Z}_{\epsilon_0}^+)$ possibly with different constants.

We will find a lower bound on $\inf_\Omega \mathcal{J}(\Omega, \varrho_\epsilon)$ and use Lemma 6.6 with $\Omega' \equiv \mathcal{Z}^+_{\epsilon_0}$ to extend that lower bound to $\Lambda_\epsilon(\mathcal{Z}^+_{\epsilon_0})$. For fixed ϵ let Ω be a subset of $\mathcal{Z}^+_{\epsilon_0}$ satisfying $|\Omega|_{\varrho_\epsilon^{p+r}} \leq \frac{1}{2}|\mathcal{Z}^+_{\epsilon_0}|_{\varrho_\epsilon^{p+r}}$. First, suppose $|\Omega \cap \mathcal{Z}^+|_{\varrho_\epsilon^{p+r}} > 0$, i.e., part of Ω lies inside $\mathcal{Z}^+$. Then since ϱ_ϵ is uniformly bounded from above in $\mathcal{Z}^+$ and for sufficiently small ϵ_0 (recall (2.9)) we have

$$\begin{aligned}
\mathcal{J}(\Omega, \varrho_\epsilon) &\geq \frac{(\varrho^-_{\epsilon_0})^q |\partial\Omega \cap \mathcal{Z}^+|}{\min\{|\Omega|_{\varrho_\epsilon^{p+r}}, |\mathcal{Z}^+_{\epsilon_0} \setminus \Omega|_{\varrho_\epsilon^{p+r}}\}} \\
&\geq \frac{(\varrho^-_{\epsilon_0})^q |\partial\Omega \cap \mathcal{Z}^+|}{(\varrho^+_{\epsilon_0})^{p+r} |\Omega \cap \mathcal{Z}^+| + |\Omega \cap \left(\mathcal{Z}^+_{\epsilon_0} \setminus \mathcal{Z}^+\right)|_{\varrho_\epsilon^{p+r}}} \\
&\geq \frac{(\varrho^-_{\epsilon_0})^q |\partial\Omega \cap \mathcal{Z}^+|}{(\varrho^+_{\epsilon_0})^{p+r} |\Omega \cap \mathcal{Z}^+|} + \mathcal{O}(|\mathcal{Z}^+_{\epsilon_0} \setminus \mathcal{Z}^+|_{\varrho_\epsilon^{p+r}}) \geq \Xi_1 \,,
\end{aligned}$$

where we used Taylor expansions to write the last line. The first ratio is uniformly bounded away from zero independent of ϵ by the standard isoperimetric inequality for the set $\Omega \cap \mathcal{Z}^+$ while the second term is small following our assumptions on ϱ_ϵ. Thus, in this case $\mathcal{J}$ is uniformly bounded from below.

Now consider the case where $|\Omega \cap \mathcal{Z}^+|_{\varrho_\epsilon^{p+r}} = 0$, and so Ω lies entirely in the strip $\mathcal{Z}^+_{\epsilon_0} \setminus \mathcal{Z}^+$ but $|\Omega \cap \mathcal{Z}^+_\epsilon|_{\varrho_\epsilon^{p+r}} > 0$. Then it is possible to have $|\partial\Omega \cap \partial\mathcal{Z}^+|_{\varrho_\epsilon^q} > 0$ or for the boundary of Ω to touch the boundary $\partial\mathcal{Z}^+$ on a null set. Then similar calculations to the above yield

$$\begin{aligned}
\mathcal{J}(\Omega, \varrho_\epsilon) &= \frac{|\partial\Omega|_{\varrho_\epsilon^q}}{\min\{|\Omega|_{\varrho_\epsilon^{p+r}}, |\mathcal{Z}^+_{\epsilon_0} \setminus \Omega|_{\varrho_\epsilon^{p+r}}\}} \\
&\geq \frac{(\varrho^-_{\epsilon_0})^q |\partial\Omega \cap \mathcal{Z}^+|}{|\Omega|_{\varrho_\epsilon^{p+r}}} \\
&= \frac{(\varrho^-_{\epsilon_0})^q |\partial\Omega \cap \mathcal{Z}^+|}{|\Omega \cap \mathcal{Z}^+_\epsilon|_{\varrho_\epsilon^{p+r}} + K_2^{p+r}\epsilon^{p+r}|\Omega \cap \left(\mathcal{Z}^+_{\epsilon_0} \setminus \mathcal{Z}^+_\epsilon\right)|} \\
&\geq \frac{(\varrho^-_{\epsilon_0})^q |\partial\Omega \cap \mathcal{Z}^+|}{\Xi_3\epsilon|\partial\Omega \cap \bar{\mathcal{Z}}^+| + K_2^{p+r}\epsilon^{p+r}|\Omega \cap \left(\mathcal{Z}^+_{\epsilon_0} \setminus \mathcal{Z}^+_\epsilon\right)|} \,,
\end{aligned}$$

and so the lower bound on $\mathcal{J}$ blows up as $\epsilon \to 0$.

Finally, we consider the case where $|\Omega \cap \mathcal{Z}^+_\epsilon|_{\varrho_\epsilon^{p+r}} = 0$, and so $\partial\Omega$ is far from $\partial\mathcal{Z}^+$. Proceeding as above, we write

$$\begin{aligned}
\mathcal{J}(\Omega, \varrho_\epsilon) &\geq \frac{|\partial\Omega|_{\varrho_\epsilon^q}}{|\Omega|_{\varrho_\epsilon^{p+r}}} = \frac{(K_2\epsilon)^q |\partial\Omega|}{(K_2\epsilon)^{p+r}|\Omega|} \\
&\geq \Xi_4 \epsilon^{q-p-r},
\end{aligned}$$

where Ξ_4 depends on K_2^{q-p-r} and the standard isoperimetric constant.

Summarizing, if $q \leq p + r$, then $\mathcal{J}$ is bounded away from zero by a uniform constant independent of ϵ, implying that (6.14) holds with a uniform constant $h > 0$. Note that $\inf_{\mathcal{Z}^+_{\epsilon_0}} \varrho_\epsilon^{p+r-q} = K_2^{p+r-q}\epsilon^{p+r-q}$ by Assumption 2.6(d). We now investigate the different cases of (p, q, r) separately:

- if $q = p + r$, we obtain a uniform lower bound on $\Lambda_\epsilon(\mathcal{Z}^+_{\epsilon_0})$ by Lemma 6.6;
- if $q < p + r$ on the other hand, the lower bound on $\Lambda_\epsilon(\mathcal{Z}^+_{\epsilon_0})$ is of order ϵ^{p+r-q};
- if $q > p + r$, then we have the lower bound $\mathcal{J} \geq \Xi_4\epsilon^{q-p-r}$ and Lemma 6.6 implies $\Lambda_\epsilon(\mathcal{Z}^+_{\epsilon_0}) \geq \Xi_4^2\epsilon^{2(q-p-r)}/4$.

Note that in the final bullet the factor $\inf_{\mathcal{Z}^+_{\epsilon_0}} \varrho_\epsilon^{p+r-q}$ does not play a role here thanks to the uniform upper bound on ϱ_ϵ guaranteed in Assumption 2.6(c). The exact same reasoning can be applied for the set $\Omega' = \mathcal{Z} \setminus \mathcal{Z}^+_{\epsilon_0}$, where $\mathcal{Z}^-$ plays the role of $\mathcal{Z}^+$, and the region around $\mathcal{Z}^-$ where ϱ_ϵ is of order ϵ is simply extended up to the boundary of $\mathcal{Z}^+_{\epsilon_0}$. Therefore, similar bounds also hold for $\Lambda_\epsilon(\mathcal{Z} \setminus \mathcal{Z}^+_{\epsilon_0})$ in each case. By combining all the above lower bounds into one expression, Proposition 6.5 yields the existence of a constant $\Xi > 0$ so that $\sigma_{3,\epsilon} \geq \Xi \epsilon^{\max\{p+r-q, 2(q-p-r)\}}$ as claimed. □

The last proposition suggests that when $q \neq p+r$ we cannot hope for a spectral gap. Indeed, we are able to obtain a vanishing upper bound on $\sigma_{3,\epsilon}$ for $q > p+r$ and quantify how fast it approaches zero in that case and ultimately obtain a spectral ratio gap.

Proposition 6.8. *Suppose the conditions of Proposition 6.7 are satisfied.*

- *If $q > p+r$ and $\epsilon_0 > 0$ is sufficiently small, then there exists a constant $\Xi_1 > 0$ depending only on $\Lambda_\Delta(\mathcal{Z} \setminus \mathcal{Z}'_{\epsilon_0})$ so that $\forall (\epsilon, \beta) \in (0, \epsilon_0) \times (0,1)$,*

$$\sigma_{3,\epsilon} \leq \Xi_1 \epsilon^{q-p-r-2\beta}.$$

- *If $q \leq p+r$ and $\epsilon_0 > 0$ is sufficiently small, then there exists a constant $\Xi_2 > 0$ depending only on $\Lambda_\Delta(\mathcal{Z} \setminus \mathcal{Z}'_{\epsilon_0})$ so that $\forall \epsilon \in (0, \epsilon_0)$,*

$$\sigma_{3,\epsilon} \leq \Xi_2 \,.$$

Note that according to Definition 2.9, $\Lambda_\Delta(\mathcal{Z} \setminus \mathcal{Z}'_{\epsilon_0})$ is the second eigenvalue of the standard Laplacian on $\mathcal{Z} \setminus \mathcal{Z}'_{\epsilon_0}$.

Proof. We apply a similar argument to the proof of Proposition 6.4 using the min-max principle. Let $\tilde{\varphi}_2 \in H^1(\mathcal{Z} \setminus \mathcal{Z}'_{\epsilon_0})$ denote the second eigenfunction of the standard Laplacian on $\mathcal{Z} \setminus \mathcal{Z}'_{\epsilon_0}$, i.e., $\tilde{\varphi}_2 \perp \mathbf{1}_{\mathcal{Z} \setminus \mathcal{Z}'_{\epsilon_0}}$ and

$$\int_{\mathcal{Z} \setminus \mathcal{Z}'_{\epsilon_0}} |\nabla \tilde{\varphi}_2|^2 dx = \Lambda_\Delta(\mathcal{Z} \setminus \mathcal{Z}'_{\epsilon_0}) \|\tilde{\varphi}_2\|^2_{L^2(\mathcal{Z} \setminus \mathcal{Z}'_{\epsilon_0})}.$$

We proceed by constructing a suitable approximation to $\tilde{\varphi}_2$. Let $\epsilon_2 := \epsilon_0 + \epsilon^\beta$ and $\epsilon_3 := \epsilon_0 + 2\epsilon^\beta$ for $0 < \beta < 1$. In a similar manner to (6.3), we define a function $\tilde{\xi}_\epsilon$ (see Fig. 4.1)

$$\begin{aligned} &\tilde{\xi}_\epsilon(x) = 1, && x \in \mathcal{Z} \setminus \mathcal{Z}'_{\epsilon_3}, \\ &0 < \tilde{\xi}_\epsilon(x) < 1, \quad |\nabla \tilde{\xi}_\epsilon(x)| \leq \vartheta \epsilon^{-\beta}, && x \in \mathcal{Z}'_{\epsilon_3} \setminus \mathcal{Z}'_{\epsilon_2}, \\ &\tilde{\xi}_\epsilon(x) = 0, && x \in \mathcal{Z}'_{\epsilon_2}. \end{aligned}$$

This allows us to define the function

$$\tilde{\varphi}_{F,\epsilon} := \tilde{\xi}_\epsilon \tilde{\varphi}_2 - \frac{\varrho_\epsilon^r}{|\mathcal{Z} \setminus \mathcal{Z}'_{\epsilon_2}|_{\varrho_\epsilon^{p+r}}} \int_{\mathcal{Z} \setminus \mathcal{Z}'_{\epsilon_2}} \tilde{\xi}_\epsilon \tilde{\varphi}_2 \varrho_\epsilon^p \, dx. \tag{6.15}$$

The shift ensures that $\tilde{\varphi}_{F,\epsilon} \in V^1(\mathcal{Z} \setminus \mathcal{Z}'_{\epsilon_2}, \varrho_\epsilon)$. The choice of ϵ_2 and ϵ_3 guarantee that the supports of $\tilde{\varphi}_{F,\epsilon}$ and $\varphi_{F,\epsilon}$ are disjoint, and so they are orthogonal in $V^1(\mathcal{Z}, \varrho_\epsilon)$. Now let $u \in \operatorname{span}\{\varphi_{F,\epsilon}, \tilde{\varphi}_{F,\epsilon}\}$. We wish to

bound $\mathcal{R}_\epsilon(u)$. A straightforward calculation shows that since $\varphi_{F,\epsilon} \perp \tilde{\varphi}_{F,\epsilon}$ it suffices to bound $\mathcal{R}_\epsilon(\varphi_{F,\epsilon})$ and $\mathcal{R}_\epsilon(\tilde{\varphi}_{F,\epsilon})$ separately.

For $\varphi_{F,\epsilon}$ we showed in the proof of Proposition 6.4 the existence of $\Xi_1 > 0$ so that

$$\mathcal{R}_\epsilon(\varphi_{F,\epsilon}) \leq \Xi_1 \epsilon^{q-\beta},$$

for any $\beta \in (0, q)$. To estimate $R_\epsilon(\tilde{\varphi}_{F,\epsilon})$, observe that for $\epsilon \in (0, \epsilon_0]$ the function $\tilde{\xi}_\epsilon \tilde{\varphi}_2$ is in $H^1(\mathcal{Z} \setminus \mathcal{Z}'_{\epsilon_2})$. Thus, following our assumptions on ϱ_ϵ we can write

$$\begin{aligned}
\|\tilde{\varphi}_{F,\epsilon}\|^2_{L^2(\mathcal{Z},\varrho_\epsilon^{p-r})} \mathcal{R}_\epsilon(\tilde{\varphi}_{F,\epsilon}) &= \int\limits_{\mathcal{Z}\setminus\mathcal{Z}'_{\epsilon_2}} \left| \nabla \left(\frac{\tilde{\varphi}_{F,\epsilon}}{\varrho_\epsilon^r} \right) \right|^2 \varrho_\epsilon^q dx \\
&= K_2^{q-2r} \epsilon^{q-2r} \int\limits_{\mathcal{Z}\setminus\mathcal{Z}'_{\epsilon_2}} \left| \nabla \left(\tilde{\xi}_\epsilon \tilde{\varphi}_2 \right) \right|^2 dx \\
&\leq 2K_2^{q-2r} \epsilon^{q-2r} \left(\int\limits_{\mathcal{Z}\setminus\mathcal{Z}'_{\epsilon_2}} \left| \tilde{\xi}_\epsilon \nabla \tilde{\varphi}_2 \right|^2 dx + \int\limits_{\mathcal{Z}\setminus\mathcal{Z}'_{\epsilon_2}} \left| \tilde{\varphi}_2 \nabla \tilde{\xi}_\epsilon \right|^2 dx \right) \\
&\leq 2K_2^{q-2r} \epsilon^{q-2r} \left(\int\limits_{\mathcal{Z}\setminus\mathcal{Z}'_{\epsilon_0}} \left| \nabla \tilde{\varphi}_2 \right|^2 dx + \int\limits_{\mathcal{Z}'_{\epsilon_3}\setminus\mathcal{Z}'_{\epsilon_2}} \left| \tilde{\varphi}_2 \nabla \tilde{\xi}_\epsilon \right|^2 dx \right) \\
&\leq 2K_2^{q-2r} \epsilon^{q-2r} \left(\int\limits_{\mathcal{Z}\setminus\mathcal{Z}'_{\epsilon_0}} \left| \nabla \tilde{\varphi}_2 \right|^2 dx + \vartheta^2 \epsilon^{-2\beta} \int\limits_{\mathcal{Z}\setminus\mathcal{Z}'_{\epsilon_0}} \left| \tilde{\varphi}_2 \right|^2 dx \right) \\
&\leq 2K_2^{q-2r} \epsilon^{q-2r} \left(\Lambda_\Delta(\mathcal{Z} \setminus \mathcal{Z}'_{\epsilon_0}) + \vartheta^2 \epsilon^{-2\beta} \right) \|\tilde{\varphi}_2\|^2_{L^2(\mathcal{Z}\setminus\mathcal{Z}'_{\epsilon_0})}.
\end{aligned} \tag{6.16}$$

Next, we bound $\|\tilde{\varphi}_{F,\epsilon}\|^2_{L^2(\mathcal{Z},\varrho_\epsilon^{p-r})}$ from below. We have

$$\begin{aligned}
\|\tilde{\xi}_\epsilon \tilde{\varphi}_2\|^2_{L^2(\mathcal{Z},\varrho_\epsilon^{p-r})} &= K_2^{p-r} \epsilon^{p-r} \int\limits_{\mathcal{Z}\setminus\mathcal{Z}'_{\epsilon_0}} |\tilde{\xi}_\epsilon \tilde{\varphi}_2|^2 \, dx \\
&\geq K_2^{p-r} \epsilon^{p-r} \left(\int\limits_{\mathcal{Z}\setminus\mathcal{Z}'_{\epsilon_0}} |\tilde{\varphi}_2|^2 \, dx - \int\limits_{\mathcal{Z}'_{\epsilon_3}\setminus\mathcal{Z}'_{\epsilon_0}} |\tilde{\varphi}_2|^2 \, dx \right),
\end{aligned}$$

and for any $k \geq 2$ by Hölder's inequality,

$$\int\limits_{\mathcal{Z}'_{\epsilon_3}\setminus\mathcal{Z}'_{\epsilon_0}} |\tilde{\varphi}_2|^2 \, dx \leq \|\tilde{\varphi}_2\|^2_{L^k(\mathcal{Z}\setminus\mathcal{Z}'_{\epsilon_0})} |\mathcal{Z}'_{\epsilon_3} \setminus \mathcal{Z}'_{\epsilon_0}|^{\frac{k-2}{k}}.$$

By the Sobolev embedding theorem [1, Thm. 4.12], $\tilde{\varphi}_2 \in L^k(\mathcal{Z} \setminus \mathcal{Z}'_{\epsilon_0})$ for $k \in [2, 2d/(d-2))$ if $d > 2$ and $k \in [2, \infty)$ if $d \leq 2$; and so using Sobolev inequalities, and the fact that $\|\tilde{\varphi}_2\|_{L^2(\mathcal{Z}\setminus\mathcal{Z}'_{\epsilon_0})} \leq 1$,

$$\begin{aligned}
\|\tilde{\varphi}_2\|^2_{L^k(\mathcal{Z}\setminus\mathcal{Z}'_{\epsilon_0})} &\leq \Xi_2 \|\tilde{\varphi}_2\|^2_{H^1(\mathcal{Z}\setminus\mathcal{Z}'_{\epsilon_0})} = \Xi_2 \left(1 + \Lambda_\Delta(\mathcal{Z} \setminus \mathcal{Z}'_{\epsilon_0}))\right) \|\tilde{\varphi}_2\|^2_{L^2(\mathcal{Z}\setminus\mathcal{Z}'_{\epsilon_0})} \\
&\leq \Xi_2 \left(1 + \Lambda_\Delta(\mathcal{Z} \setminus \mathcal{Z}'_{\epsilon_0}))\right).
\end{aligned}$$

Since $|\mathcal{Z}'_{\epsilon_3} \setminus \mathcal{Z}'_{\epsilon_0}| \leq \Xi_3 \epsilon^\beta |\partial \mathcal{Z}'_{\epsilon_0}|$, we can write

$$\begin{aligned}\|\tilde{\xi}_\epsilon \tilde{\varphi}_2\|^2_{L^2(\mathcal{Z},\varrho_\epsilon^{p-r})} &\geq K_2^{p-r}\epsilon^{p-r}\left(\|\tilde{\varphi}_2\|^2_{L^2(\mathcal{Z}\setminus\mathcal{Z}'_{\epsilon_0})} - \Xi_2\left(1+\Lambda_\Delta(\mathcal{Z}\setminus\mathcal{Z}'_{\epsilon_0})\right)|\mathcal{Z}'_{\epsilon_3}\setminus\mathcal{Z}'_{\epsilon_0}|^{\frac{k-2}{k}}\right)\\ &\geq K_2^{p-r}\epsilon^{p-r}\left(\|\tilde{\varphi}_2\|^2_{L^2(\mathcal{Z}\setminus\mathcal{Z}'_{\epsilon_0})} - \Xi_4\epsilon^{\frac{\beta(k-2)}{k}}\right). \end{aligned} \tag{6.17}$$

Furthermore, using Assumption 2.6(d), the fact that $\tilde{\varphi}_2 \perp \mathbf{1}_{\mathcal{Z}\setminus\mathcal{Z}'_{\epsilon_0}}$ in $L^2(\mathcal{Z}\setminus\mathcal{Z}'_{\epsilon_0})$, Hölder's inequality, $\|\tilde{\varphi}_2\|_{L^2(\mathcal{Z}\setminus\mathcal{Z}'_{\epsilon_0})} \leq 1$, and the estimate (2.8), in that order, we can write

$$\begin{aligned}&\frac{\varrho_\epsilon^r}{|\mathcal{Z}\setminus\mathcal{Z}'_{\epsilon_2}|_{\varrho_\epsilon^{p+r}}}\left|\int_{\mathcal{Z}\setminus\mathcal{Z}'_{\epsilon_2}} \tilde{\xi}_\epsilon\tilde{\varphi}_2\varrho_\epsilon^p\,dx\right|\\ &\quad= \frac{1}{|\mathcal{Z}\setminus\mathcal{Z}'_{\epsilon_2}|}\left|\int_{\mathcal{Z}\setminus\mathcal{Z}'_{\epsilon_2}}\tilde{\xi}_\epsilon\tilde{\varphi}_2 dx\right|\\ &\quad= \frac{1}{|\mathcal{Z}\setminus\mathcal{Z}'_{\epsilon_2}|}\left|\int_{\mathcal{Z}\setminus\mathcal{Z}'_{\epsilon_0}}\tilde{\varphi}_2 dx - \int_{\mathcal{Z}'_{\epsilon_2}\setminus\mathcal{Z}'_{\epsilon_0}}\tilde{\varphi}_2 dx + \int_{\mathcal{Z}'_{\epsilon_3}\setminus\mathcal{Z}'_{\epsilon_2}}(\tilde{\xi}_\epsilon-1)\tilde{\varphi}_2 dx\right|\\ &\quad\leq \frac{1}{|\mathcal{Z}\setminus\mathcal{Z}'_{\epsilon_2}|}\left(\int_{\mathcal{Z}'_{\epsilon_2}\setminus\mathcal{Z}'_{\epsilon_0}}|\tilde{\varphi}_2|\,dx + \int_{\mathcal{Z}'_{\epsilon_3}\setminus\mathcal{Z}'_{\epsilon_2}}|\tilde{\xi}_\epsilon-1|\,|\tilde{\varphi}_2|\,dx\right)\\ &\quad\leq \frac{1}{|\mathcal{Z}\setminus\mathcal{Z}'_{\epsilon_2}|}\int_{\mathcal{Z}'_{\epsilon_3}\setminus\mathcal{Z}'_{\epsilon_0}}|\tilde{\varphi}_2|\,dx \leq \frac{|\mathcal{Z}'_{\epsilon_3}\setminus\mathcal{Z}'_{\epsilon_0}|^{1/2}}{|\mathcal{Z}\setminus\mathcal{Z}'_{\epsilon_0}| - |\mathcal{Z}'_{\epsilon_2}\setminus\mathcal{Z}'_{\epsilon_0}|} \leq \Xi_5\epsilon^{\beta/2}. \end{aligned} \tag{6.18}$$

To bound $\tilde{\varphi}_{F,\epsilon}$ on the outside set, we write explicitly

$$\begin{aligned}&\|\tilde{\varphi}_{F,\epsilon}\|^2_{L^2(\mathcal{Z},\varrho_\epsilon^{p-r})}\\ &\quad= \int_{\mathcal{Z}}\left|\tilde{\xi}_\epsilon(x)\tilde{\varphi}_2(x) - \frac{\varrho_\epsilon^r}{|\mathcal{Z}\setminus\mathcal{Z}'_{\epsilon_2}|_{\varrho_\epsilon^{p+r}}}\int_{\mathcal{Z}\setminus\mathcal{Z}'_{\epsilon_2}}\tilde{\xi}_\epsilon(y)\tilde{\varphi}_2(y)\varrho_\epsilon^p(y)\,dy\right|^2\varrho_\epsilon^{p-r}(x)\,dx\\ &\quad\geq \|\tilde{\xi}_\epsilon\tilde{\varphi}_2\|^2_{L^2(\mathcal{Z},\varrho_\epsilon^{p-r})} - \frac{2}{|\mathcal{Z}\setminus\mathcal{Z}'_{\epsilon_2}|_{\varrho_\epsilon^{p+r}}}\left|\int_{\mathcal{Z}\setminus\mathcal{Z}'_{\epsilon_2}}\tilde{\xi}_\epsilon\tilde{\varphi}_2\varrho_\epsilon^p\,dy\right|^2\\ &\quad= \|\tilde{\xi}_\epsilon\tilde{\varphi}_2\|^2_{L^2(\mathcal{Z},\varrho_\epsilon^{p-r})} - \frac{|\mathcal{Z}\setminus\mathcal{Z}'_{\epsilon_2}|^{p+r}_{\varrho_\epsilon^{p+r}}}{\varrho_\epsilon^{2r}}\left(\frac{2\varrho_\epsilon^r}{|\mathcal{Z}\setminus\mathcal{Z}'_{\epsilon_2}|_{\varrho_\epsilon^{2(p+r)}}}\left|\int_{\mathcal{Z}\setminus\mathcal{Z}'_{\epsilon_2}}\tilde{\xi}_\epsilon\tilde{\varphi}_2\varrho_\epsilon^p\,dy\right|^2\right)\\ &\quad= \|\tilde{\xi}_\epsilon\tilde{\varphi}_2\|^2_{L^2(\mathcal{Z},\varrho_\epsilon^{p-r})} - |\mathcal{Z}\setminus\mathcal{Z}'_{\epsilon_2}|K_2^{p-r}\epsilon^{p-r}\left(\frac{2\varrho_\epsilon^r}{|\mathcal{Z}\setminus\mathcal{Z}'_{\epsilon_2}|_{\varrho_\epsilon^{2(p+r)}}}\left|\int_{\mathcal{Z}\setminus\mathcal{Z}'_{\epsilon_2}}\tilde{\xi}_\epsilon\tilde{\varphi}_2\varrho_\epsilon^p\,dy\right|^2\right). \end{aligned}$$

Together with the bounds (6.17) and (6.18), we obtain for small enough ϵ_0,

$$\begin{aligned}\|\tilde{\varphi}_{F,\epsilon}\|^2_{L^2(\mathcal{Z},\varrho_\epsilon^{p-r})} &\geq K_2^{p-r}\epsilon^{p-r}\left(\|\tilde{\varphi}_2\|^2_{L^2(\mathcal{Z}\setminus\mathcal{Z}'_{\epsilon_0})} - \Xi_4\epsilon^{\frac{\beta(k-2)}{k}} - \Xi_6\epsilon^{\beta}\right)\\ &\geq \Xi_7\epsilon^{p-r}\|\tilde{\varphi}_2\|^2_{L^2(\mathcal{Z}\setminus\mathcal{Z}'_{\epsilon_0})}\,.\end{aligned}$$

Finally, following from (6.16), we infer the existence of a constant Ξ, independent of $\epsilon \in (0,\epsilon_0)$, so that

$$\mathcal{R}_\epsilon(\tilde{\varphi}_{F,\epsilon}) \leq \Xi\epsilon^{q-p-r-2\beta},$$

which concludes the proof. □

6.3. Proof of Theorem 3.4 (geometry of the second eigenfunction)

First, we prove a key result, that allows us to translate our bounds on the third eigenvalue $\sigma_{3,\epsilon}$ into an upper bound on the error between the second eigenfunction $\varphi_{2,\epsilon}$ and the approximate Fiedler vector $\varphi_{F,\epsilon}$.

Proposition 6.9. *Suppose there exist constants $\Xi_1,\Xi_2,\Xi_3 \geq 0$, so that for all $\epsilon \in (0,\epsilon_0]$,*

$$\sigma_{3,\epsilon} \geq \Xi_1 + \Xi_2\epsilon^{q-\vartheta} + \Xi_3\epsilon^{\theta-q}.$$

Then for every $0<\beta<q$, there exists $\Xi>0$ so that

$$\left|1 - \left\langle \frac{\varphi_{2,\epsilon}}{\varrho_\epsilon^r}, \frac{\varphi_{F,\epsilon}}{\varrho_\epsilon^r}\right\rangle^2_{\varrho_\epsilon^{p+r}}\right| \leq \frac{\Xi\epsilon^{q-\beta}}{\Xi_1 + \Xi_2\epsilon^{q-\vartheta} + \Xi_3\epsilon^{\theta-q}}\,.$$

Proof. Since $\left\langle \frac{\varphi_{2,\epsilon}}{\varrho_\epsilon^r}, \frac{\varphi_{F,\epsilon}}{\varrho_\epsilon^r}\right\rangle_{\varrho_\epsilon^{p+r}} \equiv \langle\varphi_{2,\epsilon},\varphi_{F,\epsilon}\rangle_{\varrho_\epsilon^{p-r}}$ we will work with the $L^2(\mathcal{Z},\varrho_\epsilon^{p-r})$ inner product for brevity. It follows from the spectral theorem [19, Thm. D.7] that $\varphi_{j,\epsilon}$ form an orthonormal basis in $L^2(\mathcal{Z},\varrho_\epsilon^{p-r})$. Let $\varphi_{F,\epsilon} = \sum_{j=1}^\infty h_j\varphi_{j,\epsilon}$ where $h_j = \langle\varphi_{j,\epsilon},\varphi_{F,\epsilon}\rangle_{\varrho_\epsilon^{p-r}}$. Note that $h_1 = 0$ since $\varphi_{F,\epsilon}\perp\varphi_{1,\epsilon}$. It follows from the calculation in (6.5) that for $\beta\in(0,q)$,

$$\Xi\epsilon^{q-\beta} \geq \mathcal{R}_\epsilon(\varphi_{F,\epsilon}) = \langle\mathcal{L}_{\varrho_\epsilon}\varphi_{F,\epsilon},\varphi_{F,\epsilon}\rangle_{\varrho_\epsilon^{p-r}} = \sigma_{2,\epsilon}h_2^2 + \sum_{j=3}^\infty \sigma_{j,\epsilon}h_j^2\,,$$

and hence

$$\sigma_{3,\epsilon}\sum_{j=3}^\infty h_j^2 \leq \sum_{j=3}^\infty \sigma_{j,\epsilon}h_j^2 \leq \Xi\epsilon^{q-\beta} - \sigma_{2,\epsilon}h_2^2\,.$$

Since $\varphi_{F,\epsilon}$ is normalized, it follows that $h_j^2 \leq 1$ for all $j\geq 1$ and

$$1 - h_2^2 = \sum_{j=3}^\infty h_j^2 \leq \frac{\Xi\epsilon^{q-\beta} - \sigma_{2,\epsilon}h_2^2}{\sigma_{3,\epsilon}} \leq \frac{\Xi\epsilon^{q-\beta}}{\Xi_1 + \Xi_2\epsilon^{q-\vartheta} + \Xi_3\epsilon^{\theta-q}}.$$ □

Now consider

$$\bar{\varphi}_{2,0}(x) := b_\epsilon^0\varrho_0^r(x)\left[\mathbf{1}_{\mathcal{Z}^+}(x) - \mathbf{1}_{\mathcal{Z}^-}(x)\right] \in L^2(\mathcal{Z},\varrho_\epsilon^{p-r})\,,$$

obtained by zero extension of $\varphi_{2,0}$ to all of $\mathcal{Z}$, where

$$b_\epsilon^0 := 1/\|\varrho_0^r(x)\left[\mathbf{1}_{\mathcal{Z}^+} - \mathbf{1}_{\mathcal{Z}^-}\right]\|_{L^2(\mathcal{Z},\varrho_\epsilon^{p-r})} \tag{6.19}$$

is a normalization constant. Similarly, we denote

$$\varphi_{F,\epsilon} = b_\epsilon^F \varrho_\epsilon^r(x)\left[\chi_\epsilon^+(x) - \chi_\epsilon^-(x)\right] \in L^2(\mathcal{Z},\varrho_\epsilon^{p-r})\,,$$

with the normalization constant

$$b_\epsilon^F := 1/\|\varrho_\epsilon^r\left[\chi_\epsilon^+ - \chi_\epsilon^-\right]\|_{L^2(\mathcal{Z},\varrho_\epsilon^{p-r})} > 0\,.$$

We begin by providing bounds on the normalization constants b_ϵ^0 and b_ϵ^F.

Lemma 6.10. *Let $(p,q,r) \in \mathbb{R}^3$ satisfying $p+r>0$, and suppose Assumptions 2.4, 2.5 and 2.6 hold. Let $\epsilon_0 > 0$ small enough. Then there exist constants $\Xi_1, \Xi_2 > 0$, independent of ϵ so that for all $\epsilon \in (0,\epsilon_0)$,*

$$\left| b_\epsilon^0 - \left(\int_{\mathcal{Z}'} \varrho_0^{p+r}\,dx \right)^{-1/2} \right| \le \Xi_1 \epsilon\,, \qquad \left| b_\epsilon^F - \left(\int_{\mathcal{Z}'} \varrho_0^{p+r}\,dx \right)^{-1/2} \right| \le \Xi_2 \epsilon^{\min\{1,p+r\}}\,.$$

Proof. Using the explicit expression (6.19) write

$$(b_\epsilon^0)^{-2} = \int_{\mathcal{Z}'} \varrho_0^{2r} \varrho_\epsilon^{p-r} dx.$$

It follows from Assumption 2.6(c) that

$$\varrho_0(x) - K_1\epsilon \le \varrho_\epsilon(x) \ \le \varrho_0(x) + K_1\epsilon \qquad \forall\, x \in \mathcal{Z}'\,. \tag{6.20}$$

Combining with Assumption 2.5(c), we can find a constant $\Xi_3 > 0$ so that

$$\left| (b_\epsilon^0)^{-2} - \int_{\mathcal{Z}'} \varrho_0^{p+r}\,dx \right| \le \Xi_3 \epsilon\,. \tag{6.21}$$

Let $b_\epsilon^\pm$ be as in (6.1). Using Assumption 2.6(d), and the definition of the $\chi_\epsilon^\pm$, we can write

$$\begin{aligned}
(b_\epsilon^F)^{-2} &= \int_{\mathcal{Z}'_{\epsilon_1}} \varrho_\epsilon^{p+r}(x)\left[(b_\epsilon^+)^2 \xi_\epsilon^+(x) + (b_\epsilon^-)^2 \xi_\epsilon^-(x)\right] dx \\
&= (b_\epsilon^+)^2 \int_{\mathcal{Z}_\epsilon^+} \varrho_\epsilon^{p+r}\,dx + (b_\epsilon^-)^2 \int_{\mathcal{Z}_\epsilon^-} \varrho_\epsilon^{p+r}\,dx \\
&\quad + K_2^{p+r}\epsilon^{p+r} \int_{\mathcal{Z}'_{\epsilon_1}\setminus\mathcal{Z}'_\epsilon} \left[\chi_\epsilon^+ - \chi_\epsilon^-\right]^2 dx \\
&= \int_{\mathcal{Z}'} \varrho_\epsilon^{p+r} dx \\
&\quad + ((b_\epsilon^+)^2 - 1) \int_{\mathcal{Z}^+} \varrho_\epsilon^{p+r} dx + ((b_\epsilon^-)^2 - 1) \int_{\mathcal{Z}^-} \varrho_\epsilon^{p+r} dx
\end{aligned}$$

$$+ (b_\epsilon^+)^2 \int_{\mathcal{Z}_\epsilon^+ \setminus \mathcal{Z}^+} \varrho_\epsilon^{p+r}\, dx + (b_\epsilon^-)^2 \int_{\mathcal{Z}_\epsilon^- \setminus \mathcal{Z}^-} \varrho_\epsilon^{p+r}\, dx$$
$$+ K_2^{p+r} \epsilon^{p+r} \int_{\mathcal{Z}'_{\epsilon_1} \setminus \mathcal{Z}'_\epsilon} \left[\chi_\epsilon^+ - \chi_\epsilon^-\right]^2 dx.$$

The first term is close to $\int_{\mathcal{Z}'} \varrho_0^{p+r} dx$ using (6.20), whereas the terms in the second line can be controlled using Lemma 6.3 and the fact that $0 < b_\epsilon^\pm < 2$,

$$\left| (b_\epsilon^\pm - 1)(b_\epsilon^\pm + 1) \int_{\mathcal{Z}^\pm} \varrho_\epsilon^{p+r} dx \right| \leq 3|b_\epsilon^\pm - 1| \left| \int_{\mathcal{Z}^\pm} \varrho_\epsilon^{p+r} dx \right| \leq \Xi_4 \epsilon^{\min\{1,p+r\}}.$$

Finally, the last two lines can be estimated using (2.8),

$$0 \leq (b_\epsilon^+)^2 \int_{\mathcal{Z}_\epsilon^+ \setminus \mathcal{Z}^+} \varrho_\epsilon^{p+r}\, dx + (b_\epsilon^-)^2 \int_{\mathcal{Z}_\epsilon^- \setminus \mathcal{Z}^-} \varrho_\epsilon^{p+r}\, dx$$
$$+ K_2^{p+r} \epsilon^{p+r} \int_{\mathcal{Z}'_{\epsilon_1} \setminus \mathcal{Z}'_\epsilon} \left[\chi_\epsilon^+ - \chi_\epsilon^-\right]^2 dx$$
$$\leq 4|\mathcal{Z}'_\epsilon \setminus \mathcal{Z}'| \left(\varrho_{\epsilon_0}^+\right)^{p+r} + 4|\mathcal{Z}'_{\epsilon_1} \setminus \mathcal{Z}'_\epsilon| K_2^{p+r} \epsilon^{p+r} \leq \Xi_5 \epsilon^{\min\{1,p+r+\beta\}}$$

for some $\Xi_5 > 0$. Putting the above estimates together, we obtain

$$\left| (b_\epsilon^F)^{-2} - \int_{\mathcal{Z}'} \varrho_0^{p+r}\, dx \right| \leq \Xi_6 \epsilon^{\min\{1,p+r\}}. \tag{6.22}$$

The lemma then follows from (6.21) and (6.22). □

In order to prove Theorem 3.4, we aim to derive an error bound on the difference between $\bar{\varphi}_{2,0}$ and $\varphi_{2,\epsilon}$. To this end, we first estimate $\langle \varphi_{F,\epsilon}, \bar{\varphi}_{2,0} \rangle_{\varrho_\epsilon^{p-r}}$ using the explicit expressions for $\varphi_{F,\epsilon}$ and $\bar{\varphi}_{2,0}$.

Proposition 6.11. *Let $(p, q, r) \in \mathbb{R}^3$ satisfying $p + r > 0$, and suppose Assumptions 2.4, 2.5 and 2.6 hold. Let $\epsilon_0 > 0$ small enough. Then there exists a constant $\Xi > 0$, independent of ϵ so that for all $\epsilon \in (0, \epsilon_0)$,*

$$\|\bar{\varphi}_{2,0} - \varphi_{F,\epsilon}\|^2_{L^2(\mathcal{Z}, \varrho_\epsilon^{p-r})} \leq \Xi \epsilon^{\min\{1,p+r\}}.$$

Proof. Since $\mathcal{Z}^+ \cap \mathcal{Z}_{\epsilon_1}^- = \emptyset$ and $\mathcal{Z}^- \cap \mathcal{Z}_{\epsilon_1}^+ = \emptyset$, we have

$$\langle \varrho_\epsilon^r(x) \left[\chi_\epsilon^+(x) - \chi_\epsilon^-(x)\right], \varrho_0^r \left[\mathbf{1}_{\mathcal{Z}^+}(x) - \mathbf{1}_{\mathcal{Z}^-}(x)\right] \rangle_{\varrho_\epsilon^{p-r}}$$
$$= \int_{\mathcal{Z}} \varrho_0^r(x) \varrho_\epsilon^p(x) \left[b_\epsilon^+ \xi_\epsilon^+(x) \mathbf{1}_{\mathcal{Z}^+}(x) - b_\epsilon^+ \xi_\epsilon^+(x) \mathbf{1}_{\mathcal{Z}^-}(x)\right] dx$$
$$- \int_{\mathcal{Z}} \varrho_0^r(x) \varrho_\epsilon^p(x) \left[b_\epsilon^- \xi_\epsilon^-(x) \mathbf{1}_{\mathcal{Z}^+}(x) - b_\epsilon^- \xi_\epsilon^-(x) \mathbf{1}_{\mathcal{Z}^-}(x)\right] dx$$
$$= b_\epsilon^+ \int_{\mathcal{Z}^+} \varrho_0^r \varrho_\epsilon^p dx + b_\epsilon^- \int_{\mathcal{Z}^-} \varrho_0^r \varrho_\epsilon^p dx$$

$$= \int_{\mathcal{Z}'} \varrho_0^r \varrho_\epsilon^p \, dx + (b_\epsilon^+ - 1) \int_{\mathcal{Z}^+} \varrho_0^r \varrho_\epsilon^p dx + (b_\epsilon^- - 1) \int_{\mathcal{Z}^-} \varrho_0^r \varrho_\epsilon^p dx \,.$$

If $p \geq 0$ (and by a similar argument with the order of inequalities reversed if $p < 0$), (6.20) implies

$$\varrho_0^p(x) - \epsilon K_1 p \varrho_0^{p-1}(x) + O(\epsilon^2) \leq \varrho_\epsilon^p(x) \leq \varrho_0^p(x) + \epsilon K_1 p \varrho_0^{p-1}(x) + O(\epsilon^2) \,.$$

By Assumption 2.5(c), we conclude that there exists a constant $\Xi_1 > 0$ such that

$$\left| \int_{\mathcal{Z}'} \varrho_0^r \varrho_\epsilon^p \, dx - \int_{\mathcal{Z}'} \varrho_0^{p+r} \, dx \right| \leq \Xi_1 \epsilon \,.$$

The above estimate together with Lemma 6.3 implies

$$\left| \int_{\mathcal{Z}'} \varrho_0^{p+r} \, dx - \left\langle \varrho_\epsilon^r(x) \left[\chi_\epsilon^+(x) - \chi_\epsilon^-(x) \right], \varrho_0^r \left[\mathbf{1}_{\mathcal{Z}^+}(x) - \mathbf{1}_{\mathcal{Z}^-}(x) \right] \right\rangle_{\varrho_\epsilon^{p-r}} \right| \leq \Xi_2 \epsilon^{\min\{1, p+r\}}$$

for some constant $\Xi_2 > 0$. Combining this bound with Lemma 6.10, and writing

$$\left\langle \varphi_{F,\epsilon}, \bar{\varphi}_{2,0} \right\rangle_{\varrho_\epsilon^{p-r}} = b_\epsilon^0 b_\epsilon^F \left\langle \varrho_\epsilon^r(x) \left[\chi_\epsilon^+(x) - \chi_\epsilon^-(x) \right], \varrho_0^r \left[\mathbf{1}_{\mathcal{Z}^+}(x) - \mathbf{1}_{\mathcal{Z}^-}(x) \right] \right\rangle_{\varrho_\epsilon^{p-r}} ,$$

we conclude that there exists a constant $\Xi_3 > 0$ so that

$$\left| 1 - \left\langle \varphi_{F,\epsilon}, \bar{\varphi}_{2,0} \right\rangle_{\varrho_\epsilon^{p-r}} \right| \leq \Xi_3 \epsilon^{\min\{1, p+r\}} \,.$$

Finally, we obtain

$$\begin{aligned} \left\| \bar{\varphi}_{2,0} - \varphi_{F,\epsilon} \right\|^2_{L^2(\mathcal{Z}, \varrho_\epsilon^{p-r})} &= \int_{\mathcal{Z}} \left| \bar{\varphi}_{2,0} - \varphi_{F,\epsilon} \right|^2 \varrho_\epsilon^{p-r} \, dx \\ &= \left\| \bar{\varphi}_{2,0} \right\|^2_{L^2(\mathcal{Z}, \varrho_\epsilon^{p-r})} + \left\| \varphi_{F,\epsilon} \right\|^2_{L^2(\mathcal{Z}, \varrho_\epsilon^{p-r})} - 2 \left\langle \varphi_{F,\epsilon}, \bar{\varphi}_{2,0} \right\rangle_{\varrho_\epsilon^{p-r}} \\ &= 2 \left(1 - \left\langle \varphi_{F,\epsilon}, \bar{\varphi}_{2,0} \right\rangle_{\varrho_\epsilon^{p-r}} \right) \leq \Xi \epsilon^{\min\{1, p+r\}} \,. \quad \square \end{aligned}$$

We are now ready to provide a quantitative estimate on how close the perturbed second eigenfunction $\varphi_{2,\epsilon}$ is to $\bar{\varphi}_{2,0}$ by comparing both eigenfunctions to the approximate Fiedler vector $\varphi_{F,\epsilon}$.

Proof of Theorem 3.4. We apply Proposition 6.9 with the eigenvalue bounds in Theorem 3.4(ii, iii). Depending on (p, q, r), we have different lower bounds on $\sigma_{3,\epsilon}$. Writing the bounds from Theorem 3.4 in the notation of Proposition 6.9, we have

- If $q > p + r$, then $\Xi_1 = 0$, $\Xi_2 > 0$, $\Xi_3 = 0$ and $\vartheta = -q + 2(p + r)$;
- If $q = p + r$, then $\Xi_1 > 0$, $\Xi_2 = \Xi_3 = 0$;
- If $q < p + r$, then $\Xi_1 = \Xi_2 = 0$, $\Xi_3 > 0$, and $\theta = p + r$.

We obtain that there exists a constant $\Xi_4 > 0$ so that

$$\left| 1 - \left\langle \varphi_{2,\epsilon}, \varphi_{F,\epsilon} \right\rangle^2_{\varrho_\epsilon^{p-r}} \right| \leq \Xi_4 \epsilon^{-|q-p-r|+\min\{q, p+r\}-\beta}, \tag{6.23}$$

for any $(p, q, r) \in \mathbb{R}^3$ with $q > 0$ and $p + r > 0$. Combining estimate (6.23) with Proposition 6.11 gives

$$\begin{aligned}
&\left|1 - \langle \varphi_{2,\epsilon}, \bar{\varphi}_{2,0} \rangle^2_{\varrho_\epsilon^{p-r}}\right| \\
&\quad = \left|1 - \left(\langle \varphi_{2,\epsilon}, \varphi_{F,\epsilon} \rangle_{\varrho_\epsilon^{p-r}} + \langle \varphi_{2,\epsilon}, \bar{\varphi}_{2,0} - \varphi_{F,\epsilon} \rangle_{\varrho_\epsilon^{p-r}}\right)^2\right| \\
&\quad \leq \left|1 - \langle \varphi_{2,\epsilon}, \varphi_{F,\epsilon} \rangle^2_{\varrho_\epsilon^{p-r}}\right| + \left|\langle \varphi_{2,\epsilon}, \bar{\varphi}_{2,0} - \varphi_{F,\epsilon} \rangle_{\varrho_\epsilon^{p-r}}\right| \left|\langle \varphi_{2,\epsilon}, \bar{\varphi}_{2,0} + \varphi_{F,\epsilon} \rangle_{\varrho_\epsilon^{p-r}}\right| \\
&\quad \leq \left|1 - \langle \varphi_{2,\epsilon}, \varphi_{F,\epsilon} \rangle^2_{\varrho_\epsilon^{p-r}}\right| \\
&\qquad + \|\varphi_{2,\epsilon}\|^2_{L^2(\mathcal{Z}, \varrho_\epsilon^{p-r})} \|\bar{\varphi}_{2,0} - \varphi_{F,\epsilon}\|_{L^2(\mathcal{Z}, \varrho_\epsilon^{p-r})} \left(\|\bar{\varphi}_{2,0}\|_{L^2(\mathcal{Z}, \varrho_\epsilon^{p-r})} + \|\varphi_{F,\epsilon}\|_{L^2(\mathcal{Z}, \varrho_\epsilon^{p-r})}\right) \\
&\leq \Xi_4 \epsilon^{-|q-p-r|+\min\{q,p+r\}-\beta} + \Xi_5 \epsilon^{\frac{1}{2}\min\{1,p+r\}} \\
&\leq \Xi \epsilon^{\min\{\frac{1}{2}, \frac{p+r}{2}, q-2(q-p-r)-\beta, q-\beta, 2q-(p+r)-\beta\}}
\end{aligned}$$

for some $\Xi > 0$ since $\|\varphi_{2,\epsilon}\|_{L^2(\mathcal{Z}, \varrho_\epsilon^{p-r})} = \|\bar{\varphi}_{2,0}\|_{L^2(\mathcal{Z}, \varrho_\epsilon^{p-r})} = \|\varphi_{F,\epsilon}\|_{L^2(\mathcal{Z}, \varrho_\epsilon^{p-r})} = 1$. □

7. Conclusions

We have studied a three-parameter family of weighted elliptic differential operators, motivated by spectral clustering and semi-supervised learning problems in the analysis of large data sets.

We analyzed the perturbative properties of the family (1.1) of elliptic operators $\mathcal{L}$, characterizing the sensitive dependence of its low-lying spectrum with respect to the parameters p, q, r in cases where the density ϱ concentrates on two clusters. In particular, the theory suggests that there is a major change in the behavior of the spectrum of $\mathcal{L}$ when $q = p + r$ versus $q \neq p + r$. In the former regime, $\mathcal{L}$ has a uniform spectral gap between the third and second eigenvalues indicating that two clusters are present in ϱ, while in the latter regime only a spectral ratio gap may manifest.

In addition, we provided numerical evidence that exemplified and extended our analysis. Most notably, our numerics show that our bounds on the second eigenvalue are sharp and that a uniform spectral gap exists between the third and second eigenvalues of $\mathcal{L}$ when $q \leq p + r$, whereas only a ratio spectral gap is present when $q > p + r$. Therefore, in the $q > p + r$ and $q < p + r$ regimes, comparing with our theoretical predictions, our numerics indicate that our lower bounds on the third eigenvalues, and hence on the spectral ratio gap, can be sharpened. The question of spectral gaps is of interest from a practical point of view as the low-lying spectral properties govern many unsupervised and semi-supervised clustering tasks.

Further, we demonstrated a rigorous connection between the geometry of the low-lying eigenfunctions of $\mathcal{L}$ and the geometry of the density ϱ. We showed that as ϱ concentrates on two clusters, the span of the first two eigenfunctions of $\mathcal{L}$ approaches certain weighted set functions on the clusters.

In fact, the family of operators $\mathcal{L}$ arises naturally as continuum limits of graph Laplacians L_N of the form (1.2). We provided a roadmap for rigorous proof of convergence of L_N to $\mathcal{L}$ as $N \to \infty$ in the framework of [23], but for the more general family of any parameter choices (p, q, r); the full proof is the subject of future research. To support this analysis, we presented numerical evidence in the discrete graphical settings showing the manifestation of our continuum spectral analysis on discrete graph Laplacians that are weighted appropriately with respect to the continuum limits, and this can be observed even in the case of more general data densities ϱ than our theory provides for.

Finally, we provided numerical evidence that extends our analysis from the binary cluster case to three or five clusters, showing strong evidence that similar results can be proven in the setting where ϱ concentrates on any number of finitely many clusters.

Our work may be of independent interest within the spectral theory of elliptic operators. Furthermore it can be used to build on the paper [27], which studies consistency of semi-supervised learning on graphs, to develop a consistency theory for semi-supervised learning in the continuum limit.

The primary practical take-home message of this work is that graph Laplacian normalizations that satisfy $q = p + r$ are likely to be preferable in practical applications such as spectral clustering or semi-supervised learning/regression. The primary theoretical take-home message is that there remain a variety of interesting problems in analysis in this area, concerning the spectral properties of density-dependent elliptic operators with high-contrast coefficients, and their relationship to underlying graph Laplacian problems.

Acknowledgments

The authors are grateful to Nicolás García Trillos for helpful discussions regarding the results in Section 5 concerning various graph Laplacians and their continuum limits. We are also thankful to the anonymous reviewers whose comments and suggestions helped us improve an earlier version of this article. AMS is grateful to AFOSR (grant FA9550-17-1-0185) and NSF (grant DMS 18189770) for financial support. FH was partially supported by Caltech's von Kármán postdoctoral instructorship. BH was partially supported by an NSERC PDF fellowship.

Appendix A. Diffusion maps and weighted graph Laplacians

We note from Remark 2.3 that when $p = q$ and $r = 0$ the limiting graph Laplacian $\mathcal{L}$ is the generator of a reversible diffusion process with invariant density proportional to ϱ^q. The connection between the graph Laplacian L_N in (1.2) and diffusions was first established in the celebrated paper [14] by Coifman and Lafon, through the diffusion maps introduced therein. In this appendix we further elucidate these connections.

We fix a probability density $\varrho \in L^1(\Omega)$ for any set $\Omega \subset \mathbb{R}^d$ and introduce the following functions for $x, y \in \Omega$:

$$\tilde{W}(x, y) = \eta_\delta(|x - y|)$$

where η is a rotation-invariant normalized kernel, $\int_\Omega \eta_\delta(|x|)\, dx = 1$, with a fixed scale parameter δ, and with associated degree function

$$\tilde{d}(x) = \int_\Omega \tilde{W}(x, y)\varrho(y)\, dy\,.$$

Note that $\tilde{d}(x)$ approximates $\varrho(x)$ as η_δ converges weakly to the Dirac delta distribution. We suppress the dependence of $\tilde{d}$ and $\tilde{W}$ on δ for brevity. Given a parameter $\alpha \in \mathbb{R}$, we now construct the weighted kernel

$$W(x, y) = \frac{\tilde{W}(x, y)}{\tilde{d}(x)^\alpha \tilde{d}(y)^\alpha}$$

with associated degree function

$$d(x) = \int_\Omega W(x, y)\varrho(y)\, dy\,.$$

The kernel W gives rise to an integral operator $\mathcal{K} : L^1(\Omega) \to L^1(\Omega)$,

$$\mathcal{K}f(x) = \int_\Omega W(x,y) f(y) \varrho(y)\, dy\,.$$

Then $d(x) = \mathcal{K}\mathbf{1}_\Omega(x)$. Normalizing $\mathcal{K}$ gives a Markov operator $\mathcal{P} : L^1(\Omega) \to L^1(\Omega)$,

$$\mathcal{P}f(x) := \frac{1}{\mathcal{K}\mathbf{1}_\Omega(x)} \mathcal{K}f(x) = \int_\Omega p(x,y) f(y) \varrho(y)\, dy$$

with anisotropic Markov transition kernel

$$p(x,y) = \frac{W(x,y)}{d(x)}\,.$$

Observe that $\mathcal{P}\mathbf{1}_\Omega = \mathbf{1}_\Omega$, and so $\mathcal{P}$ leaves constants unchanged.

Discrete setting. Given N samples $x_j \sim \varrho$, we define analogously to the above the matrix $\tilde{W}_N$ with entries

$$\tilde{W}_{ij} = \tilde{W}(x_i, x_j)$$

with associated degree matrix $\tilde{D}_N$,

$$\tilde{D}_{ij} = diag\left(\tilde{d}_i\right)\,, \qquad \tilde{d}_i = \sum_{k=1}^N \tilde{W}_{ik}\,.$$

From the above, we construct the weighted similarity matrix W_N with entries

$$W_{ij} = \frac{\tilde{W}_{ij}}{\tilde{d}_i^\alpha \tilde{d}_j^\alpha}$$

with associated degree matrix D_N,

$$D_{ij} = diag\left(d_i\right)\,, \qquad d_i = \sum_{k=1}^N W_{ik}\,.$$

To make the connection between this discrete setting and the continuous analogue above, we use the degree functions of Subsection 5.2,

$$\tilde{d}^N(x) = \frac{1}{N}\sum_{j=1}^N \tilde{W}(x, x_j)$$

$$d^N(x) = \frac{1}{N}\sum_{j=1}^N W(x, x_j) = \frac{1}{N}\sum_{j=1}^N \frac{\tilde{W}(x,x_j)}{\left(\tilde{d}^N(x)\right)^\alpha \left(\tilde{d}^N(x_j)\right)^\alpha}\,.$$

They correspond exactly to $d(x)$ and $\tilde{d}(x)$ with ϱ substituted by the empirical density $\mu_N := \frac{1}{N}\sum_{i=1}^N \delta_{x_i}$. Then

$$\tilde{d}_i = N\tilde{d}^N(x_i)\,, \qquad d_i = N^{1-2\alpha} d^N(x_i)\,, \qquad W_{ij} = \frac{1}{N^{2\alpha}} W(x_i, x_j)\,,$$

and so $\tilde{d}_i/N$ approximates $\varrho(x_i)$ as η_δ converges to the Dirac delta distribution for large N. Finally, the operators $\mathcal{K}$ and $\mathcal{P}$ are approximated empirically by matrices W_N/N and P_N, where P_N has entries

$$P_{ij} = \frac{W(x_i, x_j)}{N d^N(x_i)} = \frac{N^{2\alpha} W_{ij}}{N^{2\alpha} d_i},$$

and so

$$P_N = D_N^{-1} W_N .$$

In [14], the graph Laplacian matrix $\bar{L}_N$ is defined as

$$\bar{L}_N = \frac{I_N - P_N}{\delta} = \frac{1}{\delta} D_N^{-1} (D_N - W_N) = \frac{1}{\delta} L_N ,$$

where I_N denotes the identity matrix, and L_N is our graph Laplacian matrix as defined in (1.2) with $p = q = 2(1-\alpha)$ and $r = 0$. Note that $\bar{L}_N$ is not symmetric.

Generator of a diffusion semi-group. Taking $\delta \to 0$, we see that

$$\tilde{W}(x, y) \to \delta_{x=y} ,$$
$$\tilde{d}(x) \to \varrho(x) , \qquad d(x) = \mathcal{K}\mathbf{1}_\Omega(x) \to \varrho(x)^{1-2\alpha} ,$$

and so $\mathcal{P}$ converges to the identity operator Id. Defining the operator

$$\mathcal{G} = \frac{\mathrm{Id} - \mathcal{P}}{\delta}$$

analogously to the discrete setting, it was shown in [14, Thm. 2] that

$$\lim_{\delta \to 0} \mathcal{G} f = -\mathcal{L} f$$

for f in any finite span of the eigenfunctions of the Laplace-Beltrami operator on a compact submanifold of Ω. Here, $\mathcal{G}$ is the infinitesimal generator of a Markov chain, and $\mathcal{L}$ is the weighted elliptic operator defined in (1.1) for the parameter choices $p = q = 2(1-\alpha)$ and $r = 0$. In this sense, the operator $\mathcal{P}$ is an approximation to the semi-group

$$e^{\delta \mathcal{L}} = \mathrm{Id} + \delta \mathcal{L} + \mathcal{O}(\delta^2)$$

associated with the infinitesimal generator $\mathcal{L}$,

$$\begin{aligned} -\mathcal{L} f &= \frac{1}{\varrho^{2(1-\alpha)}} \nabla \cdot \left(\varrho^{2(1-\alpha)} \nabla f \right) \\ &= \Delta f + 2(1-\alpha) \varrho^{-1} \nabla \varrho \cdot \nabla f \\ &= \Delta f + \nabla \log \left(\varrho^{2(1-\alpha)} \right) \cdot \nabla f . \end{aligned}$$

More precisely, the operator $\mathcal{L}$ is the infinitesimal generator of the reversible diffusion process

$$dX_t = -\nabla \Psi(X_t) dt + \sqrt{2}\, dB ,$$

where B denotes a Brownian motion in $\mathbb{R}^d$ with associated potential

$$\Psi(x) = -\log \left(\varrho(x)^{2(1-\alpha)} \right)$$

and invariant measure proportional to $\varrho^{2(1-\alpha)}$ satisfying $\mathcal{L}^* e^{-\Psi} = \mathcal{L}^* \rho^{2(1-\alpha)} = 0$. In this sense, the discrete graph Laplacian matrix $\bar{L}_N$ introduced above serves as an approximation of the generator $-\mathcal{L}$.

In [14], Coifman and Lafon discuss the cases (i) $\alpha = 0$ ($q = 2$) when the graph Laplacian has isotropic weights and $W = \tilde{W}$, (ii) $\alpha = 1/2$ ($q = 1$) when the Dirichlet energy of $\mathcal{L}$ is linear in ϱ, and (iii) $\alpha = 1$ ($q = 0$), when $-\mathcal{L}f = \Delta f$, and so the Markov chain corresponding to $\mathcal{G}$ converges (as $\delta \to 0$) to the Brownian motion in Ω with reflecting boundary conditions.

There is a well-known connection between the generator of reversible diffusion processes and Schrödinger operators [35]. Following the above connections between limiting graph Laplacians and generators of diffusion processes with invariant measures proportional to $\varrho^{(1-2\alpha)}$, we connect the operator $\mathcal{L}$ to certain Schrödinger operators as follows. Define

$$\mathcal{S}u := \Delta u - u \frac{\Delta\left(\varrho^{1-\alpha}\right)}{\varrho^{1-\alpha}},$$

then we can write for $u = f\varrho^{1-\alpha}$,

$$-\mathcal{L}f = \frac{\Delta\left(f\varrho^{1-\alpha}\right)}{\varrho^{1-\alpha}} - \frac{\Delta\left(\varrho^{1-\alpha}\right)}{\varrho^{1-\alpha}} f = \frac{\mathcal{S}u}{\varrho^{1-\alpha}}.$$

Appendix B. Function spaces

Throughout this section ϱ is taken to be a smooth probability density function with full support on a bounded open set $\Omega \subset \mathbb{R}^d$ with C^1 boundary which is bounded from above and below by positive constants as in (2.1), i.e.,

$$0 < \varrho^- \leq \varrho(x) \leq \varrho^+ < +\infty, \qquad \forall x \in \bar{\Omega}. \tag{B.1}$$

Our first task is to establish the equivalence between regular $L^p(\Omega)$ spaces and the weighted spaces $L^p(\Omega, \varrho)$. In fact, a straightforward calculation using (B.1) implies the following lemma.

Lemma B.1. *Let ϱ be a smooth probability density function on Ω satisfying (B.1) and let $u \in L^p(\Omega)$ for $p \geq 0$. Then*

$$\varrho^- \|u\|^p_{L^p(\Omega)} \leq \|u\|^p_{L^p(\Omega,\varrho)} \leq \varrho^+ \|u\|^p_{L^p(\Omega)},$$

i.e., $L^p(\Omega) = L^p(\Omega, \varrho)$.

Given constants $(p, q, r) \in \mathbb{R}^3$ we consider the weighted Sobolev spaces $H^1(\Omega, \varrho)$ introduced in section 2.1. We now have:

Lemma B.2. *Let $\varrho \in C^\infty(\bar{\Omega})$ be a smooth probability density function satisfying (B.1) and let $u \in H^1(\Omega, \varrho)$ with parameters $(p, q, r) \in \mathbb{R}^3$. Then there exist constants $C^\pm(q, \varrho^\pm) > 0$ so that*

$$C^- \left\| \frac{u}{\varrho^r} \right\|^2_{H^1(\Omega)} \leq \|u\|^2_{H^1(\Omega,\varrho)} \leq C^+ \left\| \frac{u}{\varrho^r} \right\|^2_{H^1(\Omega)}.$$

Proof. Since ϱ satisfies (B.1) then

$$(\varrho^-)^q \left| \nabla \left(\frac{u}{\varrho^r} \right) \right|^2 dx \leq \int_\Omega \varrho^q \left| \nabla \left(\frac{u}{\varrho^r} \right) \right|^2 dx \leq (\varrho^+)^q \int_\Omega \left| \nabla \left(\frac{u}{\varrho^r} \right) \right|^2 dx.$$

Then the desired result follows immediately by Lemma B.1 applied to L^2 norms. □

With the equivalence between the weighted and regular L^p and H^1 spaces established. We can present the following compact embedding as a consequence of the Rellich-Kondrachov Theorem [19, Ch. 5.7, Thm 1]:

Proposition B.3. *Let $\varrho \in C^\infty(\bar{\Omega})$ be a probability density function satisfying (B.1) and fix $(p,q,r) \in \mathbb{R}^3$. Then $H^1(\Omega,\varrho)$ is compactly embedded in $L^2(\Omega,\varrho^{p-r})$.*

Appendix C. Min-max principle

The min-max principle [29, Ch. 1 Sec. 6.10] is readily applied to our specific setting to obtain the following:

Proposition C.1. *Fix $(p,q,r) \in \mathbb{R}^3$. For any open bounded set $\Omega \subset \mathbb{R}^d$ with $\partial\Omega \in C^{1,1}$, and for a given density $\varrho \in C^\infty(\overline{\Omega})$ satisfying Assumption 2.5, let $\sigma_1 \le \sigma_2 \le \ldots \le \sigma_j \le \ldots$ be the sequence of eigenvalues of the Neumann operator*

$$\mathcal{L} = -\frac{1}{\varrho^p}\nabla\cdot\left(\varrho^q\nabla\left(\frac{\cdot}{\varrho^r}\right)\right)$$

in $V^1(\Omega,\varrho)$, repeated in accordance with their multiplicities, and let $\{\varphi_j\}_{j\in\mathbb{N}}$ be a corresponding Hilbertian basis of eigenvectors in $V^1(\Omega,\varrho)$; then

$$\left\langle \varrho^q\nabla\left(\frac{\varphi_j}{\varrho^r}\right), \nabla\left(\frac{v}{\varrho^r}\right)\right\rangle = \sigma_j\left\langle \varrho^{p-r}\varphi_j, v\right\rangle, \qquad \varphi_j, v \in V^1(\Omega,\varrho).$$

Define the Rayleigh quotient of $\mathcal{L}$ by

$$\mathcal{R}(u) := \frac{\langle \mathcal{L}u, u\rangle_{\varrho^{p-r}}}{\langle u,u\rangle_{\varrho^{p-r}}} = \frac{\int_\Omega \left|\nabla\left(\frac{u}{\varrho^r}\right)\right|^2\varrho^q dx}{\int_\Omega |u|^2\varrho^{p-r}dx}, \qquad u \in V^1(\Omega,\varrho)\,.$$

Denote by $\mathcal{S}_n$ the class of all n-dimensional linear subspaces in $V^1(\Omega,\varrho)$, and by $M^\perp$ the orthogonal subspace of M in $V^1(\Omega,\varrho)$. Then we have

$$\sigma_n = \min_{M\in\mathcal{S}_n}\max_{v\in M, v\neq 0}\mathcal{R}(v) \tag{C.1}$$

$$= \max_{M\in\mathcal{S}_{n-1}}\min_{v\in M^\perp, v\neq 0}\mathcal{R}(v)\,. \tag{C.2}$$

Appendix D. Weighted Cheeger's inequality

Given positive measures μ, ν on $\Omega' \subset \mathbb{R}^d$, define the isoperimetric function $\mathcal{J}$ for any subset $\Omega \subset \Omega'$ by

$$\mathcal{J}(\Omega,\mu,\nu) := \frac{|\partial\Omega|_\mu}{\min\{|\Omega|_\nu, |\Omega'\setminus\Omega|_\nu\}}.$$

Here, we use the notation

$$|\Omega|_\nu := \nu(\Omega)\,,$$

and define the μ-weighted Minkowski boundary measure of Ω by

$$|\partial\Omega|_\mu := \liminf_{\delta\downarrow 0} \frac{1}{\delta}\left[|\Omega_\delta|_\mu - |\Omega|_\mu\right],$$

with Ω_δ as defined in (2.7),

$$\Omega_\delta := \{x : \operatorname{dist}(x,\Omega) \le \delta\}.$$

We show the following weighted version of Cheeger's inequality.

Proposition D.1 *(Weighted Cheeger's inequality). Let μ, ν be absolutely continuous measures with respect to the Lebesgue measure with C^∞ densities that are uniformly bounded above and below with positive constants on Ω'. Suppose there exists a constant $h > 0$ so that*

$$h \le \inf_{\Omega} \mathcal{J}(\Omega,\mu,\nu), \tag{D.1}$$

where the infimum is over open subsets $\Omega \subset \Omega'$ such that $|\Omega|_\nu \le \frac{1}{2}|\Omega'|_\nu$. Then the following Poincaré inequality holds:

$$\left(\sup_{x\in\Omega'}\left|\frac{d\mu}{d\nu}(x)\right|\right)^{-1} \frac{h^2}{4}\int_{\Omega'} |f - \bar{f}_{\Omega'}|^2 d\nu \le \int_{\Omega'} |\nabla f|^2 \, d\mu,$$

where $\bar{f}_{\Omega'}$ denotes the average of f with respect to ν,

$$\bar{f}_{\Omega'} := \frac{\int_{\Omega'} f \, d\nu}{|\Omega'|_\nu}.$$

This is a generalization of the weighted Cheeger's inequality as here we may take different measures μ and ν, whereas $\mu = \nu$ in [2]. The proof can readily be generalized from [2, Prop. 8.5.2] to this setting.

Proof. It follows from the co-area formula [2, Thm. 8.5.1] that for every Lipschitz function f on Ω',

$$\int_{-\infty}^{\infty} |\partial S(f,t)|_\mu dt \le \int_{\Omega'} |\nabla f| \, d\mu, \tag{D.2}$$

where $S(f,t) := \{x \in \Omega' : f(x) > t\}$ for $t \in \mathbb{R}$. Now let g be a positive Lipschitz function on Ω such that $|S(g,t)|_\nu \le \frac{1}{2}|\Omega'|_\nu$. Then by the hypothesis (D.1) we have for $t \ge 0$,

$$h \min\{|S(g,t)|_\nu, |\Omega' \setminus S(g,t)|_\nu\} \le |\partial S(g,t)|_\mu,$$

which together with (D.2) gives

$$h\int_0^{\infty} \min\{|S(g,t)|_\nu, |\Omega' \setminus S(g,t)|_\nu\} dt \le \int_{\Omega'} |\nabla g| \, d\mu. \tag{D.3}$$

Now let $f : \Omega' \to \mathbb{R}$ be Lipschitz and denote by m a median of f with respect to ν, i.e., $m \in \mathbb{R}$ such that

$$|\{x \in \Omega' : f(x) \ge m\}|_\nu \le \frac{1}{2}|\Omega'|_\nu, \qquad \text{and} \qquad |\{x \in \Omega' : f(x) \le m\}|_\nu \le \frac{1}{2}|\Omega'|_\nu.$$

Proceeding in the same way as in proof of [2, Prop. 8.5.2] we define $F_+ = \max\{f-m,0\}$ and $F_- = \max\{m-f,0\}$ and by definition of the median we have for $t>0$,

$$|S(F_+^2,t)|_\nu \le \frac{1}{2}|\Omega'|_\nu, \qquad \text{and} \qquad |S(F_-^2,t)|_\nu \le \frac{1}{2}|\Omega'|_\nu.$$

Applying (D.3) with $g=F_+^2$ and $g=F_-^2$ and adding the two inequalities yields

$$\begin{aligned} h\int_{\Omega'} |f-m|^2\,d\nu &= h\int_{\Omega'} F_+^2\,d\nu + h\int_{\Omega'} F_-^2\,d\nu \\ &= h\int_0^\infty |S(F_+^2,t)|_\nu\,dt + h\int_0^\infty |S(F_-^2,t)|_\nu\,dt \\ &\le \int_{\Omega'} |\nabla(F_+^2)|\,d\mu + \int_{\Omega'} |\nabla(F_-^2)|\,d\mu\,. \end{aligned}$$

By the Cauchy-Schwartz inequality,

$$\begin{aligned} \int_{\Omega'} |\nabla(F_\pm^2)|\,d\mu &= 2\int_{\Omega'} F_\pm|\nabla F_\pm|\,d\mu \le 2\left(\int_{\Omega'} |F_\pm|^2\,d\mu\right)^{1/2}\left(\int_{\Omega'} |\nabla F_\pm|^2\,d\mu\right)^{1/2} \\ &\le 2\left(\int_{\Omega'} |f-m|^2\,d\mu\right)^{1/2}\left(\int_{\Omega'} |\nabla F_\pm|^2\,d\mu\right)^{1/2} \\ &\le 2\left(\sup_{x\in\Omega'}\left|\frac{d\mu}{d\nu}(x)\right|\right)^{1/2}\left(\int_{\Omega'} |f-m|^2\,d\nu\right)^{1/2}\left(\int_{\Omega'} |\nabla F_\pm|^2\,d\mu\right)^{1/2}. \end{aligned}$$

The previous estimate with the fact that $F_\pm$ have disjoint support, gives

$$\left(\sup_{x\in\Omega'}\left|\frac{d\mu}{d\nu}(x)\right|\right)^{-1}\frac{h^2}{4}\int_{\Omega'} |f-m|^2\,d\nu \le \int_{\Omega'} |\nabla f|^2\,d\mu\,,$$

for any median of f. Finally, minimizing the left-hand side over m gives the desired lower bound with $m=\bar f_{\Omega'}$, which concludes the proof. □

Proof of Lemma 6.6. Apply Theorem D.1 with $d\mu(x)=\varrho_\epsilon^q(x)dx$ and $d\nu(x)=\varrho_\epsilon^{p+r}(x)dx$. Setting $u=f\varrho_\epsilon^r$ yields

$$\left(\sup_{x\in\Omega'}\varrho_\epsilon^{q-p-r}\right)^{-1}\frac{h^2}{4}\int_{\Omega'} |u-\bar u_{\Omega'}\varrho_\epsilon^r|^2\,\varrho_\epsilon^{p-r}dx \le \int_{\Omega'}\left|\nabla\left(\frac{u}{\varrho_\epsilon^r}\right)\right|^2\varrho_\epsilon^q dx,$$

which concludes the proof for Lipschitz functions u. The desired result on $V^1(\Omega',\varrho_\epsilon)$ then follows by a density argument, and noting that $\bar u=0$ in that case. □

References

[1] R.A. Adams, J.J. Fournier, Sobolev Spaces, vol. 140, Elsevier, 2003.
[2] D. Bakry, I. Gentil, M. Ledoux, Analysis and Geometry of Markov Diffusion Operators, vol. 348, Springer Science & Business Media, New York, 2013.
[3] S. Balay, S. Abhyankar, M.F. Adams, J. Brown, P. Brune, K. Buschelman, L. Dalcin, A. Dener, V. Eijkhout, W.D. Gropp, D. Karpeyev, D. Kaushik, M.G. Knepley, D.A. May, L.C. McInnes, R.T. Mills, T. Munson, K. Rupp, P. Sanan, B.F. Smith, S. Zampini, H. Zhang, H. Zhang, PETSc users manual. Technical Report ANL-95/11 - Revision 3.11, Argonne National Laboratory, 2019.
[4] M. Belkin, P. Niyogi, Laplacian eigenmaps for dimensionality reduction and data representation, Neural Comput. 15 (6) (2003) 1373–1396.
[5] M. Belkin, P. Niyogi, Convergence of Laplacian eigenmaps, in: NIPS, 2006.
[6] M. Belkin, P. Niyogi, Towards a theoretical foundation for Laplacian-based manifold methods, J. Comput. Syst. Sci. 74 (8) (2008) 1289–1308.
[7] T. Berry, J. Harlim, Variable bandwidth diffusion kernels, Appl. Comput. Harmon. Anal. 40 (1) (2016) 68–96.
[8] A.L. Bertozzi, A. Flenner, Diffuse interface models on graphs for classification of high dimensional data, Multiscale Model. Simul. 10 (3) (2012) 1090–1118.
[9] A.L. Bertozzi, X. Luo, A.M. Stuart, K.C. Zygalakis, Uncertainty quantification in graph-based classification of high dimensional data, SIAM/ASA J. Uncertain. Quantificat. 6 (2) (2018) 568–595.
[10] A. Bovier, M. Eckhoff, V. Gayrard, M. Klein, Metastability in reversible diffusion processes i: sharp asymptotics for capacities and exit times, J. Eur. Math. Soc. 6 (4) (2004) 399–424.
[11] A. Bovier, V. Gayrard, M. Klein, Metastability in reversible diffusion processes ii: precise asymptotics for small eigenvalues, J. Eur. Math. Soc. 7 (1) (2005) 69–99.
[12] J. Calder, The game theoretic p-Laplacian and semi-supervised learning with few labels, Nonlinearity 32 (1) (2018) 301.
[13] J. Calder, N.G. Trillos, Improved spectral convergence rates for graph Laplacians on ϵ-graphs and k-NN graphs, arXiv: 1910.13476, 2019.
[14] R.R. Coifman, S. Lafon, Diffusion maps, Appl. Comput. Harmon. Anal. 21 (1) (2006) 5–30.
[15] H.-L. de Kergorlay, D.J. Higham, Consistency of anchor-based spectral clustering, arXiv preprint, arXiv:2006.13984, 2020.
[16] P. Deuflhard, M. Dellnitz, O. Junge, C. Schütte, Computation of essential molecular dynamics by subdivision techniques, in: Computational Molecular Dynamics: Challenges, Methods, Ideas, Springer, 1999, pp. 98–115.
[17] P. Deuflhard, W. Huisinga, A. Fischer, C. Schütte, Identification of almost invariant aggregates in reversible nearly uncoupled Markov chains, Linear Algebra Appl. 315 (1–3) (2000) 39–59.
[18] M.M. Dunlop, D. Slepčev, A.M. Stuart, M. Thorpe, Large data and zero noise limits of graph-based semi-supervised learning algorithms, Appl. Comput. Harmon. Anal. (2019).
[19] L.C. Evans, Partial Differential Equations, second edition, Graduate Studies in Mathematics, vol. 19, AMS, Providence, RI, 2010.
[20] N. García Trillos, M. Gerlach, M. Hein, D. Slepčev, Error estimates for spectral convergence of the graph Laplacian on random geometric graphs toward the Laplace–Beltrami operator, Found. Comput. Math. 20 (4) (2020) 827–887.
[21] N. García Trillos, F. Hoffmann, B. Hosseini, Geometric structure of graph Laplacian embeddings, J. Mach. Learn. Res. 22 (2021) 63, 55 pp.
[22] N. Garcia Trillos, D. Slepčev, Continuum limit of total variation on point clouds, Arch. Ration. Mech. Anal. 220 (1) (2016) 193–241.
[23] N. García Trillos, D. Slepčev, A variational approach to the consistency of spectral clustering, Appl. Comput. Harmon. Anal. 45 (2) (2018) 239–281.
[24] E. Giné, V. Koltchinskii, et al., Empirical graph Laplacian approximation of Laplace–Beltrami operators: large sample results, in: High Dimensional Probability, Institute of Mathematical Statistics, 2006, pp. 238–259.
[25] Y. Hafiene, J.M. Fadili, C. Chesneau, A. Elmoataz, Continuum limit of the nonlocal p-Laplacian evolution problem on random inhomogeneous graphs, ESAIM: Math. Model. Numer. Anal. 54 (2) (2020).
[26] Y. Hafiene, J.M. Fadili, A. Elmoataz, Continuum limits of nonlocal p-Laplacian variational problems on graphs, SIAM J. Imaging Sci. 12 (4) (2019) 1772–1807.
[27] F. Hoffmann, B. Hosseini, Z. Ren, A.M. Stuart, Consistency of semi-supervised learning algorithms on graphs: probit and one-hot methods, J. Mach. Learn. Res. 21 (2020) 186, 55 pp.
[28] W. Huisinga, S. Meyn, C. Schütte, Phase transitions and metastability in Markovian and molecular systems, Ann. Appl. Probab. 14 (1) (2004) 419–458.
[29] T. Kato, Perturbation Theory for Linear Operators, second edition, Classics in Mathematics, Springer, New York, 1995.
[30] D.O. Loftsgaarden, C.P. Quesenberry, et al., A nonparametric estimate of a multivariate density function, Ann. Math. Stat. 36 (3) (1965) 1049–1051.
[31] A. Logg, K.-A. Mardal, G. Wells, Automated Solution of Differential Equations by the Finite Element Method: The FEniCS Book, Lecture Notes in Computational Science and Engineering, vol. 84, Springer Science & Business Media, 2012.
[32] W. McLean, Strongly Elliptic Systems and Boundary Integral Equations, Cambridge University Press, Cambridge, 2000.
[33] A.Y. Ng, M.I. Jordan, Y. Weiss, On spectral clustering: analysis and an algorithm, in: Proceedings of the 14th International Conference on Neural Information Processing Systems: Natural and Synthetic.
[34] A.Y. Ng, M.I. Jordan, Y. Weiss, On spectral clustering: analysis and an algorithm, in: Advances in Neural Information Processing Systems, 2002, pp. 849–856.
[35] G.A. Pavliotis, Stochastic Processes and Applications: Diffusion Processes, the Fokker-Planck and Langevin Equations, Texts in Applied Mathematics, vol. 60, Springer, New York, 2014.

[36] G. Schiebinger, M.J. Wainwright, B. Yu, et al., The geometry of kernelized spectral clustering, Ann. Stat. 43 (2) (2015) 819–846.
[37] C. Schütte, W. Huisinga, P. Deuflhard, Transfer operator approach to conformational dynamics in biomolecular systems, in: F. Bernold (Ed.), Ergodic Theory, Analysis, and Efficient Simulation of Dynamical Systems, Springer, Berlin, 2001, pp. 191–223.
[38] J. Shi, J. Malik, Normalized cuts and image segmentation, IEEE Trans. Pattern Anal. Mach. Intell. 22 (8) (Aug. 2000) 888–905.
[39] T. Shi, M. Belkin, B. Yu, et al., Data spectroscopy: eigenspaces of convolution operators and clustering, Ann. Stat. 37 (6B) (2009) 3960–3984.
[40] D. Slepčev, M. Thorpe, Analysis of p-Laplacian regularization in semisupervised learning, SIAM J. Math. Anal. 51 (3) (2019) 2085–2120.
[41] D.A. Spielmat, S.-H. Teng, Spectral partitioning works: planar graphs and finite element meshes, in: Proceedings of 37th Conference on Foundations of Computer Science, IEEE, 1996, pp. 96–105.
[42] G.R. Terrell, D.W. Scott, Variable kernel density estimation, Ann. Stat. 20 (3) (1992) 1236–1265.
[43] U. von Luxburg, A tutorial on spectral clustering, Stat. Comput. 17 (4) (2007) 395–416.
[44] U. von Luxburg, M. Belkin, O. Bousquet, Consistency of spectral clustering, Ann. Stat. 36 (2) (2008) 555–586.
[45] C.L. Wormell, S. Reich, Spectral convergence of diffusion maps: improved error bounds and an alternative normalisation, 2020.
[46] L. Zelnik-Manor, P. Perona, Self-tuning spectral clustering, in: Advances in Neural Information Processing Systems, 2005, pp. 1601–1608.
[47] X. Zhu, Z. Ghahramani, J.D. Lafferty, Semi-supervised learning using Gaussian fields and harmonic functions, in: Proceedings of the 20th International Conference on Machine Learning, 2003, pp. 912–919.