SPECTRAL ANALYSIS OF WEIGHTED LAPLACIANS ARISING IN DATA CLUSTERING

FRANCA HOFFMANN , BAMDAD HOSSEINI , ASSAD A. OBERAI *, AND ANDREW M. STUART †

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 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 and extend the analysis; in particular the computations study situations with more than two clusters. 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.

Key words. Spectral clustering, graph Laplacian, large data limits, elliptic differential operators, perturbation analysis, spectral gap, differential geometry.

AMS subject classifications. 47A75, 62H30, 68T10, 35B20, 05C50

1. Introduction. This article presents a spectral analysis of differential operators of the form

$$
(1.1) \quad
\begin{cases}
\mathcal{L}u := -\dfrac{1}{\varrho^p}\operatorname{div}\left(\varrho^q \nabla\left(\dfrac{u}{\varrho^r}\right)\right), & \text{in } \Omega, \\[2mm]
\varrho^q \dfrac{\partial}{\partial n}\left(\dfrac{u}{\varrho^r}\right) = 0, & \text{on } \partial\Omega,
\end{cases}
$$

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. 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

$$
(1.2) \quad
L_N :=
\begin{cases}
D_N^{\frac{1-p}{q-1}}\left(D_N - W_N\right) D_N^{-\frac{r}{q-1}}, & \text{if } q \neq 1, \\[2mm]
D_N - W_N, & \text{if } q = 1,
\end{cases}
$$

where the symmetric weighted adjacency matrix W_N is constructed via a suitably reweighted kernel capturing the similarities between discrete data points, with associated weighted degree matrix D_N; in the special case $q = 2$ the matrix W_N is not

---

*Department of Aerospace and Mechanical Engineering, University of Southern California, Los Angeles, CA 90089, USA (aoberai@usc.edu)

†Computing and Mathematical Sciences, Caltech, Pasadena, CA (fkoh@caltech.edu , bamdadh@caltech.edu , astuart@caltech.edu).

re-weighted. When the data points are drawn i.i.d. from the density ϱ it is useful for intuition to think of the diagonal of D_N as a discrete approximation to the density ϱ.

We detail the role of the parameters p, q and r and provide an explanation of how the limiting (p, q, r)–dependent operator (1.1) emerges. The primary focus of our work is to study the effect of the parameters (p, q, r) on the spectral properties of the limiting differential operator $\mathcal{L}$ as a means to obtain insight into data clustering algorithms which employ graph Laplacians; the results may also be of independent interest in the spectral theory of elliptic differential operators.

In Subsection 1.1 we describe the background for our work, and relevant literature; in Subsection 1.2 we describe our contributions; Subsection 1.3 concludes the introduction with an outline of the paper, by section.

1.1. Background And 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 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, 31, 37, 38, 40] and semi-supervised learning uses this structure as prior information which is enhanced by labelled data [8, 9, 43]. 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. In the following note that W_N is not re-weighted if $q = 2$. 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 [21, 31, 35, 40, 41], 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 [13, 35]. 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 [20] 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 (see Result 1) is linear in ϱ. In [13], general choices of $p = q \geq 0$ and $r = 0$ are investigated in the context of diffusion maps.

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. To the best of our knowledge, apart from the investigations in [40], the role that the specific choice of graph Laplacian plays in the data analysis procedure for which it is constructed has not been studied systematically; such a systematic study is the goal of our work in this paper. Of particular interest is the case where N is large, relevant in large data applications. A series of recent papers have shown that, in the large data limit $N \to \infty$, and under suitable conditions on the sampling density ϱ and the construction of the weights, the operator L_N converges, in an appropriate sense, to an integral, differential or related operator defined on a space of functions on a subset of $\mathbb{R}^d$; see [5, 6, 21, 22, 33, 36, 40, 41]. The point of departure in all of these papers is a kernel w defined on $\mathbb{R}^d \times \mathbb{R}^d$, from which the weighted adjacency matrix defined on the edges of a graph is constructed. Indeed, the seminal paper of Coifman and Lafon [13] works directly on $\mathbb{R}^d$, not on the nodes of the graph, and constructs Markov transition kernels on $\mathbb{R}^d$ from the kernel w and from a continuum degree function also constructed from w; the generators of the resulting Markov processes approach elliptic differential operators, similar to our operator $\mathcal{L}$, in a suitable scaling limit which localizes the kernel w by sending its width to zero; follow-on papers from [13, 14] include [12, 30]. We present a detailed discussion in Appendix A connecting the Markov generators of [13] with our $\mathcal{L}$ operators; showing that they coincide for certain choices of p, q, r.

We focus on large data limits of graph Laplacians in which the limiting operator is an elliptic differential operator of the form (1.1). The works of [21, 36] derive these limiting differential operators for certain choices of (p, q, r) by simultaneously letting $N \to \infty$ and sending the width of the kernel w to 0. The resulting continuum limit operators obtained there have been used as the basis for the analysis of unsupervised [19, 20] and semi-supervised learning algorithms [17]. These papers suggest the potential for further analysis of the continuum limit of graph Laplacians as a means to advance our understanding of classification algorithms in large data limits.

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 [15, 16, 25, 34] 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; the connection between our work and diffusion processes is made in Remark 2.3.

1.2. Our Contribution. The overarching goal of our work is to understand the spectral properties of (1.1) with an emphasis on gaining insight into data clustering algorithms; the clustering of the data is reflected in the geometry of ϱ. We will study how the spectral properties of $\mathcal{L}$ depend on the parameters (p, q, r) and the density ϱ. In pursuit of this overarching goal we will also provide an informal derivation of (1.1) from a family of graph Laplacians $L_N = L_N(p, q, r)$, so that the meaning of $p, q,$ and r is understood. The tools we develop are used to study consistency of semi-supervised learning algorithms in [23]. In particular, we provide the foundations to generalize the work in [24], which studies consistency of graph-based semi-supervised learning algorithms, to the large data limit where L_N is replaced by $\mathcal{L}$ [23]. Furthermore, the spectral theory may be of interest in its own right, independently of the data clustering that motivates us.

The three primary contributions of this paper are as follows:

    1. we explain how $\mathcal{L}$ arises from L_N;

2. under assumptions on ϱ, which capture the notion of data clustered into two sets, we study the low lying spectrum of $\mathcal{L}$ under perturbations of ϱ, the attendant eigenfunctions and their dependence on (p, q, r);

3. we present numerical experiments which exemplify the analysis and extend it to multiple clusters.

Our contributions in point 1 is summarized in the following informal statement.

Result 1 (Convergence of L_N to $\mathcal{L}$). Suppose $(p, q, r) \in \mathbb{R}^3$. Let $\mathcal{L}$ be as in (1.1), L_N as in (1.2), and suppose the weights W_N localize at a suitable rate as $N \to \infty$. Then a rescaling of $\langle \mathbf{u}, L_N \mathbf{u} \rangle_{(p,q,r)}$ Γ-converges to $\langle u, \mathcal{L}u \rangle_{\varrho^{p-r}}$ as $N \to \infty$.

Here, $\langle \cdot, \cdot \rangle_{(p,q,r)}$ is the inner product in $\mathbb{R}^N$ weighted by suitable powers of the degree matrix D_N, and $\langle \cdot, \cdot \rangle_{\varrho^{p-r}}$ is the natural weighted L^2 inner product associated to the space $L^2(\Omega, \varrho^{p-r})$ defined in (2.4). We link vectors $\mathbf{u} \in \mathbb{R}^N$ to functions in $u \in L^2(\Omega, \varrho^{p-r})$ by writing $\mathbf{u} = (u(x_1), \ldots, u(x_N))^T$. The Γ-convergence is in the TL^2 topology introduced in [21] which provides a novel, and potentially widely applicable, methodology for comparison of vectors defined on unstructured grids, and the continuum limits arising as the grid nodes approximate a dense subset of $\mathbb{R}^d$.

In this paper we do not prove this result because doing so would distract from the main thrust of our work. However in Section 5 we give heuristics which demonstrate that L_N given by (1.2) is approximated by $\mathcal{L}$ defined by (1.1), in the limit $N \to \infty$. For specific choices of (p, q, r) the theorem is established in [21]; indeed it is shown that the spectrum of L_N converges to that of $\mathcal{L}$, and the eigenfunctions of L_N converge to those of $\mathcal{L}$ in the TL^2 topology. The methods of proof introduced in [21], and extensions to spectral convergence properties proved there, can be readily generalized to the $(p, q, r)-$ dependent family of graph Laplacian operators considered here.

Our contributions under points 2 and 3 are summarized in the informal statement below, Result 2; this will be substantiated through a combination of theorems (stated informally as Theorems 3.1, 3.2 and 3.3 in Section 3, and presented in full detail, and then proved in Section 6) and numerical results (given in Section 4). In order to describe our contributions in Result 2, we introduce the concepts of *perfectly separated clusters* and *nearly separated clusters*. To define the *perfectly clustered* case, we suppose that the points $\{x_1, ..., x_N\}$ lie within a bounded set $\mathcal{Z} \subset \mathbb{R}^d$ and are sampled i.i.d. from a probability density ϱ_0 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 $\mathcal{L}_0$. Nearly separated clusters comprise a class of smooth $\mathcal{O}(\epsilon)$ perturbations of the perfectly separated case, with density supported everywhere on $\mathcal{Z}$; we define this concept precisely in Section 3.

Result 2. Consider $\mathcal{L}_\epsilon$ when the density $\varrho = \varrho_\epsilon$ is given as a smooth $\mathcal{O}(\epsilon)$ perturbation of ϱ_0.

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

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

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

(ii) The second eigenvalue scales as $\sigma_{2,\epsilon} \asymp \epsilon^q$ and the corresponding eigenvector is

given, approximately in a density weighted L^2 space, by the formula

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

(iii) The third eigenvalue scales as $\sigma_{3,\epsilon} \asymp \min\{1, \epsilon^{q-p-r}\}$ and consequently we have:
- a spectral gap between $\sigma_{2,\epsilon}$ and $\sigma_{3,\epsilon}$ whenever $q \leq p+r$, in that $\sigma_{3,\epsilon} - \sigma_{2,\epsilon} = \mathcal{O}(1)$ and $\sigma_{2,\epsilon}/\sigma_{3,\epsilon} = \mathcal{O}(\epsilon^q)$ as $\epsilon \to 0$;
- a spectral ratio gap between $\sigma_{2,\epsilon}$ and $\sigma_{3,\epsilon}$ whenever $q > p + r$, in that $\sigma_{2,\epsilon}/\sigma_{3,\epsilon} = \mathcal{O}(\epsilon^{p+r})$ as $\epsilon \to 0$.

Since our primary motivation is data clustering, it is relevant to interpret our contributions in that context. We first observe that although our theory is developed under a rather specific assumption about the sampling density of the data, numerical experiments demonstrate that the conclusions are rather robust and work for a wide range of densities which exhibit clustering. Secondly, we observe that the contributions detailed above 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); specifically the spectral gap may only be present in ratio form for some choices of (p, q, r). More precisely, making the choice $q \leq p + r$ one obtains a family of operators whose second eigenvalue decays rapidly, while the gap between the second and third eigenvalues remains uniform under small perturbations. The form of the second eigenfunction, whilst always exhibiting the clusters present in the data, can have different behaviour away from the clusters, depending on (p, q, r) as we now demonstrate with an example.

Example 1.1 (Comparison of unnormalized and normalized graph Laplacians). Consider the two commonly used graph Laplacians, referred to as unnormalized and normalized. These correspond to the operator L_N with parameter choices $(1, 2, 0)$ and $(3/2, 2, 1/2)$ respectively. More precisely we have the two cases

$$\text{unnormalized graph Laplacian}: \quad L_N = D_N - W_N,$$
$$\text{normalized graph Laplacian}: \quad L'_N = D_N^{-1/2} L_N D_N^{-1/2}.$$

All our numerical experiments in this example are for the density shown in Figure 1.1(a).

As $N \to \infty$ and under appropriate conditions (see Result 1 and [21]) the unnormalized operator converges to the differential operator

$$(1.4) \qquad \mathcal{L}_\epsilon : u \mapsto -\frac{1}{\varrho_\epsilon}\mathrm{div}\left(\varrho_\epsilon^2 \nabla u\right).$$

Here $q > p+r$ and it follows from Result 2 that as $\epsilon \downarrow 0$ the second eigenvalue scales as ϵ^2 and the third as ϵ; thus a spectral gap is present only in ratio form. In Figure 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 both converge to zero like ϵ^2 and ϵ respectively. We also compute the second eigenfunction $\varphi_{2,\epsilon}$, referred to as the *Fiedler vector*, shown in Figure 1.1(d). Note that in this case the pointwise distance between $\varphi_{2,\epsilon}$ and the right hand side of (1.3) is only small within the clusters; this reflects the fact that the weighted L^2 norm is not sensitive to large pointwise values of functions in areas where ϱ_ϵ is small.
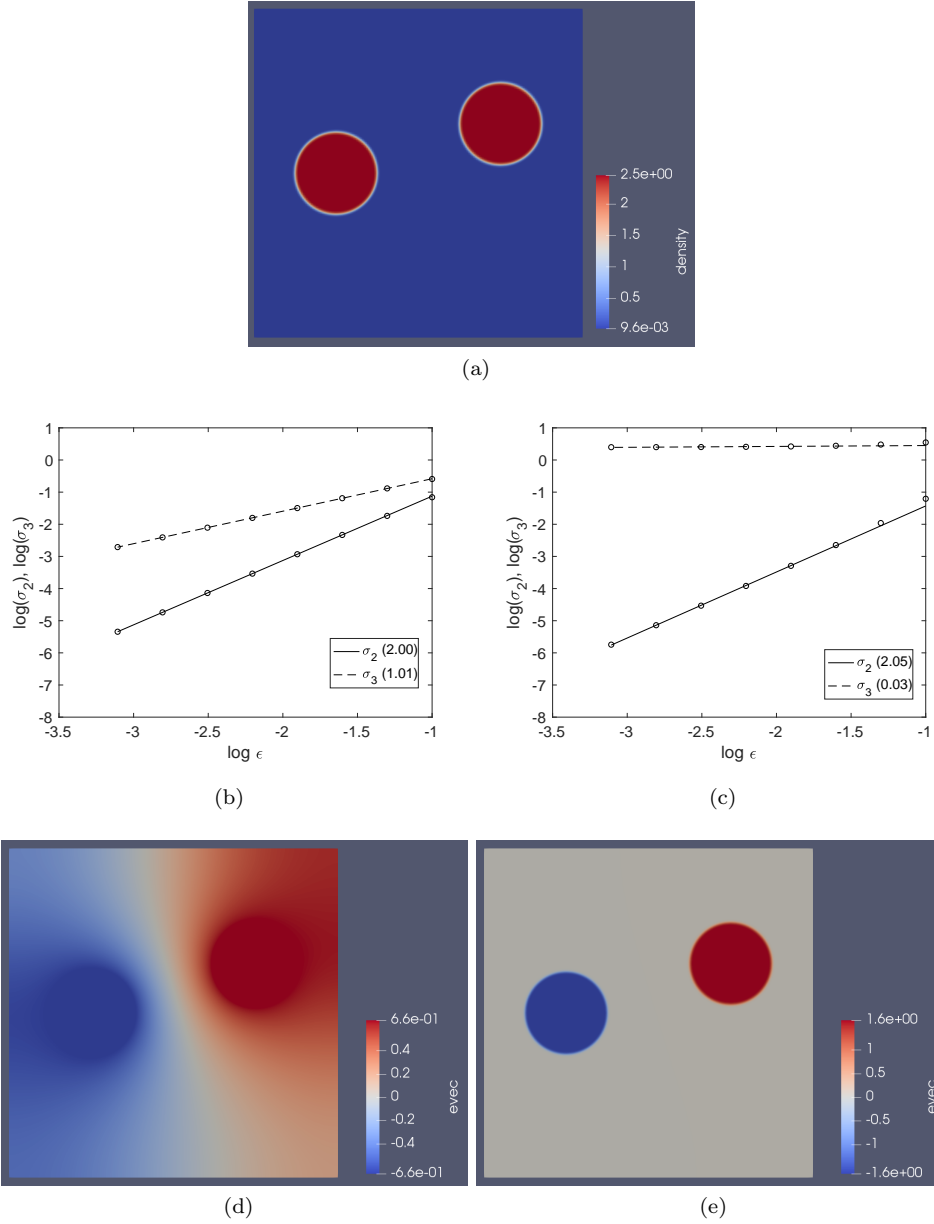
FIG. 1.1. (a) Plot of a density ϱ_ϵ of the form (4.2) with two distinct clusters. (b) Showing $\log(\sigma_2)$ and $\log(\sigma_3)$, the second and third eigenvalues of the unnormalized operator $\mathcal{L}_\epsilon$ in (1.4) as functions of ϵ. Values in brackets in the legends indicate numerical slope of the lines. (c) Showing $\log(\sigma_2)$ and $\log(\sigma_3)$ for the normalized operator $\mathcal{L}'_\epsilon$ in (1.5), as functions of ϵ. (d) The Fiedler vector of $\mathcal{L}_\epsilon$. (e) The Fiedler vector of $\mathcal{L}'_\epsilon$.

For comparison we now consider the matrix L'_N which converges (again, see Result 1 and [21]) to the differential operator

$$(1.5) \qquad \mathcal{L}'_\epsilon : u \mapsto -\frac{1}{\varrho_\epsilon^{3/2}} \mathrm{div}\left(\varrho_\epsilon^2 \nabla\left(\frac{u}{\varrho_\epsilon^{1/2}} \right) \right).$$

6

The heuristic that the diagonal of D_N approximates the density ϱ readily explains the relationship between $\mathcal{L}_\epsilon$ and $\mathcal{L}'_\epsilon$. Since $q = p + r$ this operator is balanced and so Result 2 shows that as $\epsilon \downarrow 0$ there exists a uniform spectral gap between the first two eigenvalues of $\mathcal{L}'_\epsilon$: the second eigenvalue scales as ϵ^2 and the third is of order one with respect to ϵ. In Figure 1.1(c) we also plot the second and third eigenvalues of $\mathcal{L}'_\epsilon$ against ϵ, on a log-scale, and provide best fits to the data; the results support the theory. The resulting Fiedler vector $\varphi_{2,\epsilon}$ is shown in Figure 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 [40]. However, it is noteworthy that the Fiedler vectors of $\mathcal{L}_\epsilon$ and $\mathcal{L}'_\epsilon$ differ substantially within $(\mathcal{Z}')^c$: in the unnormalized case a smooth transition is made between $\mathcal{Z}^+$ and $\mathcal{Z}^-$; in the normalized case abrupt transitions are made to near zero on the boundaries of $\mathcal{Z}^+$ and $\mathcal{Z}^-$. This fact has significant implication for uncertainty quantification studies using latent variables with Gaussian priors determined as a rational function of the graph Laplacian, such as those described in [9, 17].

The important message contained in this example is that the nature in which heuristics about clustering are manifest in the spectral properties of the graph Laplacian can be very sensitive to the specific choice of graph Laplacian used. We demonstrate this in a particular family of large graph limits. Understanding these limiting behaviours is thus fundamental to the exploitation of spectral properties of graph Laplacians for the purposes of clustering. $\qquad\qquad\qquad\qquad\qquad\diamond$

1.3. Outline. The remainder of the paper is organized as follows. Section 2 sets up the problem in detail, providing the framework and notation required in subsequent sections. Section 3 contains the precise statements of Theorems 3.2 and 3.3, relating to Result 2; proofs of these results are postponed to Section 6. Numerical results illustrating, and extending, the theory of Result 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 is the intuition underpinning Result 1; it is of independent interest within the context of clustering algorithms for learning tasks, separately from the spectral analysis contained in the main body of the paper. 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 theorem; 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

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

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

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

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

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

with inner product

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

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

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

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*

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

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 [29, Lem. 2.32]. Compactness follows from Proposition B.3. $\square$

Following the spectral theorem [18, 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 discrete spectrum with eigenvalues* $0 \le \sigma_2 \le \sigma_3 \le \ldots$ *and eigenfunctions* $\{\varphi_j\}_{j \ge 2} \in V^1(\Omega, \varrho)$ *that form an orthogonal basis in both* $V^1(\Omega, \varrho)$ *and* $V^0(\Omega, \varrho)$. *Furthermore, we may extend* $\mathcal{L} : H^1(\Omega, \varrho) \mapsto L^2(\Omega, \varrho^{p-r})$ *and then include the eigenpair* $(\sigma_1, \varphi_1) = (0, |\Omega|_{\varrho^{p+r}}^{1/2} \varrho^r \mathbf{1}_\Omega)$.

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

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

From this we see [32] 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. Fora 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 1. 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
$$\text{dist}(\mathcal{Z}^+, \mathcal{Z}^-) > l > 0, \qquad \text{and} \qquad \text{dist}(\mathcal{Z}^\pm, \partial\mathcal{Z}) > l' > 0.$$
(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 1(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. 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(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, 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

$$(2.7) \qquad \Omega_\delta := \{x : \text{dist}(x, \Omega) \le \delta\},$$

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

$$(2.8) \qquad |\Omega_\epsilon \setminus \Omega| \le \theta \epsilon |\partial \Omega|.$$

Assumption 3. Let $0 < L := \min \text{dist}(\mathcal{Z}^\pm, \partial \mathcal{Z})$. 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*) $\text{supp}\varrho_\epsilon = \bar{\mathcal{Z}}$.
(b) (*Probability density function*) $\int_{\mathcal{Z}} \varrho_\epsilon(x)dx = 1$.
(c) (*Approximation within clusters*) $\exists K_1 > 0$ so that $\|\varrho_\epsilon - \varrho_0\|_{C^\infty(\bar{\mathcal{Z}}')} \le K_1 \epsilon$ 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

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

Once again Assumption 3(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(c) and 3(c)-(e), it follows that ϱ_ϵ is uniformly bounded above and below inside $\mathcal{Z}'$: there exist constants $\varrho_{\epsilon_0}^\pm > 0$ so that

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

Note that the upper bound holds on all of $\mathcal{Z}$ as well, whereas the lower bound clearly does not in view of Assumption 3(d).

The above set of assumptions on ϱ_ϵ may seem very specific; however, the analysis we present is robust to changes in the exact construction of the perturbed densities and we can always construct a ϱ_ϵ satisfying our assumptions, given any ϱ_0.

Example 2.4. Consider the standard mollifier

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

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

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

One can directly verify that the above construction of ϱ_ϵ satisfies Assumption 3. ◇

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:

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

Similarly for ϱ_ϵ,

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

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, these operators have positive, real, discrete eigenvalues after the first eigenvalue, which is zero. For $j = 1, 2, 3, \ldots$ 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 $\varphi_{1,0} = |\mathcal{Z}'|_{\varrho_0^{p+r}}^{-1/2} \varrho_0^r \mathbf{1}_{\mathcal{Z}'}$ and $\varphi_{1,\epsilon} = |\mathcal{Z}|_{\varrho_\epsilon^{p+r}}^{-1/2} \varrho_\epsilon^r \mathbf{1}_{\mathcal{Z}}$, both with corresponding zero eigenvalues. Since we are interested in the eigenpairs for $j \ge 2$ 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 $j \ge 2$ that solve the eigenvalue problems

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

and

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

11

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 2.5 (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.,*

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

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

DEFINITION 2.6 ($\mathcal{L}_0$ 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 $\Lambda_0(\Omega) > 0$*

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

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 2.7 ($\mathcal{L}_\epsilon$ 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*

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

Remark 2.8. 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 3(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 [20].

Remark 2.9. 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 $q - p - r$ for some of our results, see 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. We start in Subsection 3.1 with an informal overview of the results proved. In Subsection 3.2 we consider the perfectly clustered case; Subsection 3.3 perturbs this setting, and considers the nearly clustered case.

3.1. Informal Overview. 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 $(p, q, r) \in \mathbb{R}^3$. Then $\mathcal{L}_0$ is positive semi-definite and self-adjoint on a suitably 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 $\varphi_{j,0}$, $j \geq 1$. Under conditions on ϱ_0 and the sets $\mathcal{Z}^\pm$, the following holds:*
  (i) *The first eigenpair is given by*

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

  (ii) *The second eigenpair is given by*

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

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

To define the nearly separated case we consider a sequence of smooth probability density functions ϱ_ϵ for $\epsilon > 0$ with full support on $\mathcal{Z}$ that converge to ϱ_0 as $\epsilon \to 0$, see Assumption 3; and let $\mathcal{L}_\epsilon$ as in (2.13). The main justification for the spectral clustering approach is the observation that under perturbations of perfect clusters, the perturbed second eigenfunction $\varphi_{2,\epsilon}$ still carries the desired information on clusters. Our aim is to make this intuition rigorous. The two main theorems comprising our second contribution can be stated in summary as follows.

THEOREM 3.2 (Low-lying spectrum of $\mathcal{L}_\epsilon$). *Let $(p, q, r) \in \mathbb{R}^3$ satisfying $p + r > 0$ and $q > 0$, and define $\beta_0 := \min\{q, p + r\}$. Then $\mathcal{L}_\epsilon$ is positive semi-definite and self-adjoint on a suitably weighted Sobolev space $H^1(\mathcal{Z}, \varrho_\epsilon)$. Denote its eigenvalues by $\sigma_{1,\epsilon} \leq \sigma_{2,\epsilon} \leq \cdots$ with corresponding eigenfunctions $\varphi_{j,\epsilon}$, $j \geq 1$. If Assumptions 1, 2, and 3 hold, then there exist $\epsilon_0 > 0$, $\Xi_k > 0$, $k = 1, 2, 3, 4$, so that for all $(\epsilon, \beta) \in (0, \epsilon_0) \times (0, \beta_0)$ it holds that:*
  (i) *The first eigenpair is given by*

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

  (ii) *The second eigenvalue tends to zero as $\epsilon \to 0$:*

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

  (iii) *If $q \leq p + r$, then $\mathcal{L}_\epsilon$ has a uniform spectral gap:*

$$\sigma_{3,\epsilon} \geq \Xi_2 > 0.$$

13

(iv) If $q > p + r$, then $\mathcal{L}_\epsilon$ does not have a uniform spectral gap, it has only a spectral ratio gap:

$$\Xi_3 \epsilon^{2(q-p-r)} \le \sigma_{3,\epsilon} \le \Xi_4 \epsilon^{q-p-r-\beta}.$$

Our numerical experiments in Section 4 demonstrate that when $q > p + r$ there exists a spectral ratio gap: $\sigma_{2,\epsilon}/\sigma_{3,\epsilon} = \mathcal{O}(\epsilon^{p+r})$. These observations suggest that our lower bound on $\sigma_{3,\epsilon}$ in Theorem 3.2(iv) could be sharpened to match the upper bound. These ideas are summarized in Result 2 and illustrated in the example following it.

Our next result, also related to point 2 of our main contributions, identifies the geometry of the Fiedler vector of $\mathcal{L}_\epsilon$; it states that $\varphi_{2,\epsilon}$ approaches $\varphi_{2,0}$ as $\epsilon \to 0$, but in a weighted L^2 space which limits the pointwise value of the approximation outside $\mathcal{Z}'$, in general; see Example 1.1.

THEOREM 3.3 (Geometry of the Fiedler vector of $\mathcal{L}_\epsilon$). *Under the same conditions as Theorem 3.2, the second eigenvalue $\sigma_{2,\epsilon}$ of $\mathcal{L}_\epsilon$ corresponds to a one-dimensional eigenspace. Furthermore, let $\bar{\varphi}_{2,0}$ and $\varphi_{2,\epsilon}$ denote the $L^2(\mathcal{Z}, \varrho_\epsilon^{p-r})$ normalizations of $\varphi_{2,0}$ and the Fiedler vector of $\mathcal{L}_\epsilon$ respectively. Then if $q < 2(p+r)$ there exist constants $\Xi > 0$ and $t > 0$ such that for ϵ small enough,*

$$\left| 1 - \langle \varphi_{2,\epsilon}, \bar{\varphi}_{2,0} \rangle_{\varrho_\epsilon^{p-r}} \right| \le \Xi \epsilon^t.$$

Our numerical experiments in Section 4 show that the condition $q < 2(p+r)$ is not necessary and can be dropped by sharpening the lower bound on $\sigma_{3,\epsilon}$ in Theorem 3.2(iv) so that it matches the upper bound.

Remark 3.4. The take-home message from these two theorems is two-fold: (i) Theorem 3.2 tells us that particular care is needed when looking for a spectral gap characterizing the number of clusters if $q > p + r$ as the gap may only be manifest in ratio form, not absolutely, leading to potential overestimation of the number of clusters; (ii) Theorem 3.3 tells us the form and geometry of the Fiedler vector which characterizes the two clusters, and its dependence on ϱ and on ϵ. Together these lead to useful rules of thumb when performing data analysis with graph Laplacians formed from large data sets.

3.2. Perfectly Separated Clusters. 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 $\sigma_{3,0}$.

THEOREM 3.5 (Spectral gap for $\mathcal{L}_0$). *Let ϱ_0 satisfy Assumption 2. If*

(3.1) $$\min\{\Lambda_\Delta(\mathcal{Z}^+), \Lambda_\Delta(\mathcal{Z}^-)\} > 0$$

then $\sigma_{3,0} > 0$.

For the proof, see Proposition 6.1. At an intuitive level the above theorem states that as long as the clusters $\mathcal{Z}^\pm$ cannot be further split into smaller clusters then $\sigma_{3,0} > 0$. In the case where ϱ_0 is constant the condition (3.1) reduces to a condition on the connectivity of the clusters $\mathcal{Z}^\pm$:

Example 3.6. Suppose $\varrho_0 > 0$ is constant on $\mathcal{Z}'$. Then the $\mathcal{L}_0$ spectral gaps $\Lambda_0(\mathcal{Z}^\pm)$ are equal to the minimum of the standard spectral gaps $\Lambda_\Delta(\mathcal{Z}^\pm)$; this minimum is non-zero as long as $\mathcal{Z}^\pm$ are connected sets of positive measure. ◇

3.3. 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 $\mathcal{Z}^\pm$. For the proof of the following result, see Section 6.4.

THEOREM 3.7 (Second eigenvalue of $\mathcal{L}_\epsilon$). *Let $(p, q, r) \in \mathbb{R}^3$ satisfying $p + r > 0$ and $q > 0$, and suppose Assumptions 1, 2, and 3 hold. Then $\exists\, \epsilon_0 > 0$ so that $\forall (\epsilon, \beta) \in (0, \epsilon_0) \times (0, q)$,*

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

where $\Xi > 0$ is a uniform constant.

We are not able to give a closed form for $\varphi_{2,\epsilon}$ in general but based on our knowledge of the perfectly clustered case, we expect that the second eigenfunction $\varphi_{2,\epsilon}$ converges to a normalization of $\varphi_{2,0}$ with $\varphi_{2,0}$ given in Theorem 3.1 (ii), under suitable conditions. To show this we first construct an explicit approximation $\varphi_{F,\epsilon}$ of $\varphi_{2,\epsilon}$ for $\epsilon > 0$. More precisely, we take $\varphi_{F,\epsilon}$ to be 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 3.7. By construction, we then have that $\varphi_{F,\epsilon}$ converges to the normalization of $\varphi_{2,0}$ as $\epsilon \to 0$.

In order to quantify how close our construction $\varphi_{F,\epsilon}$ is to the eigenfunction $\varphi_{2,\epsilon}$, we require bounds for $\sigma_{2,\epsilon}$ and $\sigma_{3,\epsilon}$. These calculations are mainly based on the min-max formulas (also known as the Courant–Fisher theorem, see Theorem C.1), a weighted version of Cheeger's inequality (see Lemma 6.5 and Theorem D.1), and standard ideas from PDE theory for weighted elliptic operators. The full analysis is postponed to Section 6.

Several interesting conclusions can be drawn from our arguments, notably desired relationships between the parameters p, q and r. The existence of spectral gaps for $\mathcal{L}_\epsilon$ inside the clusters and away from the clusters separately allows us to deduce bounds on the low-lying spectrum. Consider the set

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

and suppose that for some fixed $\epsilon_0 > 0$, we have $\Lambda_\Delta(\mathcal{Z} \smallsetminus \mathcal{Z}'_{\epsilon_0}) > 0$, that is, $\mathcal{L}_\epsilon$ has a spectral gap away from the clusters according to Definition 2.5. Since $\varrho_\epsilon(x) = K_2 \epsilon$ for $x \in \mathcal{Z} \smallsetminus \mathcal{Z}'_{\epsilon_0}$, we have for all $u \perp \mathbf{1}_{\mathcal{Z} \smallsetminus \mathcal{Z}'_{\epsilon_0}}$ in $V^1(\mathcal{Z} \smallsetminus \mathcal{Z}'_{\epsilon_0})$

$$(K_2 \epsilon)^{2r-q} \int_{\mathcal{Z} \smallsetminus \mathcal{Z}'_{\epsilon_0}} \left| \nabla \left(\frac{u}{\varrho_\epsilon^r} \right) \right|^2 \varrho_\epsilon^q dx \ge \Lambda_\Delta(\mathcal{Z} \smallsetminus \mathcal{Z}'_{\epsilon_0})(K_2 \epsilon)^{r-p} \int_{\mathcal{Z} \smallsetminus \mathcal{Z}'_{\epsilon_0}} \left| \frac{u}{\varrho_\epsilon^r} \right|^2 \varrho_\epsilon^{p-r} dx.$$

This simple calculation shows that $\Lambda_\epsilon(\mathcal{Z} \smallsetminus \mathcal{Z}'_{\epsilon_0}) = \mathcal{O}(\epsilon^{q-p-r})$, 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 $q \le p + r$ to ensure $\Lambda_\epsilon(\mathcal{Z} \smallsetminus \mathcal{Z}'_{\epsilon_0}) > 0$ uniformly which is precisely the condition in Theorem 3.2(iii).

If we assume that $\min\{\Lambda_0(\mathcal{Z}^+), \Lambda_0(\mathcal{Z}^-)\} > 0$, we then anticipate $\Lambda_\epsilon(\mathcal{Z}^\pm) > 0$ for small enough $\epsilon_0 > 0$ since ϱ_ϵ converges to ϱ_0 on $\mathcal{Z}'$. Intuitively, we expect the spectral gap condition on $\mathcal{Z}^\pm$ to hold as long as each the clusters are connected since ϱ_0 is bounded above and below on the clusters by Assumption 2(c). We will make this

intuition rigorous by making use of Theorem D.1. From the above considerations, we expect to obtain different bounds on $\sigma_{3,\epsilon}$ in cases where $q \leq p + r$ and $q > p + r$:

THEOREM 3.8 (Third eigenvalue of $\mathcal{L}_\epsilon$). *Let $(p, q, r) \in \mathbb{R}^3$ satisfying $p + r > 0$ and $q > 0$, and suppose Assumptions 1, 2, and 3 hold and $\Lambda_\Delta(\mathcal{Z} \smallsetminus \mathcal{Z}'_{\epsilon_0}) > 0$ for a sufficiently small $\epsilon_0 > 0$. Then the following holds for all $\epsilon \in (0, \epsilon_0)$:*

(i) If $q > p + r$, then $\exists \Xi_1, \Xi_2 > 0$ independent of ϵ such that $\forall \beta \in (0, (q - p - r)/2)$,

$$\Xi_1 \epsilon^{2(q-p-r)} \leq \sigma_{3,\epsilon} \leq \Xi_2 \epsilon^{q-p-r-2\beta},$$

and so $\mathcal{L}_\epsilon$ does not have a uniform spectral gap on $\mathcal{Z}$.

(ii) If $q \leq p + r$ then there exists a constant $\Xi_3 > 0$, independent of ϵ, so that

$$\sigma_{3,\epsilon} \geq \Xi_3,$$

and so $\mathcal{L}_\epsilon$ has a uniform spectral gap on $\mathcal{Z}$.

In the case $q > p + r$, we do not have an absolute spectral gap between the second and third eigenvalues. However, we are able to deduce a quantitative spectral ratio gap. The proof of the following theorem follows by combining Theorem 3.7 and Theorem 3.8.

THEOREM 3.9 (Spectral ratio gap when $q \geq p + r$). *Suppose the conditions of Theorem 3.8 hold and $q > p + r$. Then there exists a constant $\Xi > 0$ independent of ϵ, so that for all $(\epsilon, \beta) \in (0, \epsilon_0) \times (0, 2(p + r) - q)$,*

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

While this theorem suggests that there may be no spectral ratio gap when $q > 2(p + r)$, our numerical experiments in Section 4.2 indicate that when $q > p + r$ the spectral ratio gap is indeed of order $\mathcal{O}(\epsilon^{p+r})$. This discrepancy between the theory and numerical results suggests that the lower bound on $\sigma_{3,\epsilon}$ in Theorem 3.8(i) can be improved to match the upper bound so that we obtain, when $q > p + r$,

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

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

THEOREM 3.10 (Geometry of the second eigenfunction $\varphi_{2,\epsilon}$). *Suppose the conditions of Theorem 3.8 are satisfied. Let $\beta_0 := \min\{q, p + r\}$. Then there exists $\Xi, \epsilon_0 > 0$ so that $\forall (\epsilon, \beta) \in (0, \epsilon_0) \times (0, \beta_0)$*

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

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

4. Numerical Simulations. 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 where $q = p + r$ and Subsection 4.2 the unbalanced case where $q > p + r$. In

Subsection 4.3 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 Theorem 3.9, arising when $q > p + r$, is indeed of $\mathcal{O}(\epsilon^{p+r})$ 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) :

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

Our numerics are all performed in dimension $d = 2$. We solve this by the finite element method using the FEniCS software package [28]. 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

$$(4.2) \qquad \varrho_\epsilon(x) = C^{-1}\left(\epsilon + \sum_{i=1}^{K} \frac{\mathrm{erf}\left(\epsilon^{-1}(\theta_i - |x - x_i|)\right)}{4\pi\theta_i^2}\right),$$

where $|\cdot|$ is the two dimensional Euclidean norm, K is the number of circular clusters, x_i denotes the ith cluster centre, θ_i the ith 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 $x_1 = (-0.5, 0.0)$, $\theta_1 = 0.25$, $x_2 = (0.5, 0.3)$, and $\theta_2 = 0.25$ as shown in Figure 1.1(a). In Subsection 4.3 we consider three and five clusters adding the point $x_3 = (0.4, -0.5)$ with radius $\theta_3 = 0.15$, to make three clusters, and then adding $x_4 = (-0.35, 0.65)$ and $x_5 = (-0.6, -0.6)$ with radii $\theta_4 = 0.20$ and $\theta_5 = 0.15$, to generate five clusters. We plot the resulting densities in Figure 4.1.

4.1. Binary Balanced Case: $q = p + r$. In Figure 4.2(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 behaviour for $\epsilon \downarrow 0$ is observed in this regime. Theorem 3.7 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 Figure 4.2(a), and compare the numerical slopes to the analytic prediction in the first four rows of Table 4.1.

In Figure 4.2(b), we plot the ratio $\sigma_{2,\epsilon}/\sigma_{3,\epsilon}$ for different values of ϵ. By Theorem 3.8 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 Figure 4.2(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.

In Figure 4.2(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

17

similar results to Figure 4.2(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.

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 Figure 4.3(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 Theorem 3.9. As mentioned earlier, these results suggest that the lower bound on $\sigma_{3,\epsilon}$ in Theorem 3.8(i) can be sharpened to match the upper bound. In Figure 4.3(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 3.7, and $\sigma_{2,\epsilon}/\sigma_{3,\epsilon} \sim \epsilon^{p+r}$, which is better than the predicted rate in Theorem 3.9; again the results suggest that the lower bound on $\sigma_{3,\epsilon}$ in Theorem 3.8(i) 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 and, furthermore, suggest that the lower and upper bounds on the third eigenvalue should match, suggesting tighter bounds on the spectral ratio gap could be achievable.

4.3. 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 Figure 4.1. In Figures 4.4 and 4.5 we display the behaviour of the Kth 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 strongly suggest that $\sigma_{K,\epsilon} \sim \sigma_{K,\epsilon}/\sigma_{K+1,\epsilon} \sim \mathcal{O}(\epsilon^q)$. The case $q < p + r$ behaves the same way as $q = p + r$, consistent with the theory in the binary cluster case. This suggests a natural extension of the result in Theorems 3.7, 3.8 and 3.9 from the binary case to multiple clusters.

In Figures 4.6 and 4.7 we collect similar results but in the unbalanced regime where $q > p + r$. 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 $\sigma_{K,\epsilon}/\sigma_{K+1,\epsilon} \sim \epsilon^{p+r}$ in both the three and five cluster cases.

The above results lead to the conjecture that in the $K \geq 2$ cluster setting we have

$$(4.3) \qquad \sigma_{K,\epsilon} = \mathcal{O}(\epsilon^q), \qquad \frac{\sigma_{K,\epsilon}}{\sigma_{K+1,\epsilon}} = \mathcal{O}(\epsilon^{\min\{q, p+r\}}).$$

We provide further evidence for this claim in Tables 4.3 and 4.4 where we collect numerical approximations to the above rates for different choices of p, q, r in the balanced and unbalanced regimes where $q = p + r$ or $q > p + r$.

5. From Discrete To Continuum. In this section we present formal calculations showing 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 a density ϱ supported on $\mathcal{Z}$ according to which the vertices $\{x_n\}_{n=1}^N$ are i.i.d. We begin in Subsection 5.1 by introducing a weighted graph construction that leads to the correct weighting of L_N. In Section 5.2 we discuss certain properties of the L_N operators including self-adjointness and invariance of the spectrum under parameter choices. Subsection 5.3 outlines a roadmap for rigorous proof of convergence of L_N to $\mathcal{L}$ in the framework of [21, 36, 19] through study of the convergence of Dirichlet energies, using the law of large numbers and localization of the weights.

5.1. Graph Laplacian Normalizations. 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 [17], 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. We have in mind an η that assigns lower weight to edges connecting far apart points that are considered less similar. Furthermore, let $\tilde{D}_N = \text{diag}(\tilde{d}_i)$ 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 $\varrho(x_i)$. For $q \geq 0$, we introduce the matrix W_N, 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 $D_N = \text{diag}(d_i)$ where $d_i := \sum_{j=1}^N W_{ij}$. We now define the graph Laplacian L_N as in (1.2) for $(p, q, r) \in \mathbb{R}^3$,

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

5.2. Properties Of The Discrete Operator L_N. 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 an appropriately weighted inner product. For parameters (p, q, r), we can directly verify that L_N is self-adjoint with respect to the inner product

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

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

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

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

$$
\begin{aligned}
\langle \mathbf{u}, L_N \mathbf{u} \rangle_{(p,q,r)} &= \langle D_N^{\frac{p-1-r}{q-1}} \mathbf{u}, L_N \mathbf{u} \rangle \\
&= \langle D_N^{\frac{p-1}{q-1}} \mathbf{v}, D_N^{\frac{1-p}{q-1}} (D_N - W_N) \mathbf{v} \rangle \\
&= \langle \mathbf{v}, (D_N - W_N) \mathbf{v} \rangle
\end{aligned}
$$

$$(5.1) \qquad = \frac{1}{2} \sum_{i,j} W_{ij} |v_i - v_j|^2 = \frac{1}{2} \sum_{i,j} W_{ij} \left| \frac{u_i}{d_i^{r/(q-1)}} - \frac{u_j}{d_j^{r/(q-1)}} \right|^2 .$$

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. To see this, consider the situation in which the graph G 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 $\langle u, L_N u \rangle_{(p,q,r)}$ to zero. As a consequence, a simple continuity argument 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 parameter choices $(p,q,r) = (1,2,0)$, $(3/2,2,1/2)$ and $(1,1,0)$ discussed in the introduction (see Subsection 1.1), the weighted inner product $\langle \cdot, \cdot \rangle_{(p,q,r)}$ reduces to the usual Euclidean inner product.

Remark 5.1. There is a family of operators L_N corresponding to certain (p,q,r) that all produce the same spectrum. To see this, let L_N^1 denote a graph Laplacian defined by (1.2) with parameters (p_1, q_1, r_1). Then any other graph Laplacian L_N^2 with parameter choices (p_2, q_2, r_2) produces the same spectrum if

$$(5.2) \qquad \frac{p_1 + r_1 - 1}{q_1 - 1} = \frac{p_2 + r_2 - 1}{q_2 - 1} .$$

To see this, assume $(\sigma, \mathbf{u})$ 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

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

we have

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

Writing

$$\mathbf{v} := D_N^{-\frac{r_1}{q_1 - 1}} \mathbf{u}, \qquad \tilde{\mathbf{v}} := D_N^{-\frac{r_2}{q_2 - 1}} \tilde{\mathbf{u}}$$

gives $\tilde{\mathbf{v}} = \mathbf{v}$ 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

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

This idea of variable bandwidth originates in the statistical density estimation literature [27, 39] and was introduced to the machine learning community, in the context of graph based data analysis, in [42]. 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 [13] are generalized to the variable bandwidth setting.

5.3. The Large Data Limit. It is shown in the papers [21, 36, 19] 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 introduce, and use, the TL^2 topology which facilitates a Γ−convergence analysis 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 $\mathcal{Z}$). 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 [17], for example. A similar methodology can be applied to show convergence of L_N to $\mathcal{L}$ for any choice of parameters $(p, q, r) \in \mathbb{R}^3$. We now provide the framework within which to study this convergence.

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 discrete weighted Dirichlet energy can be applied 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. Γ-convergence is naturally the right tool for this analysis as it facilitates study of the convergence of the minimizers of $E_{N,\delta}$ to the minimizers of the limiting functional. In the limit, we obtain the *continuous weighted Dirichlet energy* $E : L^2(\mathcal{Z}, \varrho^{p-r}) \to [0, \infty]$ defined as

$$E(u) := \begin{cases} \frac{1}{2} \int_{\mathcal{Z}} \left| \nabla \left(\frac{u(x)}{\varrho(x)^r} \right) \right|^2 \varrho(x)^q \, dx & \text{if } u \in H^1(\mathcal{Z}, \varrho), \\ \infty & \text{if } u \in L^2(\mathcal{Z}, \varrho^{p-r}) \smallsetminus H^1(\mathcal{Z}, \varrho), \end{cases}$$

In Subsection 5.4 we detail the intuition behind this limit.

Once the convergence of the Dirichlet energies has been established, generalizations of the results in [21] immediately follow. Denote $(\sigma_k^{(N)}, \mathbf{u}_k^{(N)})$ and (σ_k, u_k) the eigenpairs of the operators L_N and $\mathcal{L}$ respectively. If either $q = 1$, or $q \neq 1$ and $p = r+1$ so that $\langle \cdot, \cdot \rangle_{(p,q,r)}$ is the usual Euclidean inner product, then the spectrum of L_N converges to the spectrum of $\mathcal{L}$, and the eigenfunctions $\mathbf{u}_k^{(N)}$ converge to u_k in the TL^2 topology. These results can be proven analogously to [21, Thm. 1.2], and provide

justification for our assertion that analysis of the spectral properties of the operator $\mathcal{L}$ given by (1.1) gives insight into the spectral properties of L_N for large data sets. Spectral analysis of $\mathcal{L}$ is the focus of Sections 3 and 6.

5.4. Convergence of Dirichlet Energies. In this subsection, we describe why we expect the discrete operators L_N to converge (in the appropriate TL^2 topology) 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.

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 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 rewrite the discrete weighted Dirichlet energy $E_{N,\delta} : \mathbb{R}^N \to [0, \infty)$ using (5.1) and the functions defined above (case $q \neq 1$):

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

$$= \frac{N^{2r-q}}{2\delta^2} \sum_{i,j} W_{ij} \left| \frac{u_i}{d_i^{r/(q-1)}} - \frac{u_j}{d_j^{r/(q-1)}} \right|^2$$

$$= \frac{N^{2r-q}}{2\delta^2} \sum_{i,j} \left(\frac{\tilde{W}_{ij}}{\tilde{d}_i^{1-q/2} \tilde{d}_j^{1-q/2}} \right) \left| \frac{u_i}{d_i^{r/(q-1)}} - \frac{u_j}{d_j^{r/(q-1)}} \right|^2$$

$$= \frac{1}{2\delta^2 N^2} \sum_{i,j} \left(\frac{\eta_\delta(|x_i - x_j|)}{\left(\tilde{d}^{N,\delta}(x_i)\right)^{1-q/2} \left(\tilde{d}^{N,\delta}(x_j)\right)^{1-q/2}} \right)$$

$$\times \left| \frac{u_i}{(d^{N,\delta}(x_i))^{r/(q-1)}} - \frac{u_j}{(d^{N,\delta}(x_j))^{r/(q-1)}} \right|^2.$$

This formulation allows us to extend $E_{N,\delta}$ from vectors to functions on $\mathcal{Z}$ as discussed in Subsection 5.3. More precisely, for $u : \mathcal{Z} \to \mathbb{R}$, we have

$$E_{N,\delta}(u) = \frac{1}{2\delta^2} \iint_{\mathcal{Z} \times \mathcal{Z}} \left(\frac{\eta_\delta(|x-y|)}{\left(\tilde{d}^{N,\delta}(x)\right)^{1-q/2} \left(\tilde{d}^{N,\delta}(y)\right)^{1-q/2}} \right)$$

(5.3)
$$\times \left| \frac{u(x)}{\left(d^{N,\delta}(x)\right)^{r/(q-1)}} - \frac{u(y)}{\left(d^{N,\delta}(y)\right)^{r/(q-1)}} \right|^2 \, d\mu_N(x) d\mu_N(y) \,.$$

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

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

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

Define

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

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

$$E_{N,\delta}(u) = \frac{1}{2\delta^2} \iint_{\mathcal{Z} \times \mathcal{Z}} \left(\frac{\eta_\delta(|x-y|)}{\left(\tilde{d}^{N,\delta}(x)\right)^{1-q/2} \left(\tilde{d}^{N,\delta}(y)\right)^{1-q/2}} \right)$$

$$\times \left| \frac{u(x)}{\left(d^{N,\delta}(x)\right)^{r/(q-1)}} - \frac{u(y)}{\left(d^{N,\delta}(y)\right)^{r/(q-1)}} \right|^2 \, d\mu_N(x) d\mu_N(y)$$

$$\overset{N \gg 1}{\approx} \frac{1}{2\delta^2} \iint_{\mathcal{Z} \times \mathcal{Z}} \left(\frac{\eta_\delta(|x-y|)}{\left(\tilde{d}^\delta(x)\right)^{1-q/2} \left(\tilde{d}^\delta(y)\right)^{1-q/2}} \right)$$

$$\times \left| \frac{u(x)}{\left(d^\delta(x)\right)^{r/(q-1)}} - \frac{u(y)}{\left(d^\delta(y)\right)^{r/(q-1)}} \right|^2 \varrho(x)\varrho(y) \, dxdy$$

$$\overset{\delta \ll 1}{\approx} \frac{1}{2\delta^2} \iint_{\mathcal{Z} \times \mathcal{Z}} \left(\frac{\eta_\delta(|x-y|)}{\left(\tilde{d}^\delta(x)\right)^{1-q/2} \left(\tilde{d}^\delta(y)\right)^{1-q/2}} \right)$$

$$\times \left| \nabla \left(\frac{u(x)}{\left(d^\delta(x)\right)^{r/(q-1)}} \right) \cdot (x-y) \right|^2 \varrho(x)\varrho(y) \, dxdy$$

$$\overset{\delta \ll 1}{\approx} \frac{1}{2} \frac{s_2}{s_0^{2r+2-q}} \int_{\mathcal{Z}} \frac{1}{\varrho(x)^{2-q}} \left| \nabla \left(\frac{u(x)}{\varrho(x)^r} \right) \right|^2 \varrho(x)^2 \, dx$$

$$= \frac{1}{2} \frac{s_2}{s_0^{2r+2-q}} \int_{\mathcal{Z}} \left| \nabla \left(\frac{u(x)}{\varrho(x)^r} \right) \right|^2 \varrho(x)^q \, dx = \frac{s_2}{s_0^{2r+2-q}} E(u) \,.$$

23

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 [21].

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, we start in Subsection 6.1 giving some informal intuition; in Subsection 6.2 we discuss the perfectly clustered case, and then consider small perturbations of this setting, the nearly clustered case, in Subsection 6.3. Theorems 3.7, 3.8 and 3.10 are then proved in Subsections 6.4, 6.5 and 6.6 respectively.

6.1. Informal Overview. There are three scaling results, with respect to ϵ, that are at the heart of our theorems, and the results on the scaling behaviour of eigenvalues in particular. The first two concern the Rayleigh quotient

$$(6.1) \qquad \frac{\int_{\mathcal{Z}'} \left| \nabla \left(\frac{u}{\varrho^r} \right) \right|^2 \varrho^q dx}{\int_{\mathcal{Z}'} \left| \frac{u}{\varrho_0^r} \right|^2 \varrho^{p+r} dx}$$

and the third arises from a weighted Cheeger-type inequality. We highlight these three scalings here, before getting into the details of the proofs.

- Theorem 3.7 (Theorem 3.2(ii)) shows an upper bound on $\sigma_{2,\epsilon}$ which scales like $\epsilon^{q-\beta}$, for any $\beta > 0$. This scaling arises from use of the min-max theorem and in particular estimation of the Rayleigh quotient (6.1) with a test function $u(x)$ which is normalized with respect to the denominator, and estimating the numerator using the fact that the test function is constructed so that $u(x)/\varrho(x)^r$ transitions smoothly from value 1 to value ϵ on length-scale ϵ^β; scaling $\varrho \to \epsilon\varrho$ in the Rayleigh quotient reveals the desired bound.
- Theorem 3.8(i) (Theorem 3.2(iv)) shows upper and lower bounds on $\sigma_{3,\epsilon}$ in the case $q > p + r$. The upper bound on $\sigma_{3,\epsilon}$ scales like $\epsilon^{q-p-r-\beta}$, while the lower bound (which our numerics from Section 4 suggest is not sharp) scales like $\epsilon^{2(q-p-r)}$. The upper bound follows from a similar argument to the upper bound on $\sigma_{2,\epsilon}$. The lower bound follows from a Cheeger-type inequality using a weighted isoperimetric function (6.14). This function has the same scaling under $\varrho \to \epsilon\varrho$ which we describe in the previous bullet, leading to a lower bound on the weighted isoperimetric function which scales like ϵ^{q-p-r}. This is squared in the Cheeger inequality leading to the result as stated.

6.2. Perfectly Separated Clusters. We begin with a precise characterization of the third eigenvalue of $\mathcal{L}_0$ in Theorem 3.5.

PROPOSITION 6.1. *Suppose Assumptions 1 and 2 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(e) ensures that $\varphi_{2,0} = |\mathcal{Z}'|^{1/2}_{\varrho_0^{p-r}} \varrho^r \left(\mathbf{1}_{\mathcal{Z}^+} - \mathbf{1}_{\mathcal{Z}^-} \right)$ belongs to $V^0(\mathcal{Z}', \varrho_0)$. Let $u \in V^1(\mathcal{Z}', \varrho_0)$ so that $u \perp \mathrm{span}\{\varphi_{1,0}, \varphi_{2,0}\}$ 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 2.6, $u|_{\mathcal{Z}^\pm}$ satisfy Poincaré inequalities of the form (2.17) on

$\mathcal{Z}^\pm$ with optimal constants $\Lambda_0^\pm$. Hence

$$\int_{\mathcal{Z}'} \left| \nabla \left(\frac{u}{\varrho_0^r} \right) \right|^2 \varrho_0^q dx = \int_{\mathcal{Z}^+} \left| \nabla \left(\frac{u}{\varrho_0^r} \right) \right|^2 \varrho_0^q dx + \int_{\mathcal{Z}^-} \left| \nabla \left(\frac{u}{\varrho_0^r} \right) \right|^2 \varrho_0^q dx$$

$$\geq \min\{\Lambda_0^+, \Lambda_0^-\} \left(\int_{\mathcal{Z}^+} \left| \frac{u}{\varrho_0^r} \right|^2 \varrho_0^{p+r} dx + \int_{\mathcal{Z}^-} \left| \frac{u}{\varrho_0^r} \right|^2 \varrho_0^{p+r} dx \right)$$

$$= \min\{\Lambda_0^+, \Lambda_0^-\} \int_{\mathcal{Z}'} \left| \frac{u}{\varrho_0^r} \right|^2 \varrho_0^{p+r} dx.$$

The result now follows from the max-min formula (C.2) in Theorem C.1. $\qquad\square$

Remark 6.2. If Assumption 2(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 $\varphi_{2,0} \in V^0(\mathcal{Z}', \varrho_0)$.

6.3. 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. Throughout this section, we routinely assume that Assumptions 1, 2 and 3 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 $\mathbf{1}_{\mathcal{Z}^\pm}$. 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 [29, Thm. 3.6]). Now define the functions $\chi_\epsilon^\pm \in C^\infty(\bar{\mathcal{Z}})$ by renormalizing $\xi_\epsilon^\pm$ in $L^2(\bar{\mathcal{Z}}, \varrho_\epsilon^{p-r})$,

$$(6.2) \qquad\qquad \chi_\epsilon^+ := b_\epsilon^+ \xi_\epsilon^+, \qquad \chi_\epsilon^- := b_\epsilon^- \xi_\epsilon^-,$$

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

$$(6.3) \qquad\qquad \int_{\mathcal{Z}_{\epsilon_1}^+} \varrho_\epsilon^{p+r} \chi_\epsilon^+ \, dx = \int_{\mathcal{Z}_{\epsilon_1}^-} \varrho_\epsilon^{p+r} \chi_\epsilon^- \, dx,$$

$$b_\epsilon^+ + b_\epsilon^- = 2.$$

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 Figure 6.1.

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

$$(6.4) \qquad\qquad \varphi_{F,\epsilon}(x) = \frac{\varrho_\epsilon^r(x) \left[\chi_\epsilon^+(x) - \chi_\epsilon^-(x) \right]}{\left\| \varrho_\epsilon^r(x) \left[\chi_\epsilon^+(x) - \chi_\epsilon^-(x) \right] \right\|_{L^2(\mathcal{Z}, \varrho_\epsilon^{p-r})}}.$$
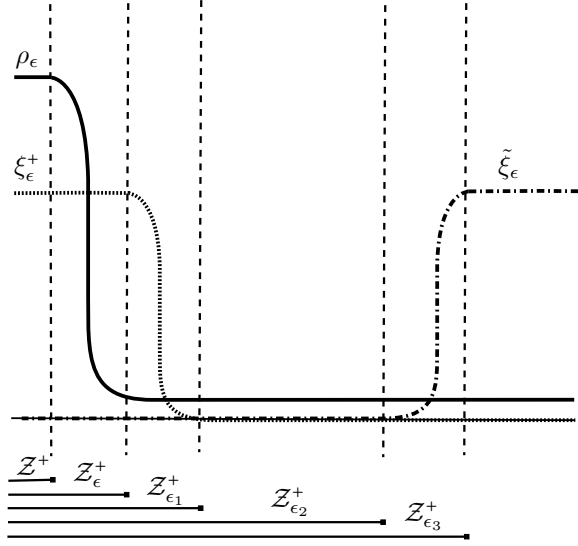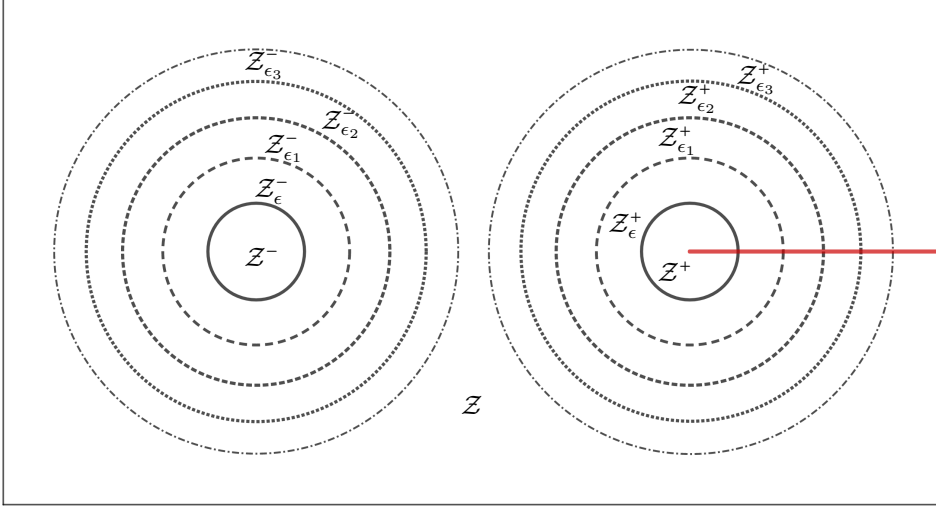
FIG. 6.1. *Schematic depiction of the different sets and functions used in construction of $\varphi_{F,\epsilon}$ from (6.4) and $\tilde{\varphi}_{F,\epsilon}$ from (6.16). Top: overhead schematic of the sets $\mathcal{Z}, \mathcal{Z}^{\pm}, \mathcal{Z}^{\pm}_{\epsilon}, \mathcal{Z}^{\pm}_{\epsilon_1}, \mathcal{Z}^{\pm}_{\epsilon_2}$, 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.*

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 $\nabla \chi^{\pm}_{\epsilon}$ 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 $p+r > 0$ and Assumptions 1, 2 and 3 hold and let $b_\epsilon^\pm$ be as in (6.2). 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*

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

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 (6.3) 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

$$(6.5) \qquad\qquad |\Xi_\epsilon - 1| \le \Xi_1 \epsilon^{\min\{1, p+r\}},$$

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

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

for some $\Xi > 0$, which concludes the proof of the lemma. It remains to show (6.5). Following Assumption 3(c, d), for sufficiently small ϵ,

$$\Xi_\epsilon \le \frac{\int_{\mathcal{Z}_\epsilon^+} \varrho_\epsilon^{p+r} dx + K_2^{p+r} \epsilon^{p+r} |\mathcal{Z}_{\epsilon_1}^+ \smallsetminus \mathcal{Z}^+|}{\int_{\mathcal{Z}^-} \varrho_\epsilon^{p+r} dx}$$

$$\le \frac{\int_{\mathcal{Z}^+} (\varrho_0 + K_1 \epsilon)^{p+r} dx + \int_{\mathcal{Z}_\epsilon^+ \smallsetminus \mathcal{Z}^+} \varrho_\epsilon^{p+r} dx + K_2^{p+r} \epsilon^{p+r} |\mathcal{Z}_{(\epsilon_0 + \epsilon_0^\beta)}^+ \smallsetminus \mathcal{Z}^+|}{\int_{\mathcal{Z}^-} (\varrho_0 - K_1 \epsilon)^{p+r} dx}.$$

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

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

Thanks to Assumption 2(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

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

since ϱ_0 is bounded below uniformly on $\mathcal{Z}^-$ by Assumption 2(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

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

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. □

6.4. Proof of Theorem 3.7 (Second Eigenvalue of $\mathcal{L}_\epsilon$). 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 (6.3) 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 $u \in \text{span}\{\varphi_{1,\epsilon}, \varphi_{F,\epsilon}\}$. 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 6.9 in Section 6.6 for a more detailed argument). Using $0 < b_\epsilon^\pm < 2$ and Assumption 3(d), we have

$$\mathcal{R}_\epsilon(\varphi_{F,\epsilon}) \leq \frac{4}{\Xi_0} \int_{\mathcal{Z}} |\nabla (\xi_\epsilon^+ - \xi_\epsilon^-)|^2 \varrho_\epsilon^q dx$$

(6.6)
$$= \frac{4}{\Xi_0} \int_{\mathcal{Z}_{\epsilon_1}' \smallsetminus \mathcal{Z}_{\epsilon}'} |\nabla (\xi_\epsilon^+ - \xi_\epsilon^-)|^2 \varrho_\epsilon^q dx$$

$$\leq \frac{16 K_2^q \vartheta^2}{\Xi_0} |\mathcal{Z}_{\epsilon_1}' \smallsetminus \mathcal{Z}_\epsilon'| \epsilon^{q-2\beta} \leq \Xi \epsilon^{q-\beta},$$

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

6.5. Proof of Theorem 3.8 (Third Eigenvalue of $\mathcal{L}_\epsilon$). We prove Theorem 3.8 in a series of propositions and corollaries. The upper and lower bounds on $\sigma_{3,\epsilon}$ in the cases $q \leq p + r$ and $q > p + r$ respectively follow from Propositions 6.6 and 6.7. 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 6.4. *Let $(p, q, r) \in \mathbb{R}^3$ satisfying $p + r > 0$ and $q > 0$, and suppose Assumptions 1, 2, and 3 hold. Let*

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

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 2.7 is not satisfied in $\mathcal{Z}_{\epsilon_0}^+$ or $\mathcal{Z} \smallsetminus \mathcal{Z}_{\epsilon_0}^+$. 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

$$\int_{\mathcal{Z}} \left| \nabla \left(\frac{u}{\varrho_\epsilon^r} \right) \right|^2 \varrho_\epsilon^q dx = \int_{\mathcal{Z}_{\epsilon_0}^+} \left| \nabla \left(\frac{u}{\varrho_\epsilon^r} \right) \right|^2 \varrho_\epsilon^q dx + \int_{\mathcal{Z} \smallsetminus \mathcal{Z}_{\epsilon_0}^+} \left| \nabla \left(\frac{u}{\varrho_\epsilon^r} \right) \right|^2 \varrho_\epsilon^q dx$$

$$\geq \Lambda_\epsilon(\mathcal{Z}_{\epsilon_0}^+) \int_{\mathcal{Z}_{\epsilon_0}^+} \left| \frac{u}{\varrho_\epsilon^r} - \bar{u}_{\mathcal{Z}_{\epsilon_0}^+} \right|^2 \varrho_\epsilon^{p+r} dx$$

$$+ \Lambda_\epsilon(\mathcal{Z} \smallsetminus \mathcal{Z}_{\epsilon_0}^+) \int_{\mathcal{Z} \smallsetminus \mathcal{Z}_{\epsilon_0}^+} \left| \frac{u}{\varrho_\epsilon^r} - \bar{u}_{\mathcal{Z} \smallsetminus \mathcal{Z}_{\epsilon_0}^+} \right|^2 \varrho_\epsilon^{p+r} dx,$$

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

(6.8)
$$\bar{u}_\Omega := \frac{1}{|\Omega|_{\varrho_\epsilon^{p+r}}} \int_\Omega \left(\frac{u}{\varrho_\epsilon^r} \right) \varrho_\epsilon^{p+r} dx.$$

After expanding the squared absolute values and rearrangement we get

(6.9)
$$\frac{1}{\Lambda(\epsilon)} \int_{\mathcal{Z}} \left| \nabla \left(\frac{u}{\varrho_\epsilon^r} \right) \right|^2 \varrho_\epsilon^q dx \geq \int_{\mathcal{Z}} \left| \frac{u}{\varrho_\epsilon^r} \right|^2 \varrho_\epsilon^{p+r} dx$$

$$+ \bar{u}_{\mathcal{Z}_{\epsilon_0}^+}^2 |\mathcal{Z}_{\epsilon_0}^+|_{\varrho_\epsilon^{p+r}} + \bar{u}_{\mathcal{Z} \smallsetminus \mathcal{Z}_{\epsilon_0}^+}^2 |\mathcal{Z} \smallsetminus \mathcal{Z}_{\epsilon_0}^+|_{\varrho_\epsilon^{p+r}}$$

$$- 2\bar{u}_{\mathcal{Z}_{\epsilon_0}^+} \int_{\mathcal{Z}_{\epsilon_0}^+} u\varrho_\epsilon^p dx - 2\bar{u}_{\mathcal{Z} \smallsetminus \mathcal{Z}_{\epsilon_0}^+} \int_{\mathcal{Z} \smallsetminus \mathcal{Z}_{\epsilon_0}^+} u\varrho_\epsilon^p dx.$$

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

$$\frac{1}{\Lambda(\epsilon)} \int_{\mathcal{Z}} \left| \nabla \left(\frac{u}{\varrho_\epsilon^r} \right) \right|^2 \varrho_\epsilon^q dx \geq \int_{\mathcal{Z}} \left| \frac{u}{\varrho_\epsilon^r} \right|^2 \varrho_\epsilon^{p+r} dx$$

$$- 2 \left(\bar{u}_{\mathcal{Z}_{\epsilon_0}^+} \int_{\mathcal{Z}_{\epsilon_0}^+} u\varrho_\epsilon^p dx + \bar{u}_{\mathcal{Z} \smallsetminus \mathcal{Z}_{\epsilon_0}^+} \int_{\mathcal{Z} \smallsetminus \mathcal{Z}_{\epsilon_0}^+} u\varrho_{\epsilon_0}^p dx \right).$$

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

$$\frac{1}{\Lambda(\epsilon)} \int_{\mathcal{Z}} \left| \nabla \left(\frac{u}{\varrho_\epsilon^r} \right) \right|^2 \varrho_\epsilon^q dx$$

$$\geq \left[1 - 2 \left(\bar{u}_{\mathcal{Z}_{\epsilon_0}^+} |\mathcal{Z}_{\epsilon_0}^+|_{\varrho_\epsilon^{p+r}}^{1/2} + \bar{u}_{\mathcal{Z} \smallsetminus \mathcal{Z}_{\epsilon_0}^+} |\mathcal{Z} \smallsetminus \mathcal{Z}_{\epsilon_0}^+|_{\varrho_\epsilon^{p+r}}^{1/2} \right) \right] \int_{\mathcal{Z}} \left| \frac{u}{\varrho_\epsilon^r} \right|^2 \varrho_\epsilon^{p+r} dx$$

$$=: [1 - 2(T_1 + T_2)].$$

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

(6.10)
$$\int_{\mathcal{Z}_{\epsilon_0}^+} u\varrho_\epsilon^p dx + \int_{\mathcal{Z}_{\epsilon_0}^-} u\varrho_\epsilon^p dx = - \int_{\mathcal{Z} \smallsetminus \mathcal{Z}_{\epsilon_0}'} u\varrho_\epsilon^p dx.$$

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

$$0 = b_\epsilon^+ \int_{\mathcal{Z}} u \xi_\epsilon^+ \varrho_\epsilon^p dx + b_\epsilon^+ \int_{\mathcal{Z}} \varrho_\epsilon^q \nabla\left(\frac{u}{\varrho_\epsilon^r}\right) \cdot \nabla \xi_\epsilon^+ dx$$

$$- b_\epsilon^- \int_{\mathcal{Z}} u \xi_\epsilon^- \varrho_\epsilon^p dx - b_\epsilon^- \int_{\mathcal{Z}} \varrho_\epsilon^q \nabla\left(\frac{u}{\varrho_\epsilon^r}\right) \cdot \nabla \xi_\epsilon^- dx.$$

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

$$\begin{aligned}
(6.11) \qquad \int_{\mathcal{Z}_{\epsilon_0}^+} u \varrho_\epsilon^p dx - \int_{\mathcal{Z}_{\epsilon_0}^-} u \varrho_\epsilon^p dx &= \int_{\mathcal{Z}_{\epsilon_0}^+ \smallsetminus \mathcal{Z}_\epsilon^+} u \varrho_\epsilon^p dx - \int_{\mathcal{Z}_{\epsilon_0}^- \smallsetminus \mathcal{Z}_\epsilon^-} u \varrho_\epsilon^p dx \\
&\quad - b_\epsilon^+ \int_{\mathcal{Z}_{\epsilon_1}^+ \smallsetminus \mathcal{Z}_\epsilon^+} u \varrho_\epsilon^p \xi_\epsilon^+ dx + b_\epsilon^- \int_{\mathcal{Z}_{\epsilon_1}^- \smallsetminus \mathcal{Z}_\epsilon^-} u \varrho_\epsilon^p \xi_\epsilon^- dx \\
&\quad + (1 - b_\epsilon^+) \int_{\mathcal{Z}_\epsilon^+} u \varrho_\epsilon^p dx - (1 - b_\epsilon^-) \int_{\mathcal{Z}_\epsilon^-} u \varrho_\epsilon^p dx \\
&\quad - b_\epsilon^+ \int_{\mathcal{Z}_{\epsilon_1}^+ \smallsetminus \mathcal{Z}_\epsilon^+} \varrho_\epsilon^q \nabla\left(\frac{u}{\varrho_\epsilon^r}\right) \cdot \nabla \xi_\epsilon^+ dx + b_\epsilon^- \int_{\mathcal{Z}_{\epsilon_1}^- \smallsetminus \mathcal{Z}_\epsilon^-} \varrho_\epsilon^q \nabla\left(\frac{u}{\varrho_\epsilon^r}\right) \cdot \nabla \xi_\epsilon^- dx.
\end{aligned}$$

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

$$(6.12) \qquad \left| \int_{\mathcal{Z}_{\epsilon_1}^\pm \smallsetminus \mathcal{Z}_\epsilon^\pm} \varrho_\epsilon^q \nabla\left(\frac{u}{\varrho_\epsilon^r}\right) \cdot \nabla \xi_\epsilon^\pm dx \right|^2 \leq \mathcal{R}_\epsilon(u) \int_{\mathcal{Z}_{\epsilon_1}^\pm \smallsetminus \mathcal{Z}_\epsilon^\pm} |\nabla \xi_\epsilon^\pm|^2 \varrho_\epsilon^q dx \leq \Xi_1 \mathcal{R}_\epsilon(u) \epsilon^{q-2\beta}$$

for a constant $\Xi_1 > 0$ independent of ϵ. Combine (6.10), (6.11), (6.12), and the fact that $0 < b_\epsilon^\pm < 2$ to get

$$2\left| \int_{\mathcal{Z}_{\epsilon_0}^+} u \varrho_\epsilon^p dx \right| \leq \int_{\mathcal{Z} \smallsetminus \mathcal{Z}_\epsilon'} |u \varrho_\epsilon^p| dx + 2 \int_{\mathcal{Z}_{\epsilon_1}' \smallsetminus \mathcal{Z}_\epsilon'} |u \varrho_\epsilon^p| dx$$

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

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

$$2\left| \int_{\mathcal{Z}_{\epsilon_0}^+} u \varrho_\epsilon^p dx \right| \leq |\mathcal{Z} \smallsetminus \mathcal{Z}_\epsilon'|_{\varrho_\epsilon^{p+r}}^{1/2} + 2 \left| \mathcal{Z}_{\epsilon_1}' \smallsetminus \mathcal{Z}_\epsilon' \right|_{\varrho_\epsilon^{p+r}}^{1/2}$$

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

Furthermore, by Assumption 3(d) and (2.8),

$$|\mathcal{Z} \smallsetminus \mathcal{Z}_\epsilon'|_{\varrho_\epsilon^{p+r}} = K_2^{p+r} \epsilon^{p+r} |\mathcal{Z} \smallsetminus \mathcal{Z}_\epsilon'| \leq \Xi_2 \epsilon^{p+r},$$

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

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 (6.10), 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

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

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. $\square$

Next, we investigate the consequences of Proposition 6.4 for different parameter choices p, q and r. The main point of interest here is to analyze how the parameter $\Lambda(\epsilon)$ in (6.7) is controlled by ϵ. We will show in Propositions 6.6 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

$$(6.13) \qquad |\partial \Omega|_{\varrho_\epsilon^q} := \liminf_{\delta \downarrow 0} \frac{1}{\delta} \left[|\Omega_\delta|_{\varrho_\epsilon^q} - |\Omega|_{\varrho_\epsilon^q} \right].$$

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

$$(6.14) \qquad \mathcal{J}(\Omega, \varrho_\epsilon) := \frac{|\partial \Omega|_{\varrho_\epsilon^q}}{\min\{|\Omega|_{\varrho_\epsilon^{p+r}}, |\Omega' \setminus \Omega|_{\varrho_\epsilon^{p+r}}\}}.$$

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

31

LEMMA 6.5. *Let $(p, q, r) \in \mathbb{R}^3$, and suppose Assumptions 1, 2 and 3 hold. Let $\Omega' \subseteq \mathcal{Z}$. Fix $\epsilon \in (0, \epsilon_0)$. Assume there exist $h(\epsilon) > 0$ so that*

$$(6.15) \qquad\qquad h(\epsilon) \leq \inf_{\Omega} \mathcal{J}(\Omega, \varrho_\epsilon),$$

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 2.7 and (2.18) holds with $\Lambda_\epsilon(\Omega') \geq \frac{h(\epsilon)^2}{4}$.

PROPOSITION 6.6. *Let $r \in \mathbb{R}$, $q > 0$, $p + r > 0$, and suppose Assumptions 1, 2 and 3 hold. Then there exists $\Xi > 0$ independent of $\epsilon \in (0, \epsilon_0]$ so that*

$$\sigma_{3,\epsilon} \geq \Xi \epsilon^{\max\{0, 2(q-p-r)\}}.$$

Proof. By Proposition 6.4 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 $\Lambda_\epsilon(\mathcal{Z} \smallsetminus \mathcal{Z}_{\epsilon_0}^+)$ separately. We only consider $\Lambda_\epsilon(\mathcal{Z}_{\epsilon_0}^+)$ and note that the same argument can be repeated for $\Lambda_\epsilon(\mathcal{Z} \smallsetminus \mathcal{Z}_{\epsilon_0}^+)$ possibly with different constants.

We will find a lower bound on $\inf_\Omega \mathcal{J}(\Omega, \varrho_\epsilon)$ and use Lemma 6.5 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

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

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 $\mathcal{Z}_{\epsilon_0}^+ \smallsetminus \mathcal{Z}^+$ 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

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

32

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

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

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, implying that (6.15) holds with a uniform constant $h > 0$. By Lemma 6.5, we obtain uniform lower bounds on $\Lambda_\epsilon(\mathcal{Z}_{\epsilon_0}^\pm)$. Since in this case, also $\Lambda_\epsilon(\mathcal{Z} \smallsetminus \mathcal{Z}_{\epsilon_0}')$ is bounded away from zero, we conclude by Proposition 6.4. If $q > p+r$, then we have the lower bound $\mathcal{J} \geq \Xi_4\epsilon^{q-p-r}$ and Lemma 6.5 implies $\Lambda_\epsilon(\mathcal{Z}_{\epsilon_0}^+) \geq \Xi_4^2\epsilon^{2(q-p-r)}/4$ with a similar bound also holding for $\Lambda_\epsilon(\mathcal{Z} \smallsetminus \mathcal{Z}_{\epsilon_0}^+)$. Therefore, Proposition 6.4 yields the existence of a constant $\Xi > 0$ so that $\sigma_{3,\epsilon} \geq \Xi\epsilon^{2(q-p-r)}$ in this case. □

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

PROPOSITION 6.7. *Suppose the conditions of Proposition* 6.6 *are satisfied. If* $q > p+r$ *and* $\epsilon_0 > 0$ *is sufficiently small, then there exists a constant* $\Xi > 0$ *depending only on* $\Lambda_\Delta(\mathcal{Z} \smallsetminus \mathcal{Z}_{\epsilon_0}')$ *so that* $\forall (\epsilon, \beta) \in (0, \epsilon_0] \times (0, (q-p-r)/2)$,

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

Note that according to Definition 2.5, $\Lambda_\Delta(\mathcal{Z} \smallsetminus \mathcal{Z}_{\epsilon_0}')$ is the second eigenvalue of the standard Laplacian on $\mathcal{Z} \smallsetminus \mathcal{Z}_{\epsilon_0}'$.

Proof. We use a similar argument to the proof of Theorem 3.7 using the min-max theorem. Let $\tilde{\varphi}_2 \in H^1(\mathcal{Z} \smallsetminus \mathcal{Z}_{\epsilon_0}')$ denote the second eigenfunction of the standard Laplacian on $\mathcal{Z} \smallsetminus \mathcal{Z}_{\epsilon_0}'$, i.e., $\tilde{\varphi}_2 \perp \mathbf{1}_{\mathcal{Z} \smallsetminus \mathcal{Z}_{\epsilon_0}'}$ and

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

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 (6.4), we define a function $\tilde{\xi}_\epsilon$ (see Figure 6.1)

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

This allows us to define the function

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

The shift ensures that $\tilde{\varphi}_{F,\epsilon} \in V^1(\mathcal{Z} \smallsetminus \mathcal{Z}_{\epsilon_2}', \varrho_\epsilon)$. 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 $V^1(\mathcal{Z}, \varrho_\epsilon)$.

33

Now let $u \in \mathrm{span}\{\varphi_{F,\epsilon}, \tilde{\varphi}_{F,\epsilon}\}$. 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 Theorem 3.7 the existence of $\Xi_1 > 0$ so that

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

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 $H^1(\mathcal{Z} \smallsetminus \mathcal{Z}'_{\epsilon_2})$. Thus, following our assumptions on ϱ_ϵ we can write

$$
\begin{aligned}
\|\tilde{\varphi}_{F,\epsilon}\|^2_{L^2(\mathcal{Z}, \varrho_\epsilon^{p-r})} \mathcal{R}_\epsilon(\tilde{\varphi}_{F,\epsilon}) &= \int_{\mathcal{Z} \smallsetminus \mathcal{Z}'_{\epsilon_2}} \left| \nabla \left(\frac{\tilde{\varphi}_{F,\epsilon}}{\varrho_\epsilon^r} \right) \right|^2 \varrho_\epsilon^q dx \\
&= K_2^{q-2r} \epsilon^{q-2r} \int_{\mathcal{Z} \smallsetminus \mathcal{Z}'_{\epsilon_2}} \left| \nabla \left(\tilde{\xi}_\epsilon \tilde{\varphi}_2 \right) \right|^2 dx \\
&\le 2K_2^{q-2r} \epsilon^{q-2r} \left(\int_{\mathcal{Z} \smallsetminus \mathcal{Z}'_{\epsilon_2}} \left| \tilde{\xi}_\epsilon \nabla \tilde{\varphi}_2 \right|^2 dx + \int_{\mathcal{Z} \smallsetminus \mathcal{Z}'_{\epsilon_2}} \left| \tilde{\varphi}_2 \nabla \tilde{\xi}_\epsilon \right|^2 dx \right) \\
&\le 2K_2^{q-2r} \epsilon^{q-2r} \left(\int_{\mathcal{Z} \smallsetminus \mathcal{Z}'_{\epsilon_0}} \left| \nabla \tilde{\varphi}_2 \right|^2 dx + \int_{\mathcal{Z}'_{\epsilon_3} \smallsetminus \mathcal{Z}'_{\epsilon_2}} \left| \tilde{\varphi}_2 \nabla \tilde{\xi}_\epsilon \right|^2 dx \right) \\
&\le 2K_2^{q-2r} \epsilon^{q-2r} \left(\int_{\mathcal{Z} \smallsetminus \mathcal{Z}'_{\epsilon_0}} \left| \nabla \tilde{\varphi}_2 \right|^2 dx + \vartheta^2 \epsilon^{-2\beta} \int_{\mathcal{Z} \smallsetminus \mathcal{Z}'_{\epsilon_0}} \left| \tilde{\varphi}_2 \right|^2 dx \right) \\
&\le 2K_2^{q-2r} \epsilon^{q-2r} \left(\Lambda_\Delta(\mathcal{Z} \smallsetminus \mathcal{Z}'_{\epsilon_0}) + \vartheta^2 \epsilon^{-2\beta} \right) \|\tilde{\varphi}_2\|^2_{L^2(\mathcal{Z} \smallsetminus \mathcal{Z}'_{\epsilon_0})}.
\end{aligned}
$$
(6.17)

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

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

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

$$\int_{\mathcal{Z}'_{\epsilon_3} \smallsetminus \mathcal{Z}'_{\epsilon_0}} |\tilde{\varphi}_2|^2 dx \le \|\tilde{\varphi}_2\|^2_{L^k(\mathcal{Z} \smallsetminus \mathcal{Z}'_{\epsilon_0})} |\mathcal{Z}'_{\epsilon_3} \smallsetminus \mathcal{Z}'_{\epsilon_0}|^{\frac{k-2}{k}}.$$

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

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

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

$$
\begin{aligned}
\|\tilde{\xi}_\epsilon \tilde{\varphi}_2\|^2_{L^2(\mathcal{Z}, \varrho_\epsilon^{p-r})} &\ge K_2^{p-r} \epsilon^{p-r} \left(\|\tilde{\varphi}_2\|^2_{L^2(\mathcal{Z} \smallsetminus \mathcal{Z}'_{\epsilon_0})} - \Xi_2 \left(1 + \Lambda_\Delta(\mathcal{Z} \smallsetminus \mathcal{Z}'_{\epsilon_0}) \right) |\mathcal{Z}'_{\epsilon_3} \smallsetminus \mathcal{Z}'_{\epsilon_0}|^{\frac{k-2}{k}} \right) \\
&\ge K_2^{p-r} \epsilon^{p-r} \left(\|\tilde{\varphi}_2\|^2_{L^2(\mathcal{Z} \smallsetminus \mathcal{Z}'_{\epsilon_0})} - \Xi_4 \epsilon^{\frac{\beta(k-2)}{k}} \right).
\end{aligned}
$$
(6.18)

34

Furthermore, using Assumption 3(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

$$\frac{\varrho_\epsilon^r}{|\mathcal{Z} \setminus \mathcal{Z}'_{\epsilon_2}|_{\varrho_\epsilon^{p+r}}} \left| \int_{\mathcal{Z} \setminus \mathcal{Z}'_{\epsilon_2}} \tilde{\xi}_\epsilon \tilde{\varphi}_2 \varrho_\epsilon^p \, dx \right|$$

$$= \frac{1}{|\mathcal{Z} \setminus \mathcal{Z}'_{\epsilon_2}|} \left| \int_{\mathcal{Z} \setminus \mathcal{Z}'_{\epsilon_2}} \tilde{\xi}_\epsilon \tilde{\varphi}_2 \, dx \right|$$

$$= \frac{1}{|\mathcal{Z} \setminus \mathcal{Z}'_{\epsilon_2}|} \left| \int_{\mathcal{Z} \setminus \mathcal{Z}'_{\epsilon_0}} \tilde{\varphi}_2 \, dx - \int_{\mathcal{Z}'_{\epsilon_2} \setminus \mathcal{Z}'_{\epsilon_0}} \tilde{\varphi}_2 \, dx + \int_{\mathcal{Z}'_{\epsilon_3} \setminus \mathcal{Z}'_{\epsilon_2}} (\tilde{\xi}_\epsilon - 1) \tilde{\varphi}_2 \, dx \right|$$

$$\leq \frac{1}{|\mathcal{Z} \setminus \mathcal{Z}'_{\epsilon_2}|} \left(\int_{\mathcal{Z}'_{\epsilon_2} \setminus \mathcal{Z}'_{\epsilon_0}} |\tilde{\varphi}_2| \, dx + \int_{\mathcal{Z}'_{\epsilon_3} \setminus \mathcal{Z}'_{\epsilon_2}} |\tilde{\xi}_\epsilon - 1| \, |\tilde{\varphi}_2| \, dx \right)$$

(6.19)
$$\leq \frac{1}{|\mathcal{Z} \setminus \mathcal{Z}'_{\epsilon_2}|} \int_{\mathcal{Z}'_{\epsilon_3} \setminus \mathcal{Z}'_{\epsilon_0}} |\tilde{\varphi}_2| \, dx \leq \frac{|\mathcal{Z}'_{\epsilon_3} \setminus \mathcal{Z}'_{\epsilon_0}|^{1/2}}{|\mathcal{Z} \setminus \mathcal{Z}'_{\epsilon_0}| - |\mathcal{Z}'_{\epsilon_2} \setminus \mathcal{Z}'_{\epsilon_0}|} \leq \Xi_5 \epsilon^{\beta/2}.$$

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

$$\|\tilde{\varphi}_{F,\epsilon}\|^2_{L^2(\mathcal{Z}, \varrho_\epsilon^{p-r})}$$

$$= \int_{\mathcal{Z}} \left| \tilde{\xi}_\epsilon(x) \tilde{\varphi}_2(x) - \frac{\varrho_\epsilon^r}{|\mathcal{Z} \setminus \mathcal{Z}'_{\epsilon_2}|_{\varrho_\epsilon^{p+r}}} \int_{\mathcal{Z} \setminus \mathcal{Z}'_{\epsilon_2}} \tilde{\xi}_\epsilon(y) \tilde{\varphi}_2(y) \varrho_\epsilon^p(y) \, dy \right|^2 \varrho_\epsilon^{p-r}(x) \, dx$$

$$\geq \|\tilde{\xi}_\epsilon \tilde{\varphi}_2\|^2_{L^2(\mathcal{Z}, \varrho_\epsilon^{p-r})} - \frac{2}{|\mathcal{Z} \setminus \mathcal{Z}'_{\epsilon_2}|_{\varrho_\epsilon^{p+r}}} \left| \int_{\mathcal{Z} \setminus \mathcal{Z}'_{\epsilon_2}} \tilde{\xi}_\epsilon \tilde{\varphi}_2 \varrho_\epsilon^p \, dy \right|^2$$

$$= \|\tilde{\xi}_\epsilon \tilde{\varphi}_2\|^2_{L^2(\mathcal{Z}, \varrho_\epsilon^{p-r})} - \frac{|\mathcal{Z} \setminus \mathcal{Z}'_{\epsilon_2}|^{p+r}_{\varrho_\epsilon^{p+r}}}{\varrho_\epsilon^{2r}} \left(\frac{2\varrho_\epsilon^r}{|\mathcal{Z} \setminus \mathcal{Z}'_{\epsilon_2}|_{\varrho_\epsilon^{2(p+r)}}} \left| \int_{\mathcal{Z} \setminus \mathcal{Z}'_{\epsilon_2}} \tilde{\xi}_\epsilon \tilde{\varphi}_2 \varrho_\epsilon^p \, dy \right|^2 \right)$$

$$= \|\tilde{\xi}_\epsilon \tilde{\varphi}_2\|^2_{L^2(\mathcal{Z}, \varrho_\epsilon^{p-r})} - |\mathcal{Z} \setminus \mathcal{Z}'_{\epsilon_2}| K_2^{p-r} \epsilon^{p-r} \left(\frac{2\varrho_\epsilon^r}{|\mathcal{Z} \setminus \mathcal{Z}'_{\epsilon_2}|_{\varrho_\epsilon^{2(p+r)}}} \left| \int_{\mathcal{Z} \setminus \mathcal{Z}'_{\epsilon_2}} \tilde{\xi}_\epsilon \tilde{\varphi}_2 \varrho_\epsilon^p \, dy \right|^2 \right).$$

Together with the bounds (6.18) and (6.19), we obtain for small enough ϵ_0,

$$\|\tilde{\varphi}_{F,\epsilon}\|^2_{L^2(\mathcal{Z}, \varrho_\epsilon^{p-r})}$$

$$\geq K_2^{p-r} \epsilon^{p-r} \left(\|\tilde{\varphi}_2\|^2_{L^2(\mathcal{Z} \setminus \mathcal{Z}'_{\epsilon_0})} - \Xi_4 \epsilon^{\frac{\beta(k-2)}{k}} - \Xi_6 \epsilon^\beta \right)$$

$$\geq \Xi_7 \epsilon^{p-r} \|\tilde{\varphi}_2\|^2_{L^2(\mathcal{Z} \setminus \mathcal{Z}'_{\epsilon_0})}.$$

Finally, following from (6.17), 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. □

Proof of Theorem 3.8. The case $q > p + r$ in Theorem 3.8(i) is a consequence of Propositions 6.6 and 6.7, whereas case $q \leq p + r$ in Theorem 3.8(ii) follows directly from Proposition 6.6. □

6.6. Proof of Theorem 3.10 (Geometry Of The Second Eigenfunction $\varphi_{2,\epsilon}$). 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 6.8. *Suppose there exist constants $\Xi_1 \geq 0$ and $\Xi_2 \geq 0$ not both zero, so that for all $\epsilon \in (0, \epsilon_0]$,*

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

Suppose $\vartheta < q$ in the case $\Xi_2 > 0$. Then for every $0 < \beta < q$, there exists $\Xi > 0$ so that

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

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 [18, 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 (6.6) 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 $h_j^2 \leq 1$ or all $j \geq 1$ and

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

$\square$

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

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

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 6.9. *Let $(p,q,r) \in \mathbb{R}^3$ satisfying $p+r > 0$, and suppose Assumptions 1, 2 and 3 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 (6.20) write

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

It follows from Assumption 3(c) that

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

Combining with Assumption 2(c), we can find a constant $\Xi_3 > 0$ so that

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

Let $b_\epsilon^\pm$ be as in (6.2). Using Assumption 3(d), and the definition of the $\chi_\epsilon^\pm$, we can write

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

The first term is close to $\int_{\mathcal{Z}'} \varrho_0^{p+r} dx$ using (6.21), whereas the terms in the second line can be controlled using Lemma 6.3 and the fact that $0 < b_\epsilon^\pm < 2$,

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

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

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

37

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

$$(6.23) \qquad \left| (b_\epsilon^F)^{-2} - \int_{\mathcal{Z}'} \varrho_0^{p+r}\, dx \right| \le \Xi_6 \epsilon^{\min\{1, p+r\}} .$$

The lemma then follows from (6.22) and (6.23). $\qquad\qquad\square$

In order to prove Theorem 3.10, 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 6.10. *Let $(p, q, r) \in \mathbb{R}^3$ satisfying $p + r > 0$, and suppose Assumptions 1, 2 and 3 hold. Let $\epsilon_0 > 0$ small enough. Then there exists a constant $\Xi > 0$, independent of ϵ so that for all $\epsilon \in (0, \epsilon_0)$,*

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

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

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

If $p \ge 0$ (and by a similar argument with the order of inequalities reversed if $p < 0$), (6.21) implies

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

By Assumption 2(c), we conclude that there exists a constant $\Xi_1 > 0$ such that

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

The above estimate together with Lemma 6.3 implies

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

for some constant $\Xi_2 > 0$. Combining this bound with Lemma 6.9, and writing

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

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

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

Finally, we obtain

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

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 3.10. We apply Proposition 6.8 with the eigenvalue bounds in Theorem 3.8 and observe that the required lower bound on $\sigma_{3,\epsilon}$ in Proposition 6.8 holds with $\Xi_1 = 0$, $\Xi_2 > 0$ and $\vartheta = -q + 2(p+r)$ if $q > p + r$, and with $\Xi_1 > 0$, $\Xi_2 = 0$ in the case $q \leq p + r$. We obtain that there exists a constant $\Xi_3 > 0$ so that

$$(6.24) \qquad \left|1 - \langle \varphi_{2,\epsilon}, \varphi_{F,\epsilon} \rangle^2_{\varrho_\epsilon^{p-r}}\right| \leq \Xi_3 \epsilon^{q-2\max\{0,q-p-r\}-\beta},$$

for any $(p,q,r) \in \mathbb{R}^3$ with $q > 0$ and $p + r > 0$. Combining estimate (6.24) with Proposition 6.10 gives

$$\left|1 - \langle \varphi_{2,\epsilon}, \bar{\varphi}_{2,0} \rangle^2_{\varrho_\epsilon^{p-r}}\right|$$

$$= \left|1 - \left(\langle \varphi_{2,\epsilon}, \varphi_{F,\epsilon} \rangle_{\varrho_\epsilon^{p-r}} + \langle \varphi_{2,\epsilon}, \bar{\varphi}_{2,0} - \varphi_{F,\epsilon} \rangle_{\varrho_\epsilon^{p-r}}\right)^2\right|$$

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

$$\leq \left|1 - \langle \varphi_{2,\epsilon}, \varphi_{F,\epsilon} \rangle^2_{\varrho_\epsilon^{p-r}}\right|$$

$$\qquad + \|\varphi_{2,\epsilon}\|^2_{L^2(\mathcal{Z},\varrho_\epsilon^{p-r})} \|\bar{\varphi}_{2,0} - \varphi_{F,\epsilon}\|_{L^2(\mathcal{Z},\varrho_\epsilon^{p-r})} \left(\|\bar{\varphi}_{2,0}\|_{L^2(\mathcal{Z},\varrho_\epsilon^{p-r})} + \|\varphi_{F,\epsilon}\|_{L^2(\mathcal{Z},\varrho_\epsilon^{p-r})}\right)$$

$$\leq \Xi_3 \epsilon^{q-\max\{0,2(q-p-r)\}-\beta} + \Xi_4 \epsilon^{\frac{1}{2}\min\{1,p+r\}}$$

$$\leq \Xi \epsilon^{\min\{\frac{1}{2},\frac{p+r}{2},q-\beta,q-2(q-p-r)-\beta\}}$$

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

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 demonstrated that the family (1.1) of elliptic operators $\mathcal{L}$ arise naturally as continuum limits of graph Laplacians L_N of the form (1.2); and provided a roadmap for rigorous proof of convergence of L_N to $\mathcal{L}$ as $N \to \infty$, in the framework of [21].

We then characterized the sensitive dependence of the low-lying spectrum of $\mathcal{L}$ with respect to the parameters p, q, r in cases where the density ϱ concentrates on two clusters. In particular, we showed that there is a major change in the behavior of the spectrum of $\mathcal{L}$ when $q \leq p + r$ versus $q > p + r$. 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 manifests. This is of interest from a practical point of view as the low-lying spectral properties govern many unsupervised and semi-supervised clustering tasks.

We also demonstrated a 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 more on two clusters the span of the first two eigenfunctions of $\mathcal{L}$ approaches certain weighted set functions on the clusters.

Finally, 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 $q \leq p+r$. In the $q > p+r$ regime our numerics indicate that our lower bounds on the third eigenvalues, and hence on the spectral ratio gap, can be sharpened. Furthermore, 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 finitely many clusters.

Our work may be of independent interest within the spectral theory of elliptic operators. Furthermore it will be used in our upcoming publication [23] to build on the paper [24], which studies consistency of semi-supervised learning on graphs, to develop a consistency theory for semi-supervised learning in the continuum limit.

<div style="text-align:center">REFERENCES</div>

[1] R. A. Adams and J. J. Fournier. *Sobolev spaces*, volume 140. Elsevier, 2003.

[2] D. Bakry, I. Gentil, and M. Ledoux. *Analysis and geometry of Markov diffusion operators*, volume 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, and H. Zhang. PETSc users manual. Technical Report ANL-95/11 - Revision 3.11, Argonne National Laboratory, 2019.

[4] M. Belkin and P. Niyogi. Laplacian eigenmaps for dimensionality reduction and data representation. *Neural computation*, 15(6):1373–1396, 2003.

[5] M. Belkin and P. Niyogi. Convergence of laplacian eigenmaps. In *NIPS*, 2006.

[6] M. Belkin and P. Niyogi. Towards a theoretical foundation for laplacian-based manifold methods. *Journal of Computer and System Sciences*, 74(8):1289–1308, 2008.

[7] T. Berry and J. Harlim. Variable bandwidth diffusion kernels. *Applied and Computational Harmonic Analysis*, 40(1):68–96, 2016.

[8] A. L. Bertozzi and A. Flenner. Diffuse interface models on graphs for classification of high dimensional data. *Multiscale Modeling and Simulation*, 10(3):1090–1118, 2012.

[9] A. L. Bertozzi, X. Luo, A. M. Stuart, and K. C. Zygalakis. Uncertainty quantification in graph-based classification of high dimensional data. *SIAM/ASA Journal on Uncertainty Quantification*, 6(2):568–595, 2018.

[10] A. Bovier, M. Eckhoff, V. Gayrard, and M. Klein. Metastability in reversible diffusion processes i: Sharp asymptotics for capacities and exit times. *Journal of the European Mathematical Society*, 6(4):399–424, 2004.

[11] A. Bovier, V. Gayrard, and M. Klein. Metastability in reversible diffusion processes ii: Precise asymptotics for small eigenvalues. *Journal of the European Mathematical Society*, 7(1):69–99, 2005.

[12] R. R. Coifman, I. G. Kevrekidis, S. Lafon, M. Maggioni, and B. Nadler. Diffusion maps, reduction coordinates, and low dimensional representation of stochastic systems. *Multiscale Modeling and Simulation*, 7(2):842–864, 2008.

[13] R. R. Coifman and S. Lafon. Diffusion maps. *Appl. Comput. Harmon. Anal.*, 21(1):5–30, 2006.

[14] R. R. Coifman, S. Lafon, A. B. Lee, M. Maggioni, B. Nadler, F. Warner, and S. W. Zucker. Geometric diffusions as a tool for harmonic analysis and structure definition of data: Diffusion maps. *Proceedings of the National Academy of Sciences*, 102(21):7426–7431, 2005.

[15] P. Deuflhard, M. Dellnitz, O. Junge, and C. Schütte. Computation of essential molecular dynamics by subdivision techniques. In *Computational molecular dynamics: challenges, methods, ideas*, pages 98–115. Springer, 1999.

[16] P. Deuflhard, W. Huisinga, A. Fischer, and C. Schütte. Identification of almost invariant aggregates in reversible nearly uncoupled Markov chains. *Linear Algebra and its Applications*, 315(1-3):39–59, 2000.

[17] M. M. Dunlop, D. Slepev, A. M. Stuart, and M. Thorpe. Large data and zero noise limits of graph-based semi-supervised learning algorithms. *Applied and Computational Harmonic Analysis*, 2019.

[18] L. C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*.

AMS, Providence, RI, second edition, 2010.

[19] N. García Trillos, M. Gerlach, M. Hein, and D. Slepčev. Error estimates for spectral convergence of the graph laplacian on random geometric graphs towards the Laplace–Beltrami operator. *arXiv preprint arXiv:1801.10108*, 2018.

[20] N. García Trillos, F. Hoffmann, and B. Hosseini. Geometric structure of graph laplacian embeddings. *arXiv preprint arXiv:1901.10651*, 2019.

[21] N. García Trillos and D. Slepčev. A variational approach to the consistency of spectral clustering. *Applied and Computational Harmonic Analysis*, 45(2):239–281, 2018.

[22] E. Giné, V. Koltchinskii, et al. Empirical graph laplacian approximation of laplace–beltrami operators: Large sample results. In *High dimensional probability*, pages 238–259. Institute of Mathematical Statistics, 2006.

[23] F. Hoffmann, B. Hosseini, A. Oberai, and A. Stuart. Consistency of graphical semi-supervised learning algorithms in the continuum limit: The probit method. In preparation, 2019.

[24] F. Hoffmann, B. Hosseini, Z. Ren, and A. M. Stuart. Consistency of semi-supervised learning algorithms on graphs: probit and one-hot methods. *arXiv preprint:1906.07658*, 2019.

[25] W. Huisinga, S. Meyn, and C. Schütte. Phase transitions and metastability in Markovian and molecular systems. *The Annals of Applied Probability*, 14(1):419–458, 2004.

[26] T. Kato. *Perturbation theory for linear operators*. Classics In Mathematics. Springer, New York, second edition, 1995.

[27] D. O. Loftsgaarden, C. P. Quesenberry, et al. A nonparametric estimate of a multivariate density function. *The Annals of Mathematical Statistics*, 36(3):1049–1051, 1965.

[28] A. Logg, K.-A. Mardal, and G. Wells. *Automated solution of differential equations by the finite element method: The FEniCS book*, volume 84 of *Lecture Notes in Computational Science and Engineering*. Springer Science & Business Media, 2012.

[29] W. McLean. *Strongly elliptic systems and boundary integral equations*. Cambridge University Press, Cambridge, 2000.

[30] B. Nadler, S. Lafon, R. R. Coifman, and I. G. Kevrekidis. Diffusion maps, spectral clustering and reaction coordinates of dynamical systems. *Applied and Computational Harmonic Analysis*, 21(1):113–127, 2006.

[31] A. Y. Ng, M. I. Jordan, and Y. Weiss. On spectral clustering: Analysis and an algorithm. In *Proceedings of the 14th International Conference on Neural Information Processing Systems: Natural and Synthetic*.

[32] G. A. Pavliotis. *Stochastic processes and applications: diffusion processes, the Fokker-Planck and Langevin equations*, volume 60 of *Texts in Applied Mathematics*. Springer, New York, 2014.

[33] K. Rohe, S. Chatterjee, and B. Yu. Spectral clustering and the high-dimensional stochastic blockmodel. *The Annals of Statistics*, 39(4):1878–1915, 2011.

[34] C. Schütte, W. Huisinga, and P. Deuflhard. Transfer operator approach to conformational dynamics in biomolecular systems. In F. Bernold, editor, *Ergodic theory, analysis, and efficient simulation of dynamical systems*, pages 191–223. Springer, Berlin, 2001.

[35] J. Shi and J. Malik. Normalized cuts and image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(8):888–905, Aug. 2000.

[36] D. Slepčev and M. Thorpe. Analysis of p-Laplacian regularization in semisupervised learning. *SIAM Journal on Mathematical Analysis*, 51(3):2085–2120, 2019.

[37] D. A. Spielman and S.-H. Teng. Spectral partitioning works: Planar graphs and finite element meshes. *Linear Algebra and its Applications*, 421(2-3):284–305, 2007.

[38] D. A. Spielmat and S.-H. Teng. Spectral partitioning works: Planar graphs and finite element meshes. In *Proceedings of 37th Conference on Foundations of Computer Science*, pages 96–105. IEEE, 1996.

[39] G. R. Terrell and D. W. Scott. Variable kernel density estimation. *The Annals of Statistics*, 20(3):1236–1265, 1992.

[40] U. von Luxburg. A tutorial on spectral clustering. *Statistics and Computing*, 17(4):395–416, 2007.

[41] U. von Luxburg, M. Belkin, and O. Bousquet. Consistency of spectral clustering. *The Annals of Statistics*, 36(2):555–586, 2008.

[42] L. Zelnik-Manor and P. Perona. Self-tuning spectral clustering. In *Advances in neural information processing systems*, pages 1601–1608, 2005.

[43] X. Zhu, Z. Ghahramani, and J. D. Lafferty. Semi-supervised learning using Gaussian fields and harmonic functions. In *Proceedings of the 20th International conference on Machine learning*, pages 912–919, 2003.

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 [13] 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,

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

To make the connection between this discrete setting and the continuous analogue above, we use the degree functions of Subsection 5.4,

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

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

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 [13], 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 [13, 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 [13], 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 [32]. 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.,

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

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 $p \ge 0$. Then*

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

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*

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

Proof. Since ϱ satisfies (B.1) then

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

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

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 [18, 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 Theorem.
The min-max theorem [26, Ch. 1 Sec. 6.10] is readily applied to our specific setting to obtain the following:

THEOREM 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 $\varrho \in C^\infty(\bar{\Omega})$ satisfying Assumption 2, 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

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

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

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

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

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),

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

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

THEOREM 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*

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

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

$$\frac{h^2}{4} \int_{\Omega'} |f - \bar{f}_{\Omega'}|^2 d\nu \leq \int_{\Omega'} |\nabla f|^2 d\mu,$$

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',

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

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 $t \ge 0$,

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

which together with (D.2) gives

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

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,

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

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

$$\frac{h^2}{4} \int_{\Omega'} |f - m|^2 \, d\nu \le \int_{\Omega'} |\nabla f|^2 \, d\mu.$$

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. $\qquad \square$

Proof of Lemma 6.5. 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

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

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. $\qquad \square$

		$\frac{\log(\sigma_{2,\epsilon})}{\log\epsilon}$		$\frac{\log(\sigma_{2,\epsilon})-\log(\sigma_{3,\epsilon})}{\log\epsilon}$	
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

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 Theorems 3.7 and 3.8 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	q	r	Analytic	Numerical	Analytic	Numerical	$p+r$
0.5	1.50	0.5	1.50	1.51	0.50	0.99	1.00
0.5	2.0	0.5	2.00	2.00	0.00	0.99	1.00
0.5	2.5	0.5	2.49	2.57	-	0.99	1.00
0.5	3.0	0.5	2.96	3.06	-	0.97	1.00
1.0	2	0.5	2.03	2.11	1.00	1.52	1.50
1.5	2.5	0.5	2.54	2.64	1.50	2.03	2.00
2.0	3.0	0.5	3.05	3.20	2.00	2.53	2.50

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 Theorems 3.7 and 3.9. The last column denotes the conjectured slope of $p + r$ for $\log(\sigma_{2,\epsilon}/\sigma_{3,\epsilon})$.

	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	1.0	0.5	1.04	1.00
	0.5	1.5	0.5	1.53	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

TABLE 4.3

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.
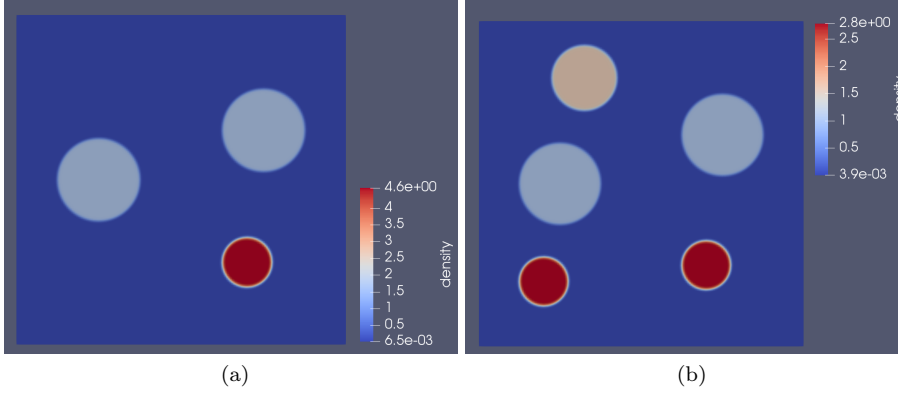
FIG. 4.1. *Plot of the densities ϱ_ϵ of the form* (4.2) *with three and five clusters.*

	p	q	r	$\dfrac{\log(\sigma_{4,\epsilon})}{\log \epsilon}$	$\dfrac{\log(\sigma_{5,\epsilon})-\log(\sigma_{4,\epsilon})}{\log \epsilon}$
$q = p + r$	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	1.0	0.5	1.04	1.00
	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

TABLE 4.4

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.
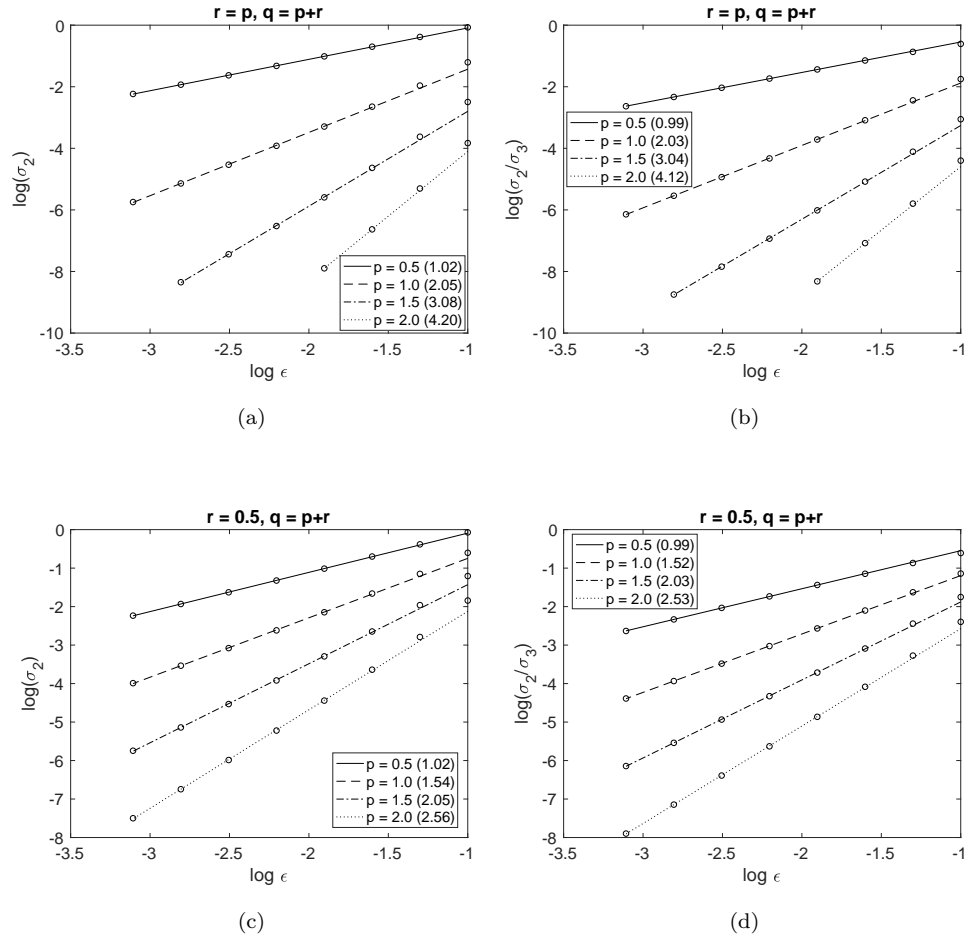
FIG. 4.2. *Variation of the second and third eigenvalues of $\mathcal{L}_\epsilon$ in the balanced case with $q = p + r$ and for various values of $p \in [0.5, 2]$. (a, b) consider $r = p$; (c, d) consider fixed $r = 0.5$. (a, c) show $\log(\sigma_{2,\epsilon})$ vs $\log(\epsilon)$ while (b, d) show $\log(\sigma_{2,\epsilon}/\sigma_{3,\epsilon})$ 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.*
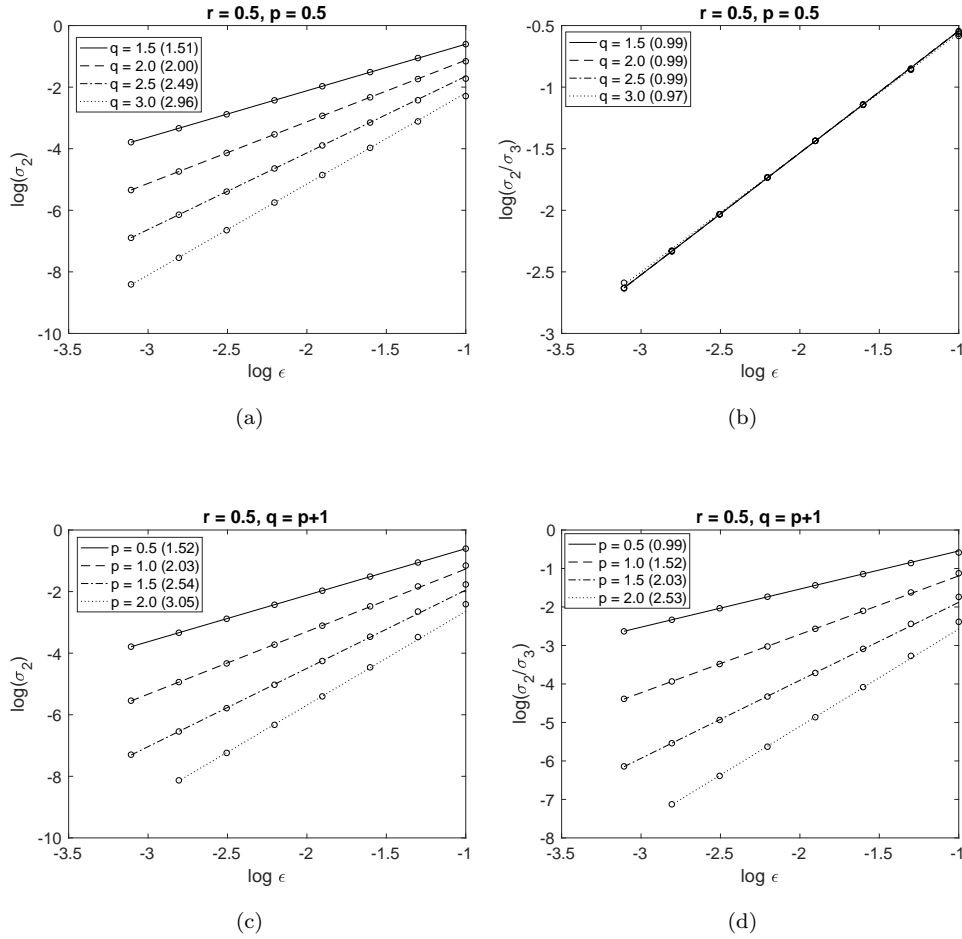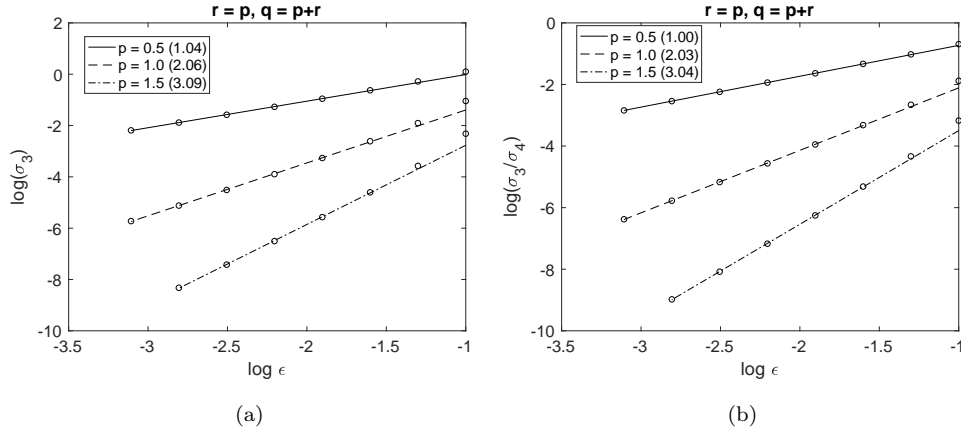
(a)

(b)

(c)

(d)

Fig. 4.3. *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.*

FIG. 4.4. *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.*
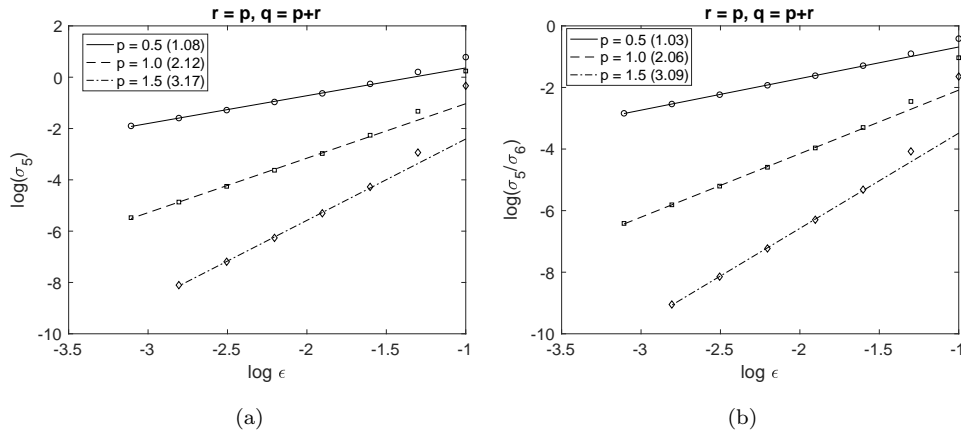

FIG. 4.5. *Variation of the fifth and sixth eigenvalues of $\mathcal{L}_\epsilon$ in the five cluster case with $q = p + r$, $r = p$ and for $p \in [0.5, 1.5]$. (a) shows $\log(\sigma_{5,\epsilon})$ vs $\log(\epsilon)$ while (b) shows $\log(\sigma_{5,\epsilon}/\sigma_{6,\epsilon})$ vs $\log(\epsilon)$. The values reported in the brackets in the legends are numerical approximations to the slope of the lines for different values of p.*
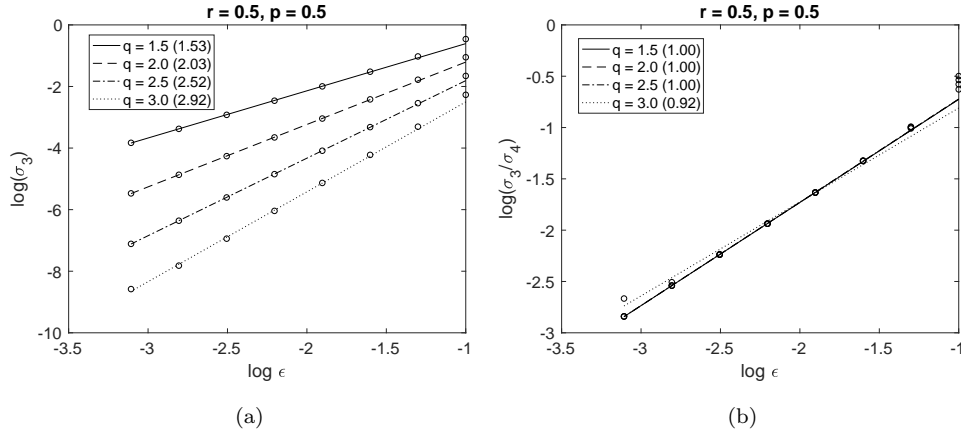
(a)   (b)

FIG. 4.6. *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.*
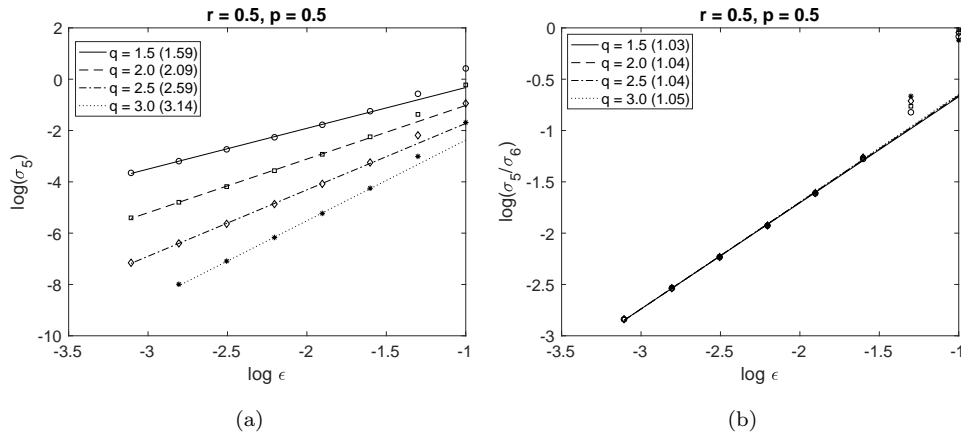

(a)   (b)

FIG. 4.7. *Variation of the fifth and sixth eigenvalues of $\mathcal{L}_\epsilon$ in the five cluster case with $q > p + r$, $r = p = 0.5$ 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.*