Supplementary Material for ”Spectral Clustering based on the graph p-Laplacian”

Thomas Bühler and Matthias Hein
Saarland University, Saarbrücken, Germany
{tb,hein}@cs.uni-sb.de

May 2009
Corrected version, June 2010

Abstract

This technical report is an appendix to the ICML 2009 submission ”Spectral Clustering based on the graph p-Laplacian” [2], containing the proofs which had to be omitted due to space restrictions. In this version, an error occuring in a previous version has been corrected. However, note that the error occurred in an additional statement which was not used in any proof, thus the correctness of the other results is not affected.

1 Overview

This technical report is an appendix to the ICML 2009 submission ”Spectral Clustering based on the graph p-Laplacian” [2], containing the proofs which had to be omitted due to space restrictions. Our proposed method and some of the results are based on the recent work by Amghibech [1]. In his very interesting article, the author proposes the variational characterization of the second eigenvector of the normalized graph p-Laplacian and derives the isoperimetric inequality in the normalized case. Due to the compressed form of the proof given in [1], some important lemmas which play a role in the proof of the variational characterization are not explicitely stated in the paper, hence we cover the proofs in more detail in the following. Moreover, we provide basic properties of the p-Laplacian and related functionals and extend the results of Amghibech to the unnormalized case. However, our main result is to show that the Cheeger cuts obtained by thresholding the second eigenvector converge to the optimal Cheeger cut as $p \to 1$.

This paper is organized as follows: We start with some basic properties of the graph p-Laplacian and related functionals in Section 2. In Section 3 we prove the variational characterization of the second eigenvalue of the unnormalized graph p-Laplacian (Theorem 3.2 in [2]), and Section 4 contains the corresponding characterizations for the first and second eigenvector in the normalized case. Section 5 contains the proof of the isoperimetric inequality from Theorem 4.3 in [2], and Section 6 an outline of the proof in the normalized case. Finally, Section 7 establishes the connection between the optimal Cheeger cut and the cut obtained by thresholding the second eigenvector (Theorem 4.4 in [2]).

As in [2], the number of points is denoted by $n = |V|$ and the complement of a set $A \subset V$ is written as $\overline{A} = V \backslash A$. The degree function $d : V \to \mathbb{R}$ of the graph is given as $d_i = \sum_{j=1}^{n} w_{ij}$ and the cut of $A \subset V$ and $B \subset V$, with $A \cap B = \emptyset$, is defined as $\mathrm{cut}(A, B) = \sum_{i \in A,\, j \in B} w_{ij}$. Moreover, we denote by $|A|$ the cardinality of the set A and by $\mathrm{vol}(A) = \sum_{i \in A} d_i$ the volume of A.

2 The graph p-Laplacian and related functionals

The **unnormalized** and **normalized** graph p-Laplacian $\Delta_p^{(\mathrm{u})}$ and $\Delta_p^{(\mathrm{n})}$ are defined for any function $f : V \to \mathbb{R}$ and $i \in V$ as

$$(\Delta_p^{(\mathrm{u})} f)_i = \sum_{j \in V} w_{ij}\, \phi_p\left(f_i - f_j\right) ,$$
$$(\Delta_p^{(\mathrm{n})} f)_i = \frac{1}{d_i} \sum_{j \in V} w_{ij}\, \phi_p\left(f_i - f_j\right) ,$$

where $\phi_p : \mathbb{R} \to \mathbb{R}$ with $\phi_p(x) = |x|^{p-1} \operatorname{sign}(x)$. As shown in [2], one can obtain the eigenvalues of the **unnormalized** p-Laplacian $\Delta_p^{(\mathrm{u})}$ as local minima of the functional $F_p : \mathbb{R}^V \to \mathbb{R}$,

$$F_p(f) = \frac{Q_p(f)}{\|f\|_p^p}$$

where

$$Q_p(f) = \left\langle f, \Delta_p^{(\mathrm{u})} f \right\rangle = \frac{1}{2} \sum_{i,j \in V} w_{ij} |f_i - f_j|^p , \tag{1}$$

and each critical point of F_p corresponds to an eigenvector of the p-Laplacian. To obtain the second eigenvalue, we consider the functional $F_p^{(2)} : \mathbb{R}^V \to \mathbb{R}$,

$$F_p^{(2)}(f) = \frac{Q_p(f)}{\mathrm{var}_p^{(\mathrm{u})}(f)} ,$$

in analogy to the functional defined by Amghibech in [1]. The **unnormalized** p-variance $\mathrm{var}_p^{(\mathrm{u})}(f)$ is defined as

$$\mathrm{var}_p^{(\mathrm{u})}(f) = \min_{m \in \mathbb{R}} \left\{ \sum_{i \in V} |f_i - m|^p \right\} . \tag{2}$$

Furthermore, we define the **unnormalized** p-mean of f as

$$\mathrm{mean}_p^{(\mathrm{u})}(f) = \operatorname*{arg\,min}_{m \in \mathbb{R}} \left\{ \sum_{i \in V} |f_i - m|^p \right\} .$$

In Theorem 3.1 in Section 3 we show that the global minimum of the functional $F_p^{(2)}$ is equal to the second eigenvalue of the graph p-Laplacian $\Delta_p^{(\mathrm{u})}$.

Analogously, in the case of the **normalized** graph p-Laplacian, one can define functionals $G_p : \mathbb{R}^V \to \mathbb{R}$ and $G_p^{(2)} : \mathbb{R}^V \to \mathbb{R}$,

$$G_p(f) = \frac{Q_p(f)}{\sum_{i \in V} d_i \, |f_i|^p} \qquad \text{and} \qquad G_p^{(2)}(f) = \frac{Q_p(f)}{\mathrm{var}_p^{(\mathrm{n})}(f)},$$

with the **normalized** p-variance defined as

$$\mathrm{var}_p^{(\mathrm{n})}(f) = \min_{m \in \mathbb{R}} \left\{ \sum_{i \in V} d_i \, |f_i - m|^p \right\} , \tag{3}$$

and the **normalized** p-mean of f as

$$\mathrm{mean}_p^{(\mathrm{n})}(f) = \arg\min_{m \in \mathbb{R}} \left\{ \sum_{i \in V} d_i \, |f_i - m|^p \right\} .$$

Theorems 4.1 and 4.2 in Section 4 establish the connection between these functionals and the eigenvalues of the normalized p-Laplacian.

2.1 Basic properties of p-Laplacian and related functionals

In the following sections we restrict ourselves in the proofs to the unnormalized case. Proofs in the normalized case are similar.

Proposition 2.1 *For any function $f : V \to \mathbb{R}$, and $c \in \mathbb{R}$, the **unnormalized** p-Laplacian $\Delta_p^{(\mathrm{u})}$ has the following properties:*

$$\begin{aligned} \Delta_p^{(\mathrm{u})}(f + c\mathbf{1}) &= \Delta_p^{(\mathrm{u})}(f) \\ \Delta_p^{(\mathrm{u})}(c \cdot f) &= \phi_p(c) \cdot \Delta_p^{(\mathrm{u})}(f) \ . \end{aligned}$$

*For any function $f : V \to \mathbb{R}$, and $c \in \mathbb{R}$, the **normalized** p-Laplacian $\Delta_p^{(\mathrm{n})}$ has the following properties:*

$$\begin{aligned} \Delta_p^{(\mathrm{n})}(f + c\mathbf{1}) &= \Delta_p^{(\mathrm{n})}(f) \\ \Delta_p^{(\mathrm{n})}(c \cdot f) &= \phi_p(c) \cdot \Delta_p^{(\mathrm{n})}(f) \ . \end{aligned}$$

Proof: These properties follow directly from the definition, as it holds $\forall i \in V$ that

$$\left(\Delta_p^{(\mathrm{u})}(f + c\mathbf{1}) \right)_i \ = \ \sum_{j \in V} w_{ij} \, \phi_p \left(f_i + c - f_j - c \right) \ = \ \left(\Delta_p^{(\mathrm{u})}(f) \right)_i$$

and

$$\begin{aligned} \left(\Delta_p^{(\mathrm{u})}(c \cdot f) \right)_i &= \sum_{j \in V} w_{ij} \, \phi_p \left(c \cdot (f_i - f_j) \right) \\ &= \sum_{j \in V} w_{ij} \, |c \cdot (f_i - f_j)|^{p-1} \, \mathrm{sign} \left(c \cdot (f_i - f_j) \right) \\ &= |c|^{p-1} \, \mathrm{sign}(c) \, \sum_{j \in V} w_{ij} \, |f_i - f_j|^{p-1} \, \mathrm{sign} \left(f_i - f_j \right) \\ &= \phi_p(c) \cdot \left(\Delta_p^{(\mathrm{u})}(f) \right)_i \ . \end{aligned}$$

$\square$

Proposition 2.2 *For any function $f : V \to \mathbb{R}$, and $c \in \mathbb{R}$, the functional $Q_p(f)$ has the following properties:*

$$\begin{aligned} Q_p(f + c\mathbf{1}) &= Q_p(f) \\ Q_p(c \cdot f) &= |c|^p \, Q_p(f) \ . \end{aligned}$$

Proof: Again the properties follow directly from the definition. $\square$

Proposition 2.3 *For any function $f : V \to \mathbb{R}$, and $c \in \mathbb{R}$, the* ***unnormalized*** *p-variance $\mathrm{var}_p^{(\mathrm{u})}(f)$ has the following properties:*

$$\begin{aligned} \mathrm{var}_p^{(\mathrm{u})}(f + c\mathbf{1}) &= \mathrm{var}_p^{(\mathrm{u})}(f) \\ \mathrm{var}_p^{(\mathrm{u})}(c \cdot f) &= |c|^p \, \mathrm{var}_p^{(\mathrm{u})}(f) \ . \end{aligned}$$

For any function $f : V \to \mathbb{R}$, and $c \in \mathbb{R}$, the ***normalized*** *p-variance $\mathrm{var}_p^{(\mathrm{n})}(f)$ has the following properties:*

$$\begin{aligned} \mathrm{var}_p^{(\mathrm{n})}(f + c\mathbf{1}) &= \mathrm{var}_p^{(\mathrm{n})}(f) \\ \mathrm{var}_p^{(\mathrm{n})}(c \cdot f) &= |c|^p \, \mathrm{var}_p^{(\mathrm{n})}(f) \ . \end{aligned}$$

Proof: a) Let the p-means of f and $f + c\mathbf{1}$ be given by $\widetilde{m_1} = \mathrm{mean}_p^{(\mathrm{u})}(f)$ and $\widetilde{m_2} = \mathrm{mean}_p^{(\mathrm{u})}(f + c\mathbf{1})$. Let now $m_2' := \widetilde{m_1} + c$. Then it follows from the definition of the p-variance that

$$\begin{aligned} \mathrm{var}_p^{(\mathrm{u})}(f + c\mathbf{1}) &= \min_{m \in \mathbb{R}} \left\{ \sum_{i \in V} |f_i + c - m|^p \right\} \\ &\leq \sum_{i \in V} |f_i + c - m_2'|^p \\ &= \sum_{i \in V} |f_i - \widetilde{m_1}|^p \\ &= \mathrm{var}_p^{(\mathrm{u})}(f) \ . \end{aligned}$$

Analogously, for $m_1' := \widetilde{m_2} - c$, we obtain $\mathrm{var}_p^{(\mathrm{u})}(f) \leq \mathrm{var}_p^{(\mathrm{u})}(f + c\mathbf{1})$ and hence $\mathrm{var}_p^{(\mathrm{u})}(f) = \mathrm{var}_p^{(\mathrm{u})}(f + c\mathbf{1})$.

b) If $c = 0$, one easily sees that $\mathrm{var}_p^{(\mathrm{u})}(c\,f) = 0 = |c|^p \, \mathrm{var}_p^{(\mathrm{u})}(f)$. Let now

$c \neq 0$. Then

$$\begin{aligned}
\operatorname{var}_p^{(\mathrm{u})}(c\,f) &= \min_{m \in \mathbb{R}} \left\{ \sum_{i \in V} |c\,f_i - m|^p \right\} \\
&= |c|^p \min_{m \in \mathbb{R}} \left\{ \sum_{i \in V} \left| f_i - \frac{m}{c} \right|^p \right\} \\
&= |c|^p \min_{m_2 \in \mathbb{R}} \left\{ \sum_{i \in V} |f_i - m_2|^p \ \mid \exists m \in \mathbb{R} : m_2 = \frac{m}{c} \right\} \\
&= |c|^p \operatorname{var}_p^{(\mathrm{u})}(f) \ .
\end{aligned}$$

□

The following property will later be used to establish a connection between the non-constant eigenvectors of the unnormalized and normalized p-Laplacian and the minimizers of the functionals $F_p^{(2)}$ resp. $G_p^{(2)}$.

Proposition 2.4 *Let $f \in \mathbb{R}^V$ and $\tilde{m} \in \mathbb{R}$. Then f has **unnormalized** p-mean value $\tilde{m} = \operatorname{mean}_p^{(\mathrm{u})}(f)$ if and only if the following condition holds:*

$$\sum_{i \in V} \phi_p\left(f_i - \tilde{m}\right) = 0.$$

*Let $f \in \mathbb{R}^V$ and $\tilde{m} \in \mathbb{R}$. Then f has **normalized** p-mean value $\tilde{m} = \operatorname{mean}_p^{(\mathrm{n})}(f)$ if and only if the following condition holds:*

$$\sum_{i \in V} d_i\ \phi_p\left(f_i - \tilde{m}\right) = 0.$$

Proof: We have

$$\begin{aligned}
\frac{\partial}{\partial m} \left(\sum_{i \in V} |f_i - m|^p \right) &= p \sum_{i \in V} |f_i - m|^{p-1} \operatorname{sign}\left(f_i - m\right)(-1) \\
&= -p \sum_{i \in V} \phi_p\left(f_i - m\right) \ ,
\end{aligned}$$

which implies that a necessary condition for any minimizer $\tilde{m}$ of the term $\sum_{i \in V} |f_i - m|^p$ is given as

$$\sum_{i \in V} \phi_p(f_i - \tilde{m}) = 0 \ .$$

Due to the convexity of the term $\sum_{i \in V} |f_i - m|^p$ for $p > 1$, this is also a sufficient condition. □

Proposition 2.5 *The derivative of the **unnormalized** variance $\operatorname{var}_p^{(\mathrm{u})}(f)$ with respect to f_k is given as*

$$\frac{\partial}{\partial f_k} \operatorname{var}_p^{(\mathrm{u})}(f) = p\ \phi_p\left(f_k - \operatorname{mean}_p^{(\mathrm{u})}(f)\right) \ .$$

*The derivative of the **normalized** variance* $\mathrm{var}_p^{(\mathrm{n})}(f)$ *with respect to* f_k *is given as*

$$\frac{\partial}{\partial f_k}\mathrm{var}_p^{(\mathrm{n})}(f) = p\, d_k\, \phi_p\left(f_k - \mathrm{mean}_p^{(\mathrm{n})}(f)\right) .$$

Proof: We have

$$\begin{aligned}
\frac{\partial}{\partial f_k}\mathrm{var}_p^{(\mathrm{u})}(f) &= \frac{\partial}{\partial f_k}\left(\sum_{i\in V}\left|f_i - \mathrm{mean}_p^{(\mathrm{u})}(f)\right|^p\right)\\
&= \sum_{i\in V} p\left|f_i - \mathrm{mean}_p^{(\mathrm{u})}(f)\right|^{p-1}\ \mathrm{sign}(f_i - \mathrm{mean}_p^{(\mathrm{u})}(f))\\
&\quad \cdot\left(\frac{\partial}{\partial f_k}(f_i - \mathrm{mean}_p^{(\mathrm{u})}(f))\right)
\end{aligned}$$

By applying the definition of ϕ_p and splitting the last term one obtains

$$\begin{aligned}
&\sum_{i\in V} p\,\phi_p\left(f_i - \mathrm{mean}_p^{(\mathrm{u})}(f)\right)\frac{\partial}{\partial f_k}f_i\\
-&\sum_{i\in V} p\,\phi_p\left(f_i - \mathrm{mean}_p^{(\mathrm{u})}(f)\right)\frac{\partial}{\partial f_k}\left(\mathrm{mean}_p^{(\mathrm{u})}f\right)\\
=\ & p\ \phi_p\left(f_k - \mathrm{mean}_p^{(\mathrm{u})}(f)\right)\\
-&\frac{\partial}{\partial f_k}\left(\mathrm{mean}_p^{(\mathrm{u})}(f)\right)\left(\sum_{i\in V} p\ \phi_p\left(f_i - \mathrm{mean}_p^{(\mathrm{u})}(f)\right)\right) .
\end{aligned}$$

Due to Prop. 2.4 it holds that

$$\sum_{i\in V} p\ \phi_p\left(f_i - \mathrm{mean}_p^{(\mathrm{u})}(f)\right) = 0 .$$

Thus we obtain

$$\frac{\partial}{\partial f_k}\mathrm{var}_p^{(\mathrm{u})}(f) = p\ \phi_p\left(f_k - \mathrm{mean}_p^{(\mathrm{u})}(f)\right) .$$

□

The following proposition provides the link between the functionals F_p and $F_p^{(2)}$ as well as G_p and $G_p^{(2)}$. Note that the p-mean inside the p-variance is a function $\mathbb{R}^V \to \mathbb{R}$, which we have to take into account when taking the derivative.

Proposition 2.6 *For any function* $f: V \to \mathbb{R}$ *let* $\tilde{f}$ *denote the* ***unnormalized*** *p-mean of f. Then it holds that*

$$\begin{aligned}
F_p^{(2)}(f) &= F_p(f - \tilde{f}\mathbf{1})\\
\left(\frac{\partial}{\partial f_k}F_p^{(2)}\right)(f) &= \left(\frac{\partial}{\partial f_k}F_p\right)(f - \tilde{f}\mathbf{1})\\
\left(\frac{\partial^2}{\partial f_k \partial f_l}F_p^{(2)}\right)(f) &= \left(\frac{\partial^2}{\partial f_k \partial f_l}F_p\right)(f - \tilde{f}\mathbf{1}) + F_p^{(2)}(f)\cdot\Omega(f)_{k,l} ,
\end{aligned}$$

where

$$\Omega(f)_{k,l} = \frac{p(p-1)\left|f_l - \tilde{f}\right|^{p-2}\left|f_k - \tilde{f}\right|^{p-2}}{\sum_i \left|f_i - \tilde{f}\right|^p \sum_i \left|f_i - \tilde{f}\right|^{p-2}} \ .$$

For any function $f : V \to \mathbb{R}$ let $\tilde{f}$ denote the ***normalized*** *p-mean of f. Then it holds that*

$$\begin{aligned} G_p^{(2)}(f) &= G_p(f - \tilde{f}\mathbf{1}) \\ \left(\frac{\partial}{\partial f_k} G_p^{(2)}\right)(f) &= \left(\frac{\partial}{\partial f_k} G_p\right)(f - \tilde{f}\mathbf{1}) \\ \left(\frac{\partial^2}{\partial f_k \partial f_l} G_p^{(2)}\right)(f) &= \left(\frac{\partial^2}{\partial f_k \partial f_l} F_p\right)(f - \tilde{f}\mathbf{1}) + F_p^{(2)}(f) \cdot \Omega(f)_{k,l} \end{aligned}$$

where

$$\Omega(f)_{k,l} = \frac{p(p-1) d_l d_k \left|f_l - \tilde{f}\right|^{p-2}\left|f_k - \tilde{f}\right|^{p-2}}{\sum_i d_i \left|f_i - \tilde{f}\right|^p \sum_i d_i \left|f_i - \tilde{f}\right|^{p-2}} \ .$$

Proof: The first statement can be seen directly by the definitions of F_p and $F_p^{(2)}$ and the fact that

$$Q_p(f) = Q_p(f - \tilde{f}\mathbf{1}) \ .$$

Using Prop. 2.5 and the definition of $\Delta_p^{(\mathrm{u})}$, the derivative of $\frac{Q_p(f)}{\mathrm{var}_p(f)}$ with respect to f_k can be written as

$$\frac{\partial}{\partial f_k}\left(\frac{Q_p(f)}{\mathrm{var}_p(f)}\right) = \frac{p}{\mathrm{var}_p(f)}(\Delta_p^{(\mathrm{u})} f)_k - \frac{Q_p(f)\, p}{\mathrm{var}_p^2(f)} \phi_p\left(f_k - \tilde{f}\right) \ .$$

By applying Prop. 2.1 and Prop. 2.2 as well as the definition of the p-variance, the above expression can be rewritten as

$$\frac{p}{\left\|f - \tilde{f}\mathbf{1}\right\|_p^p}\left(\Delta_p^{(\mathrm{u})}(f - \tilde{f}\mathbf{1})\right)_k - \frac{Q_p(f - \tilde{f}\mathbf{1})\, p}{\left\|f - \tilde{f}\mathbf{1}\right\|_p^{2p}} \phi_p\left(f_k - \tilde{f}\right) \ .$$

Comparison with the expression for $\left(\frac{\partial}{\partial f_k} F_p\right)$ now yields the second statement. For the statement for the second derivatives, one first shows that

$$\frac{\partial}{\partial f_l}\tilde{f} = \frac{|f_l - \tilde{f}|^{p-2}}{\sum_i |f_i - \tilde{f}|^{p-2}} \ \forall l = 1 \ldots n$$

and then proceeds analogously to the first and second statement. $\square$

3 Variational characterization of the second eigenvalue - Unnormalized case

Theorem 3.1 *The second eigenvalue of the* ***unnormalized*** *graph p-Laplacian $\Delta_p^{(\mathrm{u})}$ is equal to the global minimum of the functional $F_p^{(2)}$. The corresponding*

eigenvector $v_p^{(2)}$ *of* $\Delta_p^{(\mathrm{u})}$ *is then given as* $v_p^{(2)} = u^* - c^*\mathbf{1}$ *for any global minimizer* u^* *of* $F_p^{(2)}$*, where* $c^* = \arg\min_{c\in\mathbb{R}} \sum_{i=1}^n |u_i^* - c|^p$.
Furthermore, the functional $F_p^{(2)}$ *satisfies* $F_p^{(2)}(tu + c\mathbf{1}) = F_p^{(2)}(u) \quad \forall t,\, c \in \mathbb{R}$.

Lemma 3.1 *Let* f *be a critical point of the functional* $F_p^{(2)}$*. Then the vector*

$$v = f - \mathrm{mean}_p^{(\mathrm{u})}(f)\mathbf{1}$$

is an eigenfunction of $\Delta_p^{(\mathrm{u})}$ *with eigenvalue* $\lambda_p = F_p^{(2)}(f)$.

Proof: Let f be a critical point of $F_p^{(2)}$ with minimum λ_p. Then

$$\left(\frac{\partial}{\partial f_k} F_p^{(2)}\right)(f) = 0 \ .$$

By Prop. 2.6 this implies

$$\left(\frac{\partial}{\partial f_k} F_p\right)\left(f - \mathrm{mean}_p^{(\mathrm{u})}(f)\mathbf{1}\right) = 0$$

as well as

$$\lambda_p = F_p^{(2)}(f) = F_p\left(f - \mathrm{mean}_p^{(\mathrm{u})}(f)\mathbf{1}\right) \ .$$

It follows that $f - \mathrm{mean}_p^{(\mathrm{u})}(f)\mathbf{1}$ is a critical point of F_p, and by Theorem 3.1. in [2] an eigenvector of $\Delta_p^{(\mathrm{u})}$ with eigenvalue λ_p. $\square$

Before proving the other direction, let us first derive an important property of the non-constant eigenvectors of the p-Laplacian.

Lemma 3.2 *Let* v *be a non-constant eigenvector of* Δ_p*. Then*

$$\sum_{i\in V} \phi_p(v_i) = 0 \ .$$

Proof: Let v be a non-constant eigenvector of $\Delta_p^{(\mathrm{u})}$ with eigenvalue λ_p. Hence for all $i \in V$ the equation

$$(\Delta_p v)_i - \lambda_p \phi_p(v_i) = 0$$

holds. As v is not the constant vector, we know that $\lambda_p \neq 0$ and hence $\forall i \in V$,

$$\phi_p(v_i) = \frac{(\Delta_p v)_i}{\lambda_p} \ .$$

It follows that

$$\begin{aligned}
\sum_{i\in V} \phi_p(v_i) &= \frac{1}{\lambda_p}\sum_{i\in V} (\Delta_p v)_i \\
&= \frac{1}{\lambda_p}\sum_{i\in V}\sum_{j\in V} w_{ij}\, \phi_p(v_i - v_j) \\
&= \frac{1}{\lambda_p}\sum_{i,j\in V, v_i > v_j} w_{ij}\, |v_i - v_j|^{p-1} - \frac{1}{\lambda_p}\sum_{i,j\in V, v_i < v_j} w_{ij}\, |v_i - v_j|^{p-1} \\
&= \frac{1}{\lambda_p}\sum_{i,j\in V, v_i > v_j} w_{ij}\, |v_i - v_j|^{p-1} - \frac{1}{\lambda_p}\sum_{j,i\in V, v_j < v_i} w_{ji}\, |v_j - v_i|^{p-1} \\
&= 0 \ ,
\end{aligned}$$

where in the penultimate step we have performed a change of the variable names in the second term and in the last step exploited the fact that $w_{ij} = w_{ji}$. $\square$

The above property can be seen as a generalization of the fact that the larger eigenvectors of the unnormalized (standard) graph Laplacian are orthogonal to the first eigenvector.

Lemma 3.3 *Let v be a non-constant eigenvector of the p-Laplacian $\Delta_p^{(\mathrm{u})}$ with eigenvalue λ_p. Then there exists a function f which is a critical point of $F_p^{(2)}$ with $\lambda_p = F_p^{(2)}(f)$ and it holds that $v = f - \mathrm{mean}_p^{(\mathrm{u})}(f)\mathbf{1}$.*

Proof: By Theorem 3.1. in [2] we know that v is a critical point of F_p with $\lambda_p = F_p(v)$. Consider now for $k \in \mathbb{R}$ the function $f : V \to \mathbb{R}$ defined by

$$f = v + k\mathbf{1} \ .$$

By Lemma 3.2 it holds that

$$\sum_{i \in V} \phi_p(v_i) = 0 \ .$$

It follows that $\forall k$:

$$\sum_{i \in V} \phi_p\left(f_i - k\right) = \sum_{i \in V} \phi_p\Big((f - k\mathbf{1})_i \Big) = \sum_{i \in V} \phi_p(v_i) = 0 \ .$$

By Prop. 2.4 this implies that $k = \mathrm{mean}_p^{(\mathrm{u})}(f)$, and hence

$$v = f - \mathrm{mean}_p^{(\mathrm{u})}(f)\mathbf{1} \ .$$

Prop. 2.6 now implies that

$$F_p^{(2)}\left(f\right) = F_p\left(f - \mathrm{mean}_p^{(\mathrm{u})}(f)\mathbf{1}\right) = F_p(v) = \lambda_p$$

and

$$\begin{aligned}\left(\frac{\partial}{\partial f_k} F_p^{(2)}\right)(f) &= \left(\frac{\partial}{\partial f_k} F_p\right)\left(f - \mathrm{mean}_p^{(\mathrm{u})}(f)\mathbf{1}\right) \\ &= \left(\frac{\partial}{\partial f_k} F_p\right)(v) = 0 \ .\end{aligned}$$

Hence it follows that f is a minimizer of $F_p^{(2)}$ with minimum λ_p. $\square$

Proof of Theorem 3.1: Lemma 3.1 shows the forward direction of the first statement of Theorem 3.1. The reverse direction follows from Lemma 3.3. The second statement follows from Prop. 2.2 and 2.3. $\square$

4 Variational characterization of the second eigenvalue - Normalized case

The following theorems are the normalized variants of Theorem 3.1 and Theorem 3.2 in [2].

Theorem 4.1 *The functional G_p has a critical point at $v \in \mathbb{R}^V$ if and only if v is a p-eigenfunction of the* ***normalized*** *graph p-Laplacian $\Delta_p^{(\mathrm{n})}$. The corresponding eigenvalue λ_p is given as $\lambda_p = G_p(v)$. Moreover, we have $G_p(\alpha f) = G_p(f)$ for all $f \in \mathbb{R}^V$ and $\alpha \in \mathbb{R}$.*

Theorem 4.2 *The second eigenvalue of the* ***normalized*** *graph p-Laplacian $\Delta_p^{(\mathrm{n})}$ is equal to the global minimum of the functional $G_p^{(2)}$. The corresponding eigenvector $v_p^{(2)}$ of $\Delta_p^{(\mathrm{n})}$ is then given as $v_p^{(2)} = u^* - c^* \mathbf{1}$ for any global minimizer u^* of $G_p^{(2)}$, where $c^* = \underset{c \in \mathbb{R}}{\arg\min} \sum_{i=1}^n d_i \, |u_i^* - c|^p$.*
Furthermore, the functional $G_p^{(2)}$ satisfies $G_p^{(2)}(tu + c\mathbf{1}) = G_p^{(2)}(u) \quad \forall t,\, c \in \mathbb{R}$.

The proofs of the above theorems are similar to the unnormalized case. We just want to sketch the proof of Theorem 4.2 by giving the corresponding lemmas without proof.

Lemma 4.1 *Let f be a critical point of the functional $G_p^{(2)}$. Then the vector*

$$v = f - \mathrm{mean}_p^{(\mathrm{n})}(f)\mathbf{1}$$

is an eigenfunction of $\Delta_p^{(\mathrm{n})}$ with eigenvalue $\lambda_p = G_p^{(2)}(f)$.

Lemma 4.2 *Let v be a non-constant eigenvector of $\Delta_p^{(\mathrm{n})}$. Then*

$$\sum_{i \in V} d_i \, \phi_p(v_i) = 0 \ .$$

Lemma 4.3 *Let v be a non-constant eigenvector of the p-Laplacian $\Delta_p^{(\mathrm{n})}$ with eigenvalue λ_p. Then there exists a function f which is a critical point of $G_p^{(2)}$ with $\lambda_p = G_p^{(2)}(f)$ and it holds that $v = f - \mathrm{mean}_p^{(\mathrm{n})}(f)\mathbf{1}$.*

5 Isoperimetric inequality - Unnormalized case

As shown in [2], for $p > 1$ and every partition of V into $C, \overline{C}$ there exists a function $f_{p,C} \in \mathbb{R}^V$ such that the functional $F_p^{(2)}$ associated to the **unnormalized** p-Laplacian satisfies

$$F_p^{(2)}(f_{p,C}) = \mathrm{cut}(C, \overline{C}) \left| \frac{1}{|C|^{\frac{1}{p-1}}} + \frac{1}{\left|\overline{C}\right|^{\frac{1}{p-1}}} \right|^{p-1} . \tag{4}$$

Explicitely, the function $f_{p,C}$ is given as

$$(f_{p,C})_i = \begin{cases} 1/\,|C|^{\frac{1}{p-1}} & ,\ i \in C, \\ -1/\,\left|\overline{C}\right|^{\frac{1}{p-1}} & ,\ i \in \overline{C}. \end{cases} \tag{5}$$

The expression (4) can be interpreted as a balanced graph cut criterion, and we have the special cases

$$\begin{aligned} F_2^{(2)}(f_{2,C}) &= \mathrm{RCut}(C, \overline{C}), \\ \lim_{p \to 1} F_p^{(2)}(f_{p,C}) &= \mathrm{RCC}(C, \overline{C}). \end{aligned}$$

It follows that minimizing the above balanced graph cut criterion is equivalent to minimizing $F_p^{(2)}$ with the restriction to functions that have the form given in (5). As the second eigenvalue of the p-Laplacian is the minimum of the functional $F_p^{(2)}$ taken over all possible functions (without the restriction), the second eigenvalue can be seen as a relaxation of balanced graph cuts. The question is, can we make any statements about the quality of this relaxation?

The isoperimetric inequality gives upper and lower bounds on the second eigenvalue in terms of the optimal Cheeger cut value defined as

$$h_{\mathrm{RCC}} = \inf_C \mathrm{RCC}(C, \overline{C}).$$

Theorem 5.1 *Denote by $\lambda_p^{(2)}$ the second eigenvalue of the* ***unnormalized*** *graph p-Laplacian $\Delta_p^{(\mathrm{u})}$. For $p > 1$,*

$$\left(\frac{2}{\max_{i \in V} d_i}\right)^{p-1} \left(\frac{h_{\mathrm{RCC}}}{p}\right)^p \;\leq\; \lambda_p^{(2)} \;\leq\; 2^{p-1} h_{\mathrm{RCC}} \;.$$

Proof of the upper bound in Theorem 5.1: Let for any $p > 1$ the second smallest eigenvalue of the unnormalized p-Laplacian be given by λ_p. Theorem 3.1 implies that

$$\lambda_p = \min_{f \in \mathbb{R}^V} \left\{ F_p^{(2)}(f) \right\} = \min_{f \in \mathbb{R}^V} \left\{ \frac{Q_p(f)}{\mathrm{var}_p^{(\mathrm{u})}(f)} \right\} \;,$$

where $Q_p(f)$ and $\mathrm{var}_p^{(\mathrm{u})}(f)$ are defined as in (1) and (2). Consider now for a partition $(C, \overline{C})$ the function $f_{p,C} : V \to \mathbb{R}$ which we have defined in (5). Then, using (4), we have

$$\begin{aligned} \lambda_p &\leq F_p^{(2)}(f_{p,C}) \\ &= \mathrm{cut}(C, \overline{C}) \left| \frac{1}{|C|^{\frac{1}{p-1}}} + \frac{1}{|\overline{C}|^{\frac{1}{p-1}}} \right|^{p-1} \\ &\leq \mathrm{cut}(C, \overline{C}) \left| 2 \frac{1}{\min\left\{ |C|^{\frac{1}{p-1}}, |\overline{C}|^{\frac{1}{p-1}} \right\}} \right|^{p-1} \\ &= \frac{\mathrm{cut}(C, \overline{C})}{\min\left\{ |C|, |\overline{C}| \right\}} \cdot 2^{p-1} \\ &= \mathrm{RCC}(C, \overline{C}) \cdot 2^{p-1} \;. \end{aligned}$$

As this inequality holds for all partitions $(C, \overline{C})$, it follows that

$$\lambda_p \leq \inf_C \mathrm{RCC}(C, \overline{C}) = h_{\mathrm{RCC}} \;.$$

$\square$

For the proof of the lower bound we need to introduce some notation. In the following let us for any function $f : V \to \mathbb{R}$ denote by $f^+ : V \to \mathbb{R}$ the function

$$f_i^+ = \begin{cases} f_i & , \quad f_i \geq 0 \ , \\ 0 & , \quad \text{else} \ . \end{cases} \tag{6}$$

Furthermore, we use the notation $C_f^t, \overline{C_f^t}$ for the partitioning of the vertex set V into the sets

$$C_f^t = \{i | f_i > t\} \qquad \text{and} \qquad \overline{C_f^t} = V - C_f^t = \{i | f_i \leq t\} \ , \tag{7}$$

where $t \in \mathbb{R}$. Finally, for any function $f : V \to \mathbb{R}$, we denote by $h^*_{f,\mathrm{RCC}}$ the quantity

$$h^*_{f,\mathrm{RCC}} = \inf_C \left\{ \frac{\mathrm{cut}(C, \overline{C})}{\min\left\{|C|, |\overline{C}|\right\}} \ \Big| \ C = C_f^t \text{ for } t \geq 0 \right\} \ . \tag{8}$$

The value of $h^*_{f,\mathrm{RCC}}$ is the smallest possible RCC value obtained by thresholding f at some $t \geq 0$. If $C_f^0 = \emptyset$, we define $h^*_{f,\mathrm{RCC}} = \infty$.

To prove the lower bound, we proceed in analogy to [1].

Lemma 5.1 *Suppose there exists a $\lambda \geq 0$ such that it holds $\forall i \in C_f^0$ that $(\Delta_p^{(\mathrm{u})} f)_i \leq \lambda f_i^{p-1}$. Then*

$$\lambda \geq \frac{Q_p(f^+)}{\left\|f^+\right\|_p^p} \ .$$

Proof: We have

$$\lambda \left\|f^+\right\|_p^p = \lambda \sum_{i \in V} \left|f_i^+\right|^p = \lambda \sum_{i \in C_f^0} |f_i|^p = \lambda \sum_{i \in C_f^0} f_i f_i^{p-1} \ .$$

Using the assumption, it follows that

$$\lambda \left\|f^+\right\|_p^p \geq \sum_{i \in C_f^0} f_i (\Delta_p^{(\mathrm{u})} f)_i = \sum_{i \in V} f_i^+ (\Delta_p^{(\mathrm{u})} f)_i \ .$$

Applying the definition of $\Delta_p^{(\mathrm{u})}$, this can be rewritten as

$$\begin{aligned} &= \sum_{i \in V} f_i^+ \sum_{j \in V} w_{ij} \phi_p(f_i - f_j) \\ &= \frac{1}{2} \sum_{i,j \in V} w_{ij} f_i^+ \phi_p(f_i - f_j) + \frac{1}{2} \sum_{i,j \in V} w_{ij} f_i^+ \phi_p(f_i - f_j) \\ &= \frac{1}{2} \sum_{i,j \in V} w_{ij} (f_i^+ - f_j^+) \phi_p(f_i - f_j) \ . \end{aligned}$$

Let us now have a closer look at the summands in the above sum. If both i and j are in C_f^0, we have

$$\begin{aligned} (f_i^+ - f_j^+) \phi_p(f_i - f_j) &= (f_i^+ - f_j^+) \phi_p(f_i^+ - f_j^+) \\ &= \left|f_i^+ - f_j^+\right|^p \ . \end{aligned}$$

If $i \in C_f^0$ and $j \notin C_f^0$, it holds that

$$\begin{aligned}(f_i^+ - f_j^+)\phi_p(f_i - f_j) &= f_i^+ \left|f_i - f_j\right|^{p-1} \operatorname{sign}(f_i - f_j) \\ &= f_i^+ \left|f_i - f_j\right|^{p-1} \\ &\geq f_i^+ \left|f_i\right|^{p-1} \ ,\end{aligned}$$

where we have used that $f_i > 0$ and $f_j \leq 0$. The last term can be rewritten as

$$\left|f_i^+\right|^p = \left|f_i^+ - f_j^+\right|^p \ .$$

Analogously, if $i \notin C_f^0$ and $j \in C_f^0$, we have

$$\begin{aligned}(f_i^+ - f_j^+)\phi_p(f_i - f_j) &= -f_j^+ \left|f_i - f_j\right|^{p-1} \operatorname{sign}(f_i - f_j) \\ &= f_j^+ \left|f_i - f_j\right|^{p-1} \\ &\geq f_j^+ \left|f_j\right|^{p-1} \\ &= \left|f_j^+\right|^p = \left|f_i^+ - f_j^+\right|^p \ .\end{aligned}$$

Finally, in the case that both i and j are not in C_f^0, it holds that

$$(f_i^+ - f_j^+)\phi_p(f_i - f_j) = 0 = \left|f_i^+ - f_j^+\right|^p \ .$$

If we combine these results, we obtain that

$$\begin{aligned}&\frac{1}{2}\sum_{i,j \in V} w_{ij}(f_i^+ - f_j^+)\phi_p(f_i - f_j) \\ \geq\ &\frac{1}{2}\sum_{i,j \in V} w_{ij}\left|f_i^+ - f_j^+\right|^p = Q_p(f^+) \ ,\end{aligned}$$

which completes our proof. □

The following inequality, which will be used in the next lemma, has been shown by Amghibech [1].

Lemma 5.2 (Amghibech, [1]) *If $a, b \geq 0$, $p > 1$ and $\frac{1}{p} + \frac{1}{q} = 1$, then*

$$\left(\frac{1}{p}\frac{b^p - a^p}{b - a}\right)^q \leq \frac{1}{2}(a^p + b^p) \ .$$

Lemma 5.3 *For any function $f \in \mathbb{R}^V$ with $0 < \left|C_f^0\right| \leq \frac{1}{2}\left|V\right|$ and $h^*_{f,\mathrm{RCC}}$ as defined in* (8) *it holds that*

$$\frac{Q_p(f^+)}{\left\|f^+\right\|_p^p} \geq \left(\frac{h^*_{f,\mathrm{RCC}}}{p}\right)^p \left(\frac{2}{\max_{i \in V} d_i}\right)^{p-1} \ .$$

Proof: Consider the term

$$\sum_{f_j^+ > f_i^+} w_{ij}\Big((f_j^+)^p - (f_i^+)^p\Big) \ .$$

On the one hand we have

$$\begin{aligned}\sum_{f_j^+>f_i^+} w_{ij}\Big((f_j^+)^p-(f_i^+)^p\Big) &= \sum_{f_j^+>f_i^+} w_{ij}\,[t^p]_{f_i^+}^{f_j^+}\\ &= p\cdot\sum_{f_j^+>f_i^+} w_{ij}\int_{f_i^+}^{f_j^+} t^{p-1}dt\ .\end{aligned}$$

We can change the order of integration and summation and obtain

$$p\cdot\sum_{f_j^+>f_i^+} w_{ij}\int_{f_i^+}^{f_j^+} t^{p-1}dt = p\cdot\int_0^\infty t^{p-1}\sum_{f_j^+>t\geq f_i^+} w_{ij}\ dt\ .$$

Note that for $t\geq 0$,

$$\sum_{f_j^+>t\geq f_i^+} w_{ij} = \sum_{f_j>t\geq f_i} w_{ij} = \sum_{j\in C_f^t,\, i\in\overline{C_f^t}} w_{ij} = \mathrm{cut}(C_f^t,\overline{C_f^t})\ ,$$

which leads us to the following inequality

$$\begin{aligned}\mathrm{cut}(C_f^t,\overline{C_f^t}) &= \frac{\mathrm{cut}(C_f^t,\overline{C_f^t})}{\left|C_f^t\right|}\cdot\left|C_f^t\right| = \frac{\mathrm{cut}(C_f^t,\overline{C_f^t})}{\min\left\{\left|C_f^t\right|,\left|\overline{C_f^t}\right|\right\}}\cdot\left|C_f^t\right|\\ &\geq \inf_C\left\{\frac{\mathrm{cut}(C,\overline{C})}{\min\left\{|C|,\left|\overline{C}\right|\right\}}\ \Big|\ C=C_f^t \text{ for } t\geq 0\right\}\cdot\left|C_f^t\right|\\ &= h^*_{f,\mathrm{RCC}}\cdot\left|C_f^t\right|\ ,\end{aligned}$$

where in the second step we used the assumption that $0<\left|C_f^0\right|\leq\frac{1}{2}|V|$. As this inequality holds for all $t\geq 0$, we obtain

$$p\int_0^\infty t^{p-1}\sum_{f_j^+>t\geq f_i^+} w_{ij}\ dt \geq p\int_0^\infty t^{p-1}h^*_{f,\mathrm{RCC}}\left|C_f^t\right|dt\ .$$

We now use that $\left|C_f^t\right| = \sum_{f_i>t}\mathbf{1} = \sum_{f_i^+>t}\mathbf{1}$ (for $t\geq 0$), and change the order of summation and integration again, which leads us to

$$\begin{aligned}p\int_0^\infty t^{p-1}h^*_{f,\mathrm{RCC}}\sum_{f_i^+>t}\mathbf{1}dt &= h^*_{f,\mathrm{RCC}}\,p\sum_{f_j^+>0}\int_0^{f_j^+} t^{p-1}dt\\ &= h^*_{f,\mathrm{RCC}}\sum_{f_j^+>0}[t^p]_0^{f_j^+}\\ &= h^*_{f,\mathrm{RCC}}\sum_{f_j^+>0}\left(f_j^+\right)^p\\ &= h^*_{f,\mathrm{RCC}}\left\|f^+\right\|_p^p\ .\end{aligned}$$

So we have just shown the inequality

$$\sum_{f_j^+ > f_i^+} w_{ij} \Big((f_j^+)^p - (f_i^+)^p \Big) \geq h^*_{f,\mathrm{RCC}} \ \left\| f^+ \right\|_p^p \ . \tag{9}$$

On the other hand we have

$$\begin{aligned}
& \sum_{f_j^+ > f_i^+} w_{ij} \Big((f_j^+)^p - (f_i^+)^p \Big) \\
= \ & \frac{1}{2} \sum_{f_j^+ > f_i^+} w_{ij} \Big((f_j^+)^p - (f_i^+)^p \Big) + \frac{1}{2} \sum_{f_i^+ > f_j^+} w_{ij} \Big((f_i^+)^p - (f_j^+)^p \Big) \\
= \ & \frac{1}{2} \sum_{i,j \in V} w_{ij} \left| (f_j^+)^p - (f_i^+)^p \right| \ ,
\end{aligned}$$

where again we exploited the symmetry of the weights in the second step. Let q be the conjugate of p, defined by the equation $\frac{1}{p} + \frac{1}{q} = 1$. The sum can now be decomposed into

$$\begin{aligned}
& \frac{1}{2} \sum_{i,j \in V} w_{ij} \left| (f_j^+)^p - (f_i^+)^p \right| \\
= \ & \sum_{f_j^+ \neq f_i^+} \left(\frac{1}{2} w_{ij} \right)^{1/p} \left| f_j^+ - f_i^+ \right| \cdot \left(\frac{1}{2} w_{ij} \right)^{1/q} \frac{(f_j^+)^p - (f_i^+)^p}{f_j^+ - f_i^+} \\
\leq \ & \left(\sum_{f_j^+ \neq f_i^+} \frac{1}{2} w_{ij} \left| f_j^+ - f_i^+ \right|^p \right)^{1/p} \cdot \left(\sum_{f_j^+ \neq f_i^+} \frac{1}{2} w_{ij} \left| \frac{(f_j^+)^p - (f_i^+)^p}{f_j^+ - f_i^+} \right|^q \right)^{1/q} \\
= \ & Q_p(f^+)^{1/p} \cdot \left(\sum_{f_j^+ \neq f_i^+} \frac{1}{2} w_{ij} \left(\frac{(f_j^+)^p - (f_i^+)^p}{f_j^+ - f_i^+} \right)^q \right)^{1/q} \ ,
\end{aligned}$$

where we used Hölder's inequality in the second step. By applying Lemma 5.2

we obtain for the second product term (assuming that $p \geq 1$)

$$\begin{aligned}
\left(\frac{1}{2} \sum_{f_j^+ \neq f_i^+} w_{ij} \left(\frac{(f_i^+)^p - (f_j^+)^p}{f_i^+ - f_j^+} \right)^q \right)^{1/q} &\leq \left(\frac{1}{2} \sum_{f_j^+ \neq f_i^+} w_{ij} \frac{p^q}{2} \Big((f_i^+)^p + (f_j^+)^p \Big) \right)^{1/q} \\
&\leq \frac{p}{4^{1/q}} \left(\sum_{i,j \in V} w_{ij} \Big((f_i^+)^p + (f_j^+)^p \Big) \right)^{1/q} \\
&= \frac{p}{4^{1/q}} \left(2 \sum_{i \in V} d_i \, (f_i^+)^p \right)^{1/q} \\
&\leq \frac{p}{2^{1/q}} \left(\sum_{i \in V} \max_{i \in V} d_i (f_i^+)^p \right)^{1/q} \\
&= p \left(\frac{\max_{i \in V} d_i}{2} \right)^{1-1/p} \left\| f^+ \right\|_p^{p-1} \ .
\end{aligned}$$

Hence we have

$$\sum_{f_j^+ > f_i^+} \Big((f_j^+)^p - (f_i^+)^p \Big) \cdot w_{ij} \leq p \left(\frac{\max_i d_i}{2} \right)^{1-1/p} Q_p(f^+)^{1/p} \left\| f^+ \right\|_p^{p-1} \ . \tag{10}$$

By combining (9) and (10) we obtain

$$h^*_{f,\mathrm{RCC}} \left\| f^+ \right\|_p^p \leq p \left(\frac{\max_{i \in V} d_i}{2} \right)^{1-1/p} Q_p(f^+)^{1/p} \left\| f^+ \right\|_p^{p-1} \ ,$$

which can be rewritten as

$$\left(\frac{2}{\max_{i \in V} d_i} \right)^{p-1} \left(\frac{h^*_{f,\mathrm{RCC}}}{p} \right)^p \leq \frac{Q_p(f^+)}{\left\| f^+ \right\|_p^p} \ .$$

□

Proof of the lower bound in Theorem 5.1: Let f be the eigenfunction of $\Delta_p^{(\mathrm{u})}$ corresponding to the second eigenvalue λ_p. Let C_f^0 be the set of values where $f_i > 0$, as defined in (7). Without loss of generality, we can assume that

$$0 < \left| C_f^0 \right| \leq \frac{1}{2} \left| V \right| \ ,$$

otherwise we just replace f by $-f$. We know that

$$\Delta_p^{(\mathrm{u})} f = \lambda_p f^{p-1} \text{ on } C_f^0 \ ,$$

so our condition for Lemma 5.1 is fulfilled. Applying Lemma 5.1 and 5.3 yields

$$\lambda_p \geq \frac{Q_p(f^+)}{\left\| f^+ \right\|_p^p} \geq \left(\frac{2}{\max_{i \in V} d_i} \right)^{p-1} \left(\frac{h^*_{f,\mathrm{RCC}}}{p} \right)^p \ .$$

Clearly, we have $h^*_{f,\mathrm{RCC}} \geq h_{\mathrm{RCC}}$, which completes the proof. □

6 Isoperimetric inequality - Normalized case

One can show that also in the normalized case, for $p > 1$ and every partition of V into $C, \overline{C}$ there exists a function $g_{p,C} \in \mathbb{R}^V$ such that the functional $G_p^{(2)}$ associated to the **normalized** p-Laplacian satisfies

$$G_p^{(2)}(g_{p,C}) = \mathrm{cut}(C, \overline{C}) \left| \frac{1}{\mathrm{vol}(C)^{\frac{1}{p-1}}} + \frac{1}{\mathrm{vol}(\overline{C})^{\frac{1}{p-1}}} \right|^{p-1} . \tag{11}$$

Explicitely, the function $g_{p,C}$ is given as

$$(g_{p,C})_i = \begin{cases} 1/\operatorname{vol}(C)^{\frac{1}{p-1}} & , \ i \in C, \\ -1/\operatorname{vol}(\overline{C})^{\frac{1}{p-1}} & , \ i \in \overline{C}. \end{cases} \tag{12}$$

As in the unnormalized case, the expression (11) can be interpreted as a balanced graph cut criterion, and we have the special cases

$$\begin{aligned} G_2^{(2)}(g_{2,C}) &= \mathrm{NCut}(C, \overline{C}), \\ \lim_{p \to 1} G_p^{(2)}(g_{p,C}) &= \mathrm{NCC}(C, \overline{C}). \end{aligned}$$

It follows that minimizing the above balanced graph cut criterion is equivalent to minimizing $G_p^{(2)}$ with the restriction to functions that have the form given in (12). With the same argument as in the unnormalized case, the second eigenvector can be seen as a relaxation of balanced graph cuts. As in the unnormalized case, the isoperimetric inequality gives upper and lower bounds on the second eigenvalue in terms of the optimal Cheeger cut value defined as

$$h_{\mathrm{NCC}} = \inf_C \mathrm{NCC}(C, \overline{C}).$$

Theorem 6.1 (Amghibech, [1]) *Denote by $\lambda_p^{(2)}$ the second eigenvalue of the* ***normalized*** *graph p-Laplacian $\Delta_p^{(\mathrm{n})}$. For $p > 1$,*

$$2^{p-1} \left(\frac{h_{\mathrm{NCC}}}{p} \right)^p \ \leq \ \lambda_p^{(2)} \ \leq \ 2^{p-1} \, h_{\mathrm{NCC}} \ .$$

The proof of the upper bound is similar to the unnormalized case. For the proof of the lower bound, we use again the notation f^+ for the restriction of the function f to positive values, as well as $C_f^t, \overline{C_f^t}$ for a partitioning of the vertex set by thresholding, as introduced in (6) and (7). Furthermore, for any function $f : V \to \mathbb{R}$, we denote by $h^*_{f,\mathrm{NCC}}$ the quantity

$$h^*_{f,\mathrm{NCC}} = \inf_C \left\{ \frac{\mathrm{cut}(C, \overline{C})}{\min\left\{\mathrm{vol}(C), \mathrm{vol}(\overline{C})\right\}} \ \Big| \ C = C_f^t \text{ for } t \geq 0 \right\} \ .$$

Analogously to the unnormalized case, the value of $h^*_{f,\mathrm{NCC}}$ is the smallest possible NCC value obtained by thresholding f at some $t \geq 0$. Again, we set $h^*_{f,\mathrm{NCC}} = \infty$ in the case $C_f^0 = \emptyset$.

Lemma 6.1 *Suppose there exists a $\lambda \geq 0$ such that it holds $\forall i \in C_f^0$ that $(\Delta_p^{(\mathrm{n})} f)_i \leq \lambda f_i^{p-1}$. Then*

$$\lambda \geq \frac{Q_p(f^+)}{\sum_{i \in V} d_i \left| f_i^+ \right|^p} \ .$$

Lemma 6.2 *For any function $f \in \mathbb{R}^V$ with $0 < \mathrm{vol}(C_f^0) \leq \frac{1}{2}\mathrm{vol}(V)$ and $h^*_{f,\mathrm{NCC}}$ as defined above it holds that*

$$\frac{Q_p(f^+)}{\sum_{i \in V} d_i \left|f_i^+\right|^p} \geq \left(\frac{h^*_{f,\mathrm{NCC}}}{p}\right)^p 2^{p-1} \ .$$

Using the above lemmas the lower bound can now be proven in a similar way to the unnormalized case.

7 Convergence to the optimal Cheeger cut

In p-spectral clustering, a partitioning of the graph is obtained by thresholding the real-valued second eigenvector $v_p^{(2)}$ of the graph p-Laplacian. The optimal threshold is determined by minimizing the corresponding Cheeger cut, i.e. in the case of the **unnormalized** graph p-Laplacian $\Delta_p^{(\mathrm{u})}$ one determines

$$\underset{C_t = \{i \in V \,|\, v_p^{(2)}(i) > t\}}{\arg\min} \quad \mathrm{RCC}(C_t, \overline{C_t}), \tag{13}$$

and similarly for the second eigenvector of the **normalized** graph p-Laplacian $\Delta_p^{(\mathrm{n})}$ one computes

$$\underset{C_t = \{i \in V \,|\, v_p^{(2)}(i) > t\}}{\arg\min} \quad \mathrm{NCC}(C_t, \overline{C_t}). \tag{14}$$

One can now establish a connection between the cut obtained by thresholding according to the above scheme and the optimal Cheeger cut.

Theorem 7.1 *Denote by h^*_{RCC} the ratio Cheeger cut value obtained by tresholding the second eigenvector $v_p^{(2)}$ of the* ***unnormalized*** *p-Laplacian via* (13)*. Then for $p > 1$,*

$$h_{\mathrm{RCC}} \ \leq \ h^*_{\mathrm{RCC}} \ \leq \ p\left(\max_{i \in V} d_i\right)^{\frac{p-1}{p}} \left(h_{\mathrm{RCC}}\right)^{\frac{1}{p}} \ .$$

Denote by h^*_{NCC} the normalized Cheeger cut value obtained by tresholding the second eigenvector $v_p^{(2)}$ of the ***normalized*** *p-Laplacian via* (14)*. Then for $p > 1$,*

$$h_{\mathrm{NCC}} \ \leq \ h^*_{\mathrm{NCC}} \ \leq \ p\left(h_{\mathrm{NCC}}\right)^{\frac{1}{p}} \ .$$

Interestingly, the inequalities become tight for $p \to 1$. This implies that the cut found by thresholding converges to the optimal Cheeger cut, which provides the main motivation for p-spectral clustering.

Proof of Theorem 7.1: Clearly, the lower bound holds. Let now f be the eigenfunction of $\Delta_p^{(\mathrm{u})}$ corresponding to the second eigenvalue λ_p. Let C_f^0 be the set of values where $f_i > 0$, as defined in (7). Without loss of generality, we can assume that

$$0 < \left|C_f^0\right| \leq \frac{1}{2}\left|V\right| \ ,$$

otherwise we just replace f by $-f$. We know that

$$\Delta_p^{(\mathrm{u})} f = \lambda_p f^{p-1} \text{ on } C_f^0 \ ,$$

so our condition for Lemma 5.1 is fulfilled. Applying Lemma 5.1 and 5.3 yields

$$\lambda_p \geq \frac{Q_p(f^+)}{\|f^+\|_p^p} \geq \left(\frac{2}{\max_{i\in V} d_i}\right)^{p-1} \left(\frac{h^*_{f,\mathrm{RCC}}}{p}\right)^p .$$

Note that $h^*_{f,\mathrm{RCC}} = h^*_{\mathrm{RCC}}$, and hence we obtain

$$\left(\frac{2}{\max_{i\in V} d_i}\right)^{p-1} \left(\frac{h^*_{\mathrm{RCC}}}{p}\right)^p \leq \lambda_p .$$

(Note that this bound is tighter than the lower bound from Theorem 5.1). The above inequality can be reformulated as

$$h^*_{\mathrm{RCC}} \leq p \left(\max_{i\in V} d_i\right)^{\frac{p-1}{p}} \left(\frac{\lambda_p}{2^{p-1}}\right)^{\frac{1}{p}} .$$

Using that $\lambda_p \leq 2^{p-1} h_{\mathrm{RCC}}$, we obtain

$$h^*_{\mathrm{RCC}} \leq p \left(\max_{i\in V} d_i\right)^{\frac{p-1}{p}} \left(h_{\mathrm{RCC}}\right)^{\frac{1}{p}} .$$

As shown by Amghibech [1], in the normalized case one has the inequality

$$\lambda_p \geq 2^{p-1} \left(\frac{h_{\mathrm{NCC}}}{p}\right)^p .$$

Analogously to the unnormalized case one can show the stronger statement

$$\lambda_p \geq 2^{p-1} \left(\frac{h^*_{\mathrm{NCC}}}{p}\right)^p \geq 2^{p-1} \left(\frac{h_{\mathrm{NCC}}}{p}\right)^p .$$

The first inequality can be reformulated as

$$h^*_{\mathrm{NCC}} \leq p \left(\frac{\lambda_p}{2^{p-1}}\right)^{\frac{1}{p}} ,$$

and with $\lambda_p \leq 2^{p-1} h_{\mathrm{NCC}}$ one obtains the result in the normalized case. □

References

[1] S. Amghibech. Eigenvalues of the discrete p-Laplacian for graphs. *Ars Combin.*, 67:283–302, 2003.

[2] Thomas Bühler and Matthias Hein. Spectral Clustering based on the graph p-Laplacian. In *Proceedings of the 26th International Conference on Machine Learning (ICML)*. To appear, 2009.