Discrete p -bilaplacian Operators on Graphs

Imad El Bouchairi^(✉), Abderrahim El Moataz^(✉), and Jalal Fadili^(✉)

Normandie Univ, ENSICAEN, UNICAEN, CNRS, GREYC, Caen, France

imad.elbouchairi@gmail.com, abderrahim.elmoataz@unicaen.fr,

Jalal.Fadili@greyc.ensicaen.fr

Abstract. In this paper, we first introduce a new family of operators on weighted graphs called p -bilaplacian operators, which are the analogue on graphs of the continuous p -bilaplacian operators. We then turn to study regularized variational and boundary value problems associated to these operators. For instance, we study their well-posedness (existence and uniqueness). We also develop proximal splitting algorithms to solve these problems. We finally report numerical experiments to support our findings.

Keywords: p -bilaplacian · Weighted graphs · Regularization · Boundary value problems · Proximal splitting

1 Introduction

Regularized variational problems and partial differential equations (PDEs) play an important role in mathematical modeling throughout applied and natural sciences. For instance, many variational problems and PDEs have been studied to model and solve important problems in a variety of areas such as, e.g., in physics, economy, data processing, computer vision. In particular they have been very successful in image and signal processing to solve a wide spectrum of applications such as isotropic and anisotropic filtering, inpainting or segmentation.

In many real-world problems, such as in machine learning and mathematical image processing, the data is discrete, and graphs constitute a natural structure suited to their representation. Each vertex of the graph corresponds to a datum, and the edges encode the pairwise relationships or similarities among the data. For the particular case of images, pixels (represented by nodes) have a specific organization expressed by their spatial connectivity. Therefore, a typical graph used to represent images is a grid graph. For the case of unorganized data such as point clouds, a graph can also be built by modeling neighborhood relationships between the data elements. For these reasons, there has been recently a wave of interest in adapting and solving nonlocal variational problems and PDEs on data which is represented by arbitrary graphs and networks. Using this framework, problems are directly expressed in a discrete setting where an appropriate discrete differential calculus can be proposed; see e.g., [4, 5] and references therein.

This mimetic approach consists of replacing continuous differential operators, e.g., gradient or divergence, by reasonable discrete analogues, which makes it possible to transfer many important tools and results from the continuous setting.

Contributions. In this work, we introduce a novel class of p -bilaplacian operators on weighted graphs, which can be seen as proper discretizations on graphs of the classical p -bilaplacian operators [9]. Building upon this definition, we study a corresponding regularized variational problem as well as a boundary value problem. The latter naturally gives rise to p -biharmonic functions on graphs and equivalent definitions of p -biharmonicity [8]. For these two problems, we start by establishing their well-posedness (existence and uniqueness). We then turn to developing proximal splitting algorithms to solve them, appealing to sophisticated tools from non-smooth optimization. Numerical results are reported to support the viability of our approach.

2 Notations and Preliminary Results

Throughout this paper, we assume that $G = (V, E, \omega)$ is a finite connected undirected weighted graph without loops and parallel edges, where V is the set of vertices, E is the set of edges, and the symmetric function $\omega : V \times V \rightarrow [0, 1]$ is the weight function. We denote by $(x, y) \in E$ the edge that connects the vertices x and y , and we write $x \sim y$ for two adjacent vertices. For two vertices $x, y \in V$ with $x \not\sim y$ we set $\omega(x, y) = \omega(y, x) = 0$ and thus the set of edges E can be characterized by the support of the weight function ω , i.e., $E = \{(x, y) | \omega(x, y) > 0\}$.

Let $\mathcal{H}(V) \stackrel{\text{def}}{=} \{u : x \in V \mapsto u(x) \in \mathbb{R}\}$ be the vector space of real-valued functions on the vertices of the graph. For a function $u \in \mathcal{H}(V)$ the $\ell^p(V)$ -norm of u is given by

$$\|u\|_p = \left(\sum_{x \in V} |u(x)|^p \right)^{\frac{1}{p}}, \quad 1 \leq p < \infty, \quad \text{and} \quad \|u\|_\infty = \max_{x \in V} |u(x)|.$$

We define in a similar way $\mathcal{H}(E)$ as the vector space of all real-valued functions on the edges of the graph.

Let $u \in \mathcal{H}(V)$ and $x, y \in V$. The (nonlocal) gradient operator is defined as

$$\begin{aligned} \nabla_\omega u : \mathcal{H}(V) &\rightarrow \mathcal{H}(E) \\ u &\mapsto U, \quad U(x, y) = \sqrt{\omega(x, y)}(u(y) - u(x)), \forall (x, y) \in V. \end{aligned}$$

This is a linear antisymmetric operator whose adjoint is the (nonlocal) weighted divergence operator denoted div_ω . It is easy to show that

$$\begin{aligned} \text{div}_\omega : \mathcal{H}(E) &\rightarrow \mathcal{H}(V) \\ U &\mapsto u, \quad u(x) = \sum_{y \sim x} \sqrt{\omega(x, y)}(U(y, x) - U(x, y)), \forall x \in V. \end{aligned}$$

For $1 < p < \infty$, the anisotropic graph p -Laplacian operator $\Delta_{\omega,p} : \mathcal{H}(V) \rightarrow \mathcal{H}(V)$ is thus defined by

$$\begin{aligned}\Delta_{\omega,p}u(x) &\stackrel{\text{def}}{=} \operatorname{div}_{\omega}(|\nabla_{\omega}u|^{p-2}\nabla_{\omega}u)(x) \\ &= 2 \sum_{y \sim x} (\omega(x,y))^{\frac{p}{2}} |u(y) - u(x)|^{p-2} (u(y) - u(x)), \forall x \in V.\end{aligned}$$

Unless stated otherwise, in the rest of the paper, we assume $p \in]1, +\infty[$.

3 p -biharmonic Functions on Graphs

We define p -biharmonic functions on graphs inspired by the way p -harmonic functions were introduced in [8] for networks. Let's consider the following functional

$$\mathcal{F}_d(u; p) \stackrel{\text{def}}{=} \frac{1}{p} \|\Delta_{\omega,2}u\|_p^p. \quad (1)$$

Observe that $\Delta_{\omega,2}$ is the standard Laplacian on a graphs, which is a self-adjoint operator.

Definition 1. We define the p -bilaplacian operator for a function $u \in \mathcal{H}(V)$ by

$$\Delta_p^2 u(x) \stackrel{\text{def}}{=} \Delta_{\omega,2}(|\Delta_{\omega,2}u|^{p-2}\Delta_{\omega,2}u)(x), \quad x \in V.$$

Definition 2. Let $A \subset V$. We say that a function u is p -biharmonic on A if it is a minimiser of the functional $\mathcal{F}_d(\cdot; p)$ among functions in V with the same values in $A^c = V \setminus A$, that is, if

$$\mathcal{F}_d(u; p) \leq \mathcal{F}_d(v; p)$$

for every function $v \in \mathcal{H}(V)$, with $u = v$ in A^c .

Inspired by [8], existence and uniqueness of p -biharmonic functions can be established using standard arguments.

Proposition 1. Let A subset of V and $u \in \mathcal{H}(V)$. The following assertions are equivalent:

- (i) the function u is p -biharmonic on A .
- (ii) the function u satisfies

$$\sum_{x \in V} |\Delta_{\omega,2}(u)(x)|^{p-2} \Delta_{\omega,2}(u)(x) \Delta_{\omega,2}w(x) = 0, \quad x \in A, \quad (2)$$

for every function $w \in \mathcal{H}(V)$, with $w = 0$ in A^c .

- (iii) the function u solves

$$\Delta_p^2 u(x) = 0, \quad \text{for all } x \in A.$$

4 p -bilaplacian Dirichlet Problem on Graphs

Consider the following boundary value (Dirichlet) problem

$$\begin{cases} \Delta_p^2 u = 0, & \text{on } A \\ u = g, & \text{on } A^c, \end{cases} \quad (3)$$

where $g \in \mathcal{H}(V)$, $p \in]1, +\infty[$, $A \subset V$ and $A^c = V \setminus A$. Observe that since the graph G is connected, there always exists a path connecting any pair vertices in $A \times A^c$. Our goal now is to establish well-posedness of (3). This will be derived using Dirichlet's variational principe (hence the subscript d in $\mathcal{F}_d$), which, in view of Proposition 1, amounts to equivalently studying the minimization problem

$$\min\{\mathcal{F}_d(u; p) : u \in \mathcal{H}_g(V)\}, \quad (4)$$

where $\mathcal{H}_g(V) = \{u \in \mathcal{H}(V) : u = g \text{ on } A^c\}$ is the subspace of the functions with a zero "trace".

Theorem 1. *The problem (3) has a unique solution.*

Proof. Let $\iota_{\mathcal{H}_g(V)}$ be the indicator function of $\mathcal{H}_g(V)$, i.e. it is 0 on $\mathcal{H}_g(V)$ and $+\infty$ otherwise. By the Poincaré-type inequality established in Lemma 1, we get that $\mathcal{F}_d(\cdot; p) + \iota_{\mathcal{H}_g(V)}$ is coercive. Since this objective is lower semicontinuous (lsc) by closedness of $\mathcal{H}_g(V)$ and continuity of $\mathcal{F}_d(\cdot; p)$, (4) has a minimizer. This together with strict convexity of $\mathcal{F}_d(\cdot; p)$ on $\mathcal{H}_g(V)$ (see Lemma 2) then entails uniqueness.

Lemma 1. *There is $\lambda = \lambda(\omega, V, A^c) > 0$ such that*

$$\lambda \sum_{x \in V \setminus A^c} |u(x)|^2 \leq \sum_{x \in V} \sum_{y \sim x} \omega(x, y) |u(y) - u(x)|^2 + \sum_{x \in A^c} |g(x)|^p, \quad (5)$$

for all $u \in \mathcal{H}_g(V)$. Thus $\mathcal{F}_d(\cdot; p)$ is coercive on $\mathcal{H}_g(V)$.

Proof. Set

$$\begin{aligned} S_0 &= A^c; \\ S_1 &= \{x \in V \setminus S_0 : \exists y \in S_0; y \sim x\}, \\ S_{j+1} &= \{x \in V \setminus (\cup_{k=0}^j S_k) : \exists y \in S_j; y \sim x\}, \quad j = 1, 2, \dots \end{aligned}$$

Since the graph G is connected, there is $l \in \mathbb{N}$ such that $\{S_j\}_{j=0}^l$ forms a partition of V . By Jensen's inequality, we have

$$\begin{aligned} \sum_{x \in V} \sum_{y \sim x} \omega(x, y) |u(y) - u(x)|^2 &\geq \sum_{x \in S_j} \sum_{y \in S_{j-1}} \omega(x, y) |u(y) - u(x)|^2, \\ &\geq \alpha \sum_{x \in S_j} |u(x)|^2 - \beta \sum_{y \in S_{j-1}} |u(y)|^2, \end{aligned}$$

$j = 1, \dots, l$, where $\alpha = \frac{1}{2} \min\{\omega(x, y) : (x, y) \in E\}$ and $\beta = \sum_{x, y \in V} \omega(x, y)$. Since $u = g$ on $S_0 = A^c$ and $\{S_j\}_{j=0}^l$ forms a partition of V , it is easy to see that there exists $\hat{\lambda} = \hat{\lambda}(\omega, V, A^c) > 0$ such that

$$\sum_{x \in V \setminus A^c} |u(x)|^2 \leq \hat{\lambda} \sum_{x \in V} \sum_{y \sim x} \omega(x, y) |u(y) - u(x)|^2 + \hat{\lambda} \sum_{x \in A^c} |g(x)|^2.$$

We arrive at the coercivity result by taking $\lambda = \hat{\lambda}^{-1}$.

Let $C_g = 2 \sum_{x \in A^c} |g(x)|^2$. We then have from Hölder and Young inequalities

$$\begin{aligned} \lambda \|u\|_2^2 &\leq \sum_{x \in V} \sum_{y \sim x} \omega(x, y) |u(y) - u(x)|^2 + C_g \\ &= - \sum_{x \in V} u(x) \Delta_{\omega, 2} u(x) + C_g \\ &\leq \frac{\varepsilon}{2} \|u\|_q^2 + \frac{p}{2\varepsilon} \mathcal{F}_d(u; p)^{2/p} + C_g \end{aligned}$$

where $1/p + 1/q = 1$. Since the norms are equivalent in any finite-dimensional vector space, there exists $C(n) > 0$ with $n = |V|$, such that

$$\lambda \|u\|_2^2 \leq C(n)^2 \frac{\varepsilon}{2} \|u\|_2^2 + \frac{p}{2\varepsilon} \mathcal{F}_d(u; p)^{2/p} + C_g.$$

Choosing $\varepsilon = 2\rho\lambda/C(n)^2$, $\rho \in]0, 1[$, we get

$$(1 - \rho)\lambda \|u\|_2^2 \leq \frac{pC(n)^2}{4\rho\lambda} \mathcal{F}_d(u; p)^{2/p} + C_g,$$

whence coercivity of $\mathcal{F}_d(\cdot; p)$ follows immediately.

Lemma 2. *The functional $\mathcal{F}_d(\cdot; p)$ is strictly convex on $\mathcal{H}_g(V)$.*

Proof. Assume that $\mathcal{F}_d(\cdot; p)$ is not strictly convex on $\mathcal{H}_g(V)$. Then there exist $u, v \in \mathcal{H}_g(V)$ with $u \neq v$ such that $\tau \mathcal{F}_d(u; p) + (1 - \tau) \mathcal{F}_d(v; p) = \mathcal{F}_d(\tau u + (1 - \tau)v; p)$ for all $\tau \in]0, 1[$. But since the function $t \mapsto t^p$ is strictly convex on $\mathbb{R}^+$ for $p \in]1, +\infty[$, this equality entails that $\Delta_{\omega, 2} u = \Delta_{\omega, 2} v$ on V , hence on A . Clearly $w = u - v$ satisfies

$$\begin{cases} \Delta_{\omega, 2} w = 0, & \text{on } A \\ w = 0, & \text{on } A^c. \end{cases}$$

But we know from [8, Theorem 3.11 and Corollary 3.16] that $w = 0$ on V , i.e., $u = v$ on V , leading to a contradiction.

5 p -bilaplacian Variational Problem on Graphs

In this section, we consider the following minimization problem, which is valid for any $p \in [1, +\infty]$ ¹,

$$\min_{u \in \mathcal{H}(V)} \left\{ E(u; p) \stackrel{\text{def}}{=} \frac{1}{2} \|f - Au\|_2^2 + \lambda \mathcal{F}_d(u; p) \right\}, \quad (6)$$

¹ Obviously $\lim_{p \rightarrow +\infty} \frac{1}{p} \|\cdot\|_p^p = \iota_{\|u\|_\infty \leq 1}$.

where $A : \mathcal{H}(V) \rightarrow \mathcal{H}(V)$ is a linear operator, $f \in \mathcal{H}(V)$, $\lambda > 0$ is the regularization parameter, and $\mathcal{F}_d(\cdot; p)$ is given by (1). Problems of the form (6) can be of great interest for graph-based regularization in machine learning and inverse problems in imaging; see [7] and references therein. Problem (6) is well-posed under standard assumptions.

Theorem 2. *The set of minimizers of $E(\cdot; p)$ is non-empty and compact if and only if $\ker(A) \cap \ker(\Delta_{\omega,2}) = \{0\}$. If, moreover, either A is injective or $p \in]1, +\infty[$, then $E(\cdot; p)$ has a unique minimizer.*

Proof. For any proper lsc convex function f , recall its recession function from [11, Chapter 2], denoted f_∞ . We have from the calculus rules in [11, Chapter 2] that

$$E_\infty(d; p) = \lambda \left(\frac{1}{p} \|\cdot\|_p^p \right)_\infty (\Delta_{\omega,2}d) + \frac{1}{2} (\|f - \cdot\|_2^2)_\infty (Ad).$$

Since $\frac{1}{p} \|\cdot\|_p^p$ and $\|f - \cdot\|_2^2$ are non-negative and coercive, we have from [11, Proposition 3.1.2] that their recession functions are positive for any non-zero argument. Equivalently,

$$E_\infty(d; p) > 0, \quad \forall d \notin \ker(A) \cap \ker(\Delta_{\omega,2}).$$

Thus $E_\infty(d; p) > 0$ for all $d \neq 0$ if and only if $\ker(A) \cap \ker(\Delta_{\omega,2}) = \{0\}$. Equivalence with the existence and compactness assertion follows from [11, Proposition 3.1.3].

Let's turn to uniqueness. When A is injective, the claim follows from strict (in fact strong convexity) of the data fidelity term. Suppose now that $p \in]1, +\infty[$. By strict convexity of $\frac{1}{p} \|\cdot\|_p^p$ and $\|f - \cdot\|_2^2$, a standard contradiction argument shows that for any pair of minimizers u^* and v^* , we have $u^* - v^* \in \ker(A) \cap \ker(\Delta_{\omega,2})$. This yields the uniqueness claim under the stated assumption.

6 Algorithms and Numerical Results

To solve both (3) and (6), we adopt a primal-dual proximal splitting (PDS) framework with an appropriate splitting of the functions and linear operators.

6.1 A PDS for the Boundary Value Problem (3)

Problem (3) is equivalent to (4). The latter takes the form

$$\min_{u \in \mathcal{H}_V} F(\Delta_{\omega,2}u) + G(u), \quad \text{where} \quad F(u) = \frac{1}{p} \|u\|_p^p, \quad G(u) = \iota_{\mathcal{H}_g(V)}(u). \quad (7)$$

The latter can be solved with the following PDS iterative scheme [3], which reads in this case

$$\begin{aligned} u^{k+1} &= \text{proj}_{\mathcal{H}_g(V)}(u^k - \tau \Delta_{\omega,2}v^k) \\ v^{k+1} &= \text{prox}_{\frac{\sigma}{q} \|\cdot\|_q^q}(v^k + \sigma \Delta_{\omega,2}(2u^{k+1} - u^k)), \end{aligned} \quad (8)$$

where $\tau, \sigma > 0$, $\text{proj}_{\mathcal{H}_g(V)}$ is the orthogonal projector on the subspace $\mathcal{H}_g(V)$ (which has a trivial closed form), $1/p + 1/q = 1$, and $\text{prox}_{\frac{\sigma}{q}\|\cdot\|_q^q}$ is the proximal mapping of the proper lsc convex function $\frac{\sigma}{q}\|\cdot\|_q^q$. The latter can be computed easily, see [7] for details. Combining [3, Theorem 1], Proposition 1 and Theorem 1, the convergence guarantees of (8) are summarized in the following proposition.

Proposition 2. *If $\tau\sigma\|\Delta_{\omega,2}\|^2 < 1$, then the sequence $(u^k, v^k)_{k \in \mathbb{N}}$ provided by (8) converges to (u^*, v^*) , where u^* is a solution to (3), which is unique if $p \in]1, +\infty[$.*

6.2 A PDS for the Variational Problem (6)

For simplicity and space limitation, we restrict ourselves here to the case where A is the identity. In this case, inspired by the work in [6], we use the (accelerated) FISTA iterative scheme [2, 10] to solve the Fenchel-Rockafellar dual problem of (6), and recover the primal solution by standard extremality relationships. Our scheme reads in this case

$$\begin{aligned} y^k &= v^k + \frac{k-1}{k+b}(v^k - v^{k-1}) \\ v^{k+1} &= \text{prox}_{\gamma \frac{\Delta}{q}\|\cdot\|_q^q / \lambda \|\cdot\|_q^q} (y^k + \gamma \Delta_{\omega,2}(f - \Delta_{\omega,2}y^k)) \\ u^{k+1} &= f - \Delta_{\omega,2}v^{k+1} \end{aligned} \quad (9)$$

where $\gamma \in]0, \|\Delta_{\omega,2}\|^{-2}]$, $b > 2$.

Combining Theorem 2, [6, Theorem 2], [1, Theorem 1.1], the scheme (9) has the following convergence guarantees.

Proposition 3. *The sequence $(u^k)_{k \in \mathbb{N}}$ converges to u^* , the unique minimizer of (6), at the rate $\|u^k - u^*\|_2 = o(1/k)$.*

6.3 Numerical Results

We apply the scheme (9) to solve (6) in order to denoise a function f defined on a 2-D point cloud. We apply (3) in a semisupervised classification problem which amounts to finding the missing labels of a label function defined on a 2-D point cloud. The nodes of the graph are the points in the 2-D cloud and $u(x)$ is the value at a point/vertex x . We choose the nearest neighbour graph with the standard weighting kernel $\exp(-|\mathbf{x} - \mathbf{y}|)$ when $|\mathbf{x} - \mathbf{y}| \leq \delta$ and 0 otherwise, where $\mathbf{x}$ and $\mathbf{y}$ are the 2-D spatial coordinates of the points in the cloud. The original point cloud used in our numerical experiments consists of $N = 2500$ points that are not on a regular grid. For the variational problem, the noisy observation is generated by adding a white Gaussian noise of standard deviation 0.5 to the original data, see Fig. 1(a). For the Dirichlet problem, the initial label function takes the values of the original data on a set of points/vertices where this set corresponds to the boundary data, it is chosen randomly and is equal to $N/4$ of the original points/vertices, see Fig. 1(b).

Fig. 1. (a): results for denoising with $p = 1$. (b): results for a semisupervised classification problem with $p = 1.1$. For each setting, we show the original function on the point cloud, its observed version, and the result provided by each of our algorithms.

References

1. Attouch, H., Peypouquet, J.: The rate of convergence of Nesterov's accelerated forward-backward method is actually faster than $1/k^2$. *SIAM J. Optim.* **26**(3), 1824–1834 (2016)

2. Chambolle, A., Dossal, C.: On the convergence of the iterates of the “fast iterative shrinkage/thresholding algorithm”. *J. Optim. Theory Appl.* **166**(3), 968–982 (2015). <https://doi.org/10.1007/s10957-015-0746-4>
3. Chambolle, A., Pock, T.: A first-order primal-dual algorithm for convex problems with applications to imaging. *J. Math. Imaging Vis.* **40**(1), 120–145 (2011). <https://doi.org/10.1007/s10851-010-0251-1>
4. Elmoataz, A., Lézoray, O., Bogleux, S.: Nonlocal discrete regularization on weighted graphs: a framework for image and manifold processing. *IEEE Trans. Image Process.* **17**(7), 1047–1060 (2008)
5. Elmoataz, A., Toutain, M., Tenbrinck, D.: On the p -Laplacian and ∞ -Laplacian on graphs with applications in image and data processing. *SIAM J. Imaging Sci.* **8**(4), 2412–2451 (2015)
6. Fadili, M.J., Peyré, G.: Total variation projection with first order schemes. *IEEE Trans. Image Process.* **20**(3), 657–669 (2010)
7. Hafiene, Y., Fadili, M.J., Elmoataz, A.: Continuum limits of nonlocal p -Laplacian variational problems on graphs. *SIAM J. Imaging Sci.* **12**(4), 1772–1807 (2019)
8. Holopainen, I., Soardi, P.M.: p -Harmonic functions on graphs and manifolds. *Manuscripta Mathematica* **94**(1), 95–110 (1997)
9. Katzourakis, N., Pryer, T.: On the numerical approximation of p -biharmonic and ∞ -biharmonic functions. *Num. Methods Partial Differ. Equ.* **35**(1), 155–180 (2019)
10. Nesterov, Y.: A method for solving the convex programming problem with convergence rate $O(1/k^2)$. *Dokl. Akad. Nauk SSSR* **269**(3), 543–547 (1983)
11. Auslender, A., Teboulle, M.: *Asymptotic Cones and Functions in Optimization and Variational Inequalities*. Springer, New York (2003). <https://doi.org/10.1007/b97594>