

Rayleigh quotient minimization for absolutely one-homogeneous functionals

Tal Feld, Jean-François Aujol, Guy Gilboa, Nicolas Papadakis

► **To cite this version:**

Tal Feld, Jean-François Aujol, Guy Gilboa, Nicolas Papadakis. Rayleigh quotient minimization for absolutely one-homogeneous functionals. 2018. <hal-01864129v2>

HAL Id: hal-01864129

<https://hal.archives-ouvertes.fr/hal-01864129v2>

Submitted on 30 Aug 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Rayleigh quotient minimization for absolutely one-homogeneous functionals

Tal Feld*, Jean-François Aujol †, Guy Gilboa ‡, Nicolas Papadakis §

August 30, 2018

Abstract

In this paper we examine the problem of minimizing generalized Rayleigh quotients of the form $J(u)/H(u)$, where both J and H are absolutely one-homogeneous functionals. This can be viewed as minimizing J where the solution is constrained to be on a generalized sphere with $H(u) = 1$, where H is any norm or semi-norm. The solution admits a nonlinear eigenvalue problem, based on the subgradients of J and H . We examine several flows which minimize the ratio. This is done both by time-continuous flow formulations and by discrete iterations. We focus on a certain flow, which is easier to analyze theoretically, following the theory of Brezis on flows with maximal monotone operators. A comprehensive theory is established, including convergence of the flow. We then turn into a more specific case of minimizing graph total variation on the L^1 sphere, which approximates the Cheeger-cut problem. Experimental results show the applicability of such algorithms for clustering and classification of images.

1 Introduction

In various fields of science and engineering one seeks to minimize a ratio of two functionals (Schwetlick & Schreiber 2012, Güttel & Tisseur 2017). In this work we consider the generalized nonlinear Rayleigh quotient problem

$$\min_u R(u) := \frac{J(u)}{H(u)}, \quad (1.1)$$

*Technion - IIT, Haifa, stmfeld@gmail.com

†IMB, Univ. Bordeaux, Jean-Francois.Aujol@math.u-bordeaux.fr

‡Technion - IIT, Haifa, Guy.Gilboa@ee.technion.ac.il

§CNRS, IMB, Nicolas.Papadakis@math.u-bordeaux.fr

where the two functionals $J : X \rightarrow \mathbb{R}$ and $H : X \rightarrow \mathbb{R}$ are with full domain, proper, convex, lower semi-continuous (lsc) and absolutely one-homogeneous. X is a finite dimensional hilbert space.

For J being the total variation (TV) functional and H the L^2 norm, a substantial theory has been developed. It was shown in (Andreu, Ballester, Caselles & Mazón 2001, Bellettini, Caselles & Novaga 2002) that local minimizers admit the nonlinear eigenvalue problem

$$\lambda u \in \partial J(u). \quad (1.2)$$

Moreover, in the continuous setting, in an unbounded domain \mathbb{R}^2 , a family of such nonlinear eigenfunctions was precisely characterized geometrically. It was shown that certain characteristic functions of a convex set C , referred to as a *calibrable set*, admit (1.2). These sets have smooth boundaries, where the maximal curvature on the boundary κ_{max} is bounded by

$$\kappa_{max} \leq \frac{P(C)}{|C|},$$

where $P(C)$ is the perimeter of the set C and $|C|$ is its area. For instance, disks admit this condition and are stable structures, preserved within the TV regularization (Meyer 2001). We will examine below more general notions of calibrable sets on graphs.

The problem of (1.1) can be recast as a constrained optimization problem

$$\min_u J(u) \quad s.t. \quad H(u) = 1. \quad (1.3)$$

When H is a norm, such minimization amounts to seeking a minimizer of J on the unit sphere defined by H . For some analysis of variational minimization on spheres see (Giaquinta, Modica & Souček 1989, Rivière 1995). Such problems are common in image processing and computer vision in various contexts. For example, one can denoise normals based on total-variation on the unit sphere \mathcal{S}^2 as in (Lellmann, Strekalovskiy, Koetter & Cremers 2013) or minimize second order energies on the sphere (Bacák, Bergmann, Steidl & Weinmann 2016). In (Sun, Qu & Wright 2017) it is shown how efficient dictionary learning can be formulated as a minimization problem over either the ℓ^∞ or ℓ^2 spheres.

By using Lagrange multipliers, the optimality condition with respect to u yields

$$p = \lambda q, \quad p \in \partial J(u), \quad q \in \partial H(u), \quad (1.4)$$

where $\partial J(u)$, $\partial H(u)$ denote the subdifferential of $J(u)$ and $H(u)$, respectively. Eq. (1.4) can be viewed as double-nonlinear eigenvalue problem, with λ the eigenvalue. Thus, as in classical linear algebra, extremal points of generalized linear Rayleigh quotients are eigenfunctions with respect to the associated operators.

Hein and Buhler (Hein & Bühler 2010) proposed a nonlinear inverse-power-method (IPM), based on a series of iterations, to minimize Rayleigh quotients of

convex functionals. This was applied to classification problems on graphs. However, not all the theoretical aspects were fully developed. Moreover, the procedure did not include any associated time-step. We found this is a crucial factor to gain good performance in such non-convex minimization algorithms. Bresson et al (Bresson, Laurent, Uminsky & Brecht 2012) pointed out some drawbacks of (Hein & Bühler 2010) and proposed a modified version. They also introduced an algorithm for solving the Rayleigh quotient for the case $J = TV$, $H = L^1$, which approximates the Cheeger cut problem. In (Nossek & Gilboa 2018) a continuous flow for minimizing Rayleigh quotients with J being one-homogeneous and H the square L^2 norm was proposed. A comprehensive theoretical analysis was recently performed for that flow in (Aujol, Gilboa & Papadakis 2018) and an analog flow with full convergence proof was proposed.

A classical way to minimize $R(u)$ is by using a gradient descent flow

$$u_t = -\nabla R(u).$$

Taking the variational derivative of $R(u)$, with $q \in \partial H(u)$, $p \in \partial J(u)$, the gradient descent flow is

$$u_t = \frac{J(u)q - H(u)p}{H^2(u)}. \quad (1.5)$$

The flow can also be written as

$$u_t = \frac{R(u)q - p}{H(u)}.$$

This flow is hard to analyze theoretically, mainly due to the division by $H(u)$.

Our study is inspired by the book of Brezis (Brezis 1973). In particular, it gives existence and uniqueness results for flows governed by a maximal monotone operator with possible Lipschitz perturbations. The analysis is valid for flows of the form

$$u_t = A(u) - p. \quad (1.6)$$

We therefore propose the following flow to minimize $R(u)$,

$$u_t = R(u)q - p. \quad (1.7)$$

This is essentially a gradient-descent type flow, without the division by $H(u)$, which can be interpreted as a dynamic rescaling of the time parameter. We show that the flow reduces monotonically the quotient $R(u)$ and that the steady state admits the nonlinear eigenvalue problem (1.4). We present in the paper a full time discrete analysis and provide certain conditions under which a time continuous flow converges.

In addition, we examine another flow, that minimizes the log of the Rayleigh quotient,

$$u_t = -\nabla(\log R(u)),$$

which can be written as

$$u_t = \frac{q}{H(u)} - \frac{p}{J(u)}. \quad (1.8)$$

This is motivated by a widely used practice of using the log of a function involving multiplicative expressions. It is commonly employed in statistics and machine learning algorithms, such as maximum likelihood estimation and policy learning. The flow is essentially a time rescaling of (1.7) by $1/J(u)$. We note that it is not in the form of (1.6) and therefore harder to analyze. Nevertheless, in a time discrete setting one can show the decrease of $R(u)$. We mention this flow in the context of the Cheeger cut problem as we found out that numerically it has very stable properties. In particular, its good performance is highly resilient to the choice of the discrete time step. We can therefore increase the time step and speed up numerical convergence.

We farther examine a modification of (1.7) by removing the mean of q , denoted by \bar{q} , such that the flow becomes

$$u_t = R(u)(q - \bar{q}) - p. \quad (1.9)$$

This modification avoids converging to constant functions, and maintains the validity of our analysis of (1.7). We note that for the Cheeger cut problem (1.9) is realized by removing the median of u . This ensures the existence of q with zero mean. Similar ideas were suggested in (Bühler & Hein 2009, Bresson et al. 2012).

As both J and H are one-homogeneous, the Rayleigh quotient R is unaffected by a multiplication by a constant, $R(\alpha u) = R(u)$, $\forall \alpha \in \mathbb{R}_{/0}$. We found it convenient to have solutions of unit L^2 norm. The minimization problem (1.1) is in some cases replaced by

$$\min_{\|u\|_2=1} R(u) := \frac{J(u)}{H(u)}. \quad (1.10)$$

We turn now to the main contributions and outline the plan of the paper.

1.1 Main contributions and plan of paper

The main contributions of the paper are:

1. An iterative algorithm for minimizing the Rayleigh quotient for general one-homogeneous functionals is presented and is fully analyzed. It is understood as a time discretization of the continuous flow of (1.7). We prove convergence of the iterative scheme.
2. We examine in more details the case of J being a difference -based regularizer on graphs, such as the nonlocal (graph) total-variation, and H an L^p norm (with $p \geq 1$). We propose an algorithm that is able to generate bi-valued functions (calibrable sets) for functionals J admitting a co-area formula.

3. The above theory and iterative method are used to estimate the optimal Cheeger cut by minimizing the TV to L^1 ratio, subject to certain constraints, using a modified algorithm.
4. We show experimentally the ability of our algorithm to obtain high quality classifications by the Cheeger cut estimate. Moreover, we show that the proposed log-flow (1.8) is highly robust to the time-step choice (compared to (1.7) and (1.5)), where (Bresson et al. 2012) can be understood as a discrete realization of (1.5) specifically for the Cheeger cut problem.
5. Finally, we study the existence and uniqueness of the solution of (1.7) in the time continuous setting. We need the mapping $q(u) \in \partial H(u)$ to be locally Lipschitz continuous. For $u \in X$, $H(u) = \|u\|_\alpha$ with $\alpha \in [2, +\infty)$ we readily have this. In the more general case, Moreau-Yosida regularization is used to perform the analysis.

The plan of the paper is as follows: the introductory part is concluded with some definitions and properties related to one-homogeneous functionals, nonlinear eigenfunctions and calibrable sets. In Section 2 we introduce the main numerical algorithm for minimizing Rayleigh quotients and analyze it. In Section 3 we turn to analyzing functionals J which are regularizers (based on differences) and then more specifically to TV-type regularization admitting a co-area formula. In this case we show how to numerically compute calibrable sets. In Section 4 we focus on the practical problem of seeking an optimal Cheeger cut on a graph. We provide two algorithms for this purpose, motivated by the flows (1.7) and (1.8). In Section 4.1 experimental results for digit classification using our algorithms are shown. In Section 5 the time-continuous flow is analyzed. Additional proofs and a short analysis of the log-flow are given in the appendices.

1.2 Mathematical notations and setting

Absolutely one-homogeneous functionals. We consider two absolutely one homogeneous functionals J and H which are with full domain, proper, convex and lower semi-continuous (lsc). They take as input a function $u : x \in \Omega \rightarrow \mathbb{R}$ defined on a discrete domain Ω of size $|\Omega| = N$. We assume the function u to be an element in some finite dimensional Hilbert space X , embedded with an inner product $\langle \cdot, \cdot \rangle$. We use ℓ_2 and ℓ_1 norms of u defined as $\|u\|_2 = \sqrt{\langle u, u \rangle}$ and $\|u\|_1 = \langle u, \text{sign}(u) \rangle$. Recall that the subgradients of a functional $J : x \in \Omega \rightarrow \mathbb{R}$ is defined as

$$p \in \partial J(u) \Leftrightarrow J(v) - J(u) \geq \langle p, v - u \rangle, \forall v.$$

We summarize several known properties of absolutely one-homogeneous functionals.

Properties 1.1. *An absolutely one homogeneous functional $J : u \in \Omega \rightarrow \mathbb{R}$ admits*

1. $J(\alpha u) = |\alpha|J(u)$
2. $J(u) = \sup_p \langle u, p \rangle - J^*(p)$ where the convex conjugate J^* is the characteristic function of the convex set $K_J = \partial J(0)$.
3. $J(u) = \langle p, u \rangle, \forall p \in \partial J(u)$.
4. If $p \in \partial J(u)$, then $J(v) \geq \langle p, v \rangle, \forall v$.
5. If $\langle p, u \rangle = J(u)$ and $J^*(p) = 0$ then $p \in \partial J(u)$.

In the finite dimensional case, the elements of $K_H = \partial H(0)$ have a norm bounded in ℓ^2 (Burger, Gilboa, Möller, Eckardt & Cremers 2016):

$$\exists 0 < L_J < +\infty \text{ s.t. } \|q\|_2 \leq L_J, \forall q \in \partial J(0). \quad (1.11)$$

Moreover one can characterize the so-called kernel of J as (see e.g. (Burger, Gilboa, Möller, Eckardt & Cremers 2016))

$$K_J^\perp = \{u, \text{ such that } J(u) = 0\} \quad (1.12)$$

which is a linear vectorial space. Finally, from the equivalence of norms, for u of zero mean, there exists a constant $\kappa_J > 0$ for which

$$\|u\|_2 \leq \kappa_J J(u), \forall u \text{ such that } \langle u, \mathbf{1} \rangle = 0. \quad (1.13)$$

Eigenfunctions and calibrable sets. We are interested in eigenfunctions with respect to ∂J that are defined as follows.

Definition 1 (Eigenfunction of ∂J). *An eigenfunction u of ∂J is a function that satisfies the nonlinear eigenvalue problem*

$$\lambda u \in \partial J(u),$$

so that $J(u) = \langle \lambda u, u \rangle = \lambda \|u\|_2^2$ and $\lambda = \frac{J(u)}{\|u\|_2^2} \geq 0$.

We are also interested in what was termed by (Andreu et al. 2001, Bellettini et al. 2002) in the TV context as *calibrable sets*. We generalize this notion by the following definition.

Definition 2 (Calibrable set of J). *A calibrable set u of J is a bivalued eigenfunction of ∂J , such that given two real constants c_1, c_2 and a subdomain $\Omega_1 \subset \Omega$, $u(x) = c_1$ for all $x \in \Omega_1$ and c_2 otherwise.*

2 Rayleigh quotient minimization

In this section the functionals J and H are in the most general form as described in (1.1). We propose an iterative process which follows (1.7), maintaining a unit ℓ^2 norm. Recall our main flow:

$$u_t = R(u)q - p, \quad q \in \partial H(u), p \in \partial J(u).$$

In what follows we focus on the following semi-explicit scheme of the flow,

$$\begin{cases} (u_{k+1/2} - u_k)/dt = R(u_k)q_k - p_{k+1/2}, & q_k \in \partial H(u_k), p_{k+1/2} \in \partial J(u_{k+1/2}) \\ u_{k+1} = u_{k+1/2}/\|u_{k+1/2}\|_2. \end{cases} \quad (2.1)$$

This scheme is associated with the minimization of a convex functional since

$$u_{k+1/2} = \operatorname{argmin}_{u \in X} F(u) := \frac{1}{2dt} \|u - u_k\|_2^2 - R(u_k) \langle q_k, u \rangle + J(u). \quad (2.2)$$

where $u_{k+1/2}$ being a minimizer of F implies that there exist $p_{k+1/2} \in \partial J(u_{k+1/2})$ such that

$$\frac{1}{dt}(u_{k+1/2} - u_k) - R(u_k)q_k + p_{k+1/2} = 0.$$

This leads directly to Algorithm 1.

Algorithm 1: Rayleigh quotient minimization of absolutely one-homogeneous functionals

Data: Absolutely one-homogeneous functionals J, H , positive dt, ϵ .

Result: Local Minimizer u of the Rayleigh quotient $R = J/H$.

Initialization: Initial u_0 with $\|u_0\|_2 = 1$.

while $\|R(u_{k+1}) - R(u_k)\| > \epsilon$ **do**

$$\left| \begin{array}{l} u_{k+1/2} = \operatorname{argmin}_{u \in X} F(u) := \frac{1}{2dt} \|u - u_k\|_2^2 - R(u_k) \langle q_k, u \rangle + J(u). \\ u_{k+1} = u_{k+1/2}/\|u_{k+1/2}\|_2 \end{array} \right.$$

end while

Remark 2.1. Notice that since J and H are absolutely one-homogeneous their subgradients do not change by the normalization step of the flow, i.e $q_{k+1} = q_{k+1/2}$ and $p_{k+1} = p_{k+1/2}$. We also have $R(u_{k+1}) = R(u_{k+1/2})$ as a quotient of two one-homogeneous functionals.

We start our analysis by presenting the following properties.

Lemma 2.1 (Properties of flow (2.1)). *The sequence u_k of Algorithm 1 satisfies the following properties:*

1. $1 = \|u_k\|_2^2 \leq \langle u_{k+1/2}, u_k \rangle \leq \|u_{k+1/2}\|_2^2$.

$$2. \|u_{k+1} - u_k\|_2 \leq \|u_{k+1/2} - u_k\|_2.$$

$$3. 0 < H(u_{k+1/2}) \leq C = L_H + dtL_H (R(u_0)L_H + L_J).$$

Proof. 1. The optimality condition for minimizing $F(u)$ yields $u_{k+1/2} = u_k + dt (R(u_k)q_k - p_{k+1/2})$. Taking the scalar product of $u_{k+1/2}$ with u_k we have

$$\langle u_{k+1/2}, u_k \rangle = \|u_k\|_2^2 + dt (J(u_k) - \langle p_{k+1/2}, u_k \rangle) \geq \|u_k\|_2^2. \quad (2.3)$$

On the other hand, by Cauchy-Schwarz inequality, we have

$$\langle u_{k+1/2}, u_k \rangle \leq \|u_{k+1/2}\|_2 \|u_k\|_2 = \|u_{k+1/2}\|_2, \quad (2.4)$$

so that we get $\|u_k\|_2 \leq \|u_{k+1/2}\|_2$ by combining (2.3) and (2.4).

2. Observing that u_k and u_{k+1} are the orthogonal projections of u_k and $u_{k+1/2}$ on the convex ℓ_2 ball, we conclude that

$$\|u_{k+1} - u_k\|_2 \leq \|u_{k+1/2} - u_k\|_2. \quad (2.5)$$

3. Notice that $\|u_{k+1/2}\|_2 \geq \|u_k\|_2 > 0$ so that $H(u_{k+1/2}) > 0$ from the equivalence of norms. Moreover, using Remark 2.1 we can take the inner product of the equality $u_{k+1/2} = u_k + dt (R(u_k)q_k - p_{k+1/2})$ with q_{k+1} , to have

$$\begin{aligned} H(u_{k+1/2}) &= \langle u_k, q_{k+1} \rangle + dt (R(u_k) \langle q_k, q_{k+1} \rangle - \langle p_{k+1/2}, q_{k+1} \rangle) \\ &\leq H(u_k) + dt (R(u_k) \langle q_k, q_{k+1} \rangle - \langle p_{k+1/2}, q_{k+1} \rangle) \end{aligned}$$

From 1.11 there exists $L_J, L_H < \infty$ such that $\|p\|_2 \leq L_J$ and $\|q\|_2 \leq L_H$, thus

$$\begin{aligned} H(u_{k+1/2}) &\leq L_H + dtL_H (R(u_k)L_H + L_J) \\ &\leq L_H + dtL_H (R(u_0)L_H + L_J), \end{aligned}$$

where the last inequality involves the following Lemma. □

Lemma 2.2 (Monotonicity). *The sequence u_k of Algorithm 1 satisfies $R(u_{k+1}) \leq R(u_k)$.*

Proof. Recalling that $u_{k+1/2}$ minimizes F defined in (2.2), we get

$$\begin{aligned}
F(u_{k+1/2}) &\leq F(u_k) \\
\frac{1}{2dt} \|u_{k+1/2} - u_k\|_2^2 - R(u_k) \langle q_k, u_{k+1/2} \rangle + J(u_{k+1/2}) &\leq -R(u_k) \langle q_k, u_k \rangle + J(u_k) \\
\frac{1}{2dt} \|u_{k+1/2} - u_k\|_2^2 - R(u_k) \langle q_k, u_{k+1/2} \rangle + J(u_{k+1/2}) &\leq 0 \\
\frac{1}{2dt} \|u_{k+1/2} - u_k\|_2^2 + J(u_{k+1/2}) &\leq R(u_k) \langle q_k, u_{k+1/2} \rangle \\
\frac{1}{2dt} \|u_{k+1/2} - u_k\|_2^2 + J(u_{k+1/2}) &\leq R(u_k) H(u_{k+1/2}) \\
\frac{1}{2dt H(u_{k+1/2})} \|u_{k+1/2} - u_k\|_2^2 + R(u_{k+1/2}) &\leq R(u_k) \\
\frac{1}{2dt H(u_{k+1/2})} \|u_{k+1/2} - u_k\|_2^2 + R(u_{k+1}) &\leq R(u_k).
\end{aligned} \tag{2.6}$$

□

Lemma 2.3 (Compactness). *The sequence u_k of Algorithm 1 satisfies $\|u_{k+1} - u_k\|_2^2 \rightarrow 0$.*

Proof. From (2.6) and point (3) in Proposition 2.1 it follows that

$$\begin{aligned}
\frac{1}{2dt H(u_{k+1/2})} \|u_{k+1/2} - u_k\|_2^2 + R(u_{k+1}) &\leq R(u_k) \\
\frac{1}{2dt} \|u_{k+1/2} - u_k\|_2^2 &\leq C (R(u_k) - R(u_{k+1})).
\end{aligned}$$

Summing the previous relation from $k = 0$ to K we have

$$\sum_{k=0}^K \|u_{k+1/2} - u_k\|_2^2 \leq 2dtC (R(u_0) - R(u_{K+1})),$$

which means that $\sum \|u_{k+1/2} - u_k\|_2^2$ converges, hence $\|u_{k+1/2} - u_k\|_2^2 \rightarrow 0$. Then, using inequality (2.5) we conclude that $\|u_{k+1} - u_k\|_2^2 \rightarrow 0$.

□

We finally can show convergence of Algorithm 1.

Theorem 2.1 (Convergence). *Let u_0 in X and u_k is computed by Algorithm 1. Then there exist u , p and q in X such that up to a subsequence $u_k \rightarrow u$, $p_{k+1/2} \rightarrow p$, $q_k \rightarrow q$, $\|u\|_2 = 1$, and*

$$p = R(u)q, \quad q \in \partial H(u), \quad p \in \partial J(u). \tag{2.7}$$

Proof. The goal is to pass to the limit $k \rightarrow +\infty$ in Algorithm 1

$$\begin{cases} u_{k+1/2} &= u_k + dt (R(u_k)q_k - p_{k+1/2}), \quad q_k \in \partial H(u_k), \quad p_{k+1/2} \in \partial J(u_{k+1/2}), \\ u_{k+1} &= \frac{u_{k+1/2}}{\|u_{k+1/2}\|_2}. \end{cases} \quad (2.8)$$

From the previous results, we know that u_k , q_k and $p_{k+1/2}$ are bounded in X . Hence there exist u, q and p in X such that up to a subsequence $u_k \rightarrow u$, $q_k \rightarrow q$, $p_{k+1/2} \rightarrow p$. Since H and J are continuous functions we also have $H(u_{k+1}) \rightarrow H(u)$ and $J(u_k) \rightarrow J(u)$. Then, using Lemma 2.3, we see that we can pass to the limit in Equation (2.8) to get

$$p = R(u)q, \quad q \in \partial H(u), \quad p \in \partial J(u).$$

We have $q_k \in \partial H(u_k)$. This means that for all $v : H(v) - H(u_k) \geq \langle q_k, v - u_k \rangle$. Passing to the limit, we get $H(v) - H(u) \geq \langle q, v - u \rangle$, i.e. $q \in \partial H(u)$. Similarly, we get that $p \in \partial J(u)$, which concludes the proof. \square

We thus have shown that for any initial condition u_0 Algorithm 1 monotonically decreases the Rayleigh quotient until it converges to a local minimizer u of (1.1), which admits the nonlinear eigenvalue problem (1.4).

3 Eigenfunctions and Calibrable sets

In this section we focus on absolutely one homogeneous functionals for which their nullspace contains only constant functions. We thus get the following properties:

Properties 3.1. 1. *Mean-shift invariance:* $J(u + c) = J(u)$.

2. *Kernel of constants:* $J(u) = 0$ iff $u = \text{constant}$.

3. *Subgradients of zero mean:* $\langle p, \mathbf{1} \rangle = 0$ for all p such that $J^*(p) = 0$.

More details on these properties can be seen in (Burger, Gilboa, Moeller, Eckardt & Cremers 2016).

Remark 3.1. *Properties 3.1 are satisfied for functionals $J = J_\beta$ of the form*

$$J_\beta(u) = \sum_{i=1}^N \left(\sum_{j=1}^N w_{ij} |u_i - u_j|^\beta \right)^{1/\beta} \quad (3.1)$$

with $\beta \geq 1$, $w_{ij} = w_{ji}$, $w_{ij} \geq 0$, $\sum_{i'=1}^N w_{i'j} > 0$, $\sum_{j'=1}^N w_{ij'} > 0$, $\forall i, j \in [1; N]^2$.

The flow might converge to a constant function. To avoid this, for functionals satisfying the Properties 3.1, we use the modified flow 1.9, i.e

$$u_t = R(u)(q - \bar{q}) - p, \quad q \in \partial H(u), p \in \partial J(u).$$

As detailed in A, Lemma 2.1, Lemma 2.2 and Theorem 2.3 stay valid for this modified flow. Nonetheless, for J_β functionals, we now have the additional property $\bar{p} = 0$ from point (3) of Properties 3.1. The discrete scheme now reads,

$$\begin{cases} u_{k+1/2} &= u_k + dt \left(R(u_k)(q_k - \bar{q}_k) - p_{k+1/2} \right), \\ u_{k+1} &= \frac{u_{k+1/2}}{\|u_{k+1/2}\|_2}. \end{cases} \quad (3.2)$$

Hence, for u_k of zero mean, $u_{k+1/2}$ will remain of zero mean. This new property, together with the renormalization $\|u_{k+1}\|_2 = 1$, prevents the flow from converging to trivial constant steady states.

3.1 Generating Eigenfunctions

In the case $H_2(u) = \|u\|_2$, we can take $q = \frac{u}{\|u\|_2} \in \partial\|u\|_2 = \nabla\|u\|_2$, since $\|u\|_2 > 0$. The flow (3.2) then becomes

$$\begin{cases} u_{k+1/2} &= u_k + \delta t \left(\frac{J(u_k)}{\|u_k\|_2^2} u_k - p_{k+1} \right), \\ u_{k+1} &= \frac{u_{k+1/2}}{\|u_{k+1/2}\|_2}, \end{cases} \quad (3.3)$$

where we got rid of $\bar{q}_k = \frac{1}{N} \langle u_k, \mathbf{1} \rangle = 0$. At convergence $u_k \rightarrow u$, we have

$$\frac{J(u)}{\|u\|_2^2} u - p,$$

so that $p = \frac{J(u)}{\|u\|_2^2} u \in \partial J(u)$ and u is an eigenfunction of ∂J from Definition 1. This scheme is strongly related to the work of (Aujol et al. 2018), which is dedicated to the flow with $H_2(u) = \|u\|_2$. In (Aujol et al. 2018), the discretization of the flow is more implicit, since $q_{k+1/2}$ is considered instead of q_k .

With the discretization $q_{k+1/2}$, it is nevertheless difficult to generalize convergence results to other functions H other than $\|\cdot\|_2$. The results we get with the new flow (3.2) are thus more general.

3.2 Generating calibrable sets

We now describe how we can construct calibrable sets of J with the flow based on the choice $H_1(u) = \|u\|_1$. In this case, we have $\text{sign}(u) \in \partial\|u\|_1$. For this we will also assume the co-area formula, i.e

$$J(u) = \int_{u_{min}}^{u_{max}} J(\mathbf{1}_{u>\epsilon}) d\epsilon = \int_{u_{min}}^{u_{max}} J(\mathbf{1}_{u<\epsilon}) d\epsilon, \quad (3.4)$$

where u_{min} and u_{max} stand for the minimal and maximal values of u , while $\mathbb{1}_{u>\epsilon}$ (resp. $\mathbb{1}_{u<\epsilon}$) is the characteristic function of the upper (resp. lower) level set of u with respect to level $\epsilon \in [u_{min}, u_{max}]$. We can then deduce the following result.

Proposition 3.1 (Co-area). *Let $p \in \partial J(u)$, then $p \in \partial J(\mathbb{1}_{u>\epsilon})$ for all $\epsilon \in [u_{min}, u_{max}]$.*

Proof. Using the coarea formula, we have

$$\begin{aligned} \int_{u_{min}}^{u_{max}} J(\mathbb{1}_{u>\epsilon}) d\epsilon &= J(u) \\ &= \langle p, u \rangle \\ &= \langle p, \int_{u_{min}}^{u_{max}} \mathbb{1}_{u>\epsilon} d\epsilon + u_{min} \rangle \\ &= \langle p, \int_{u_{min}}^{u_{max}} \mathbb{1}_{u>\epsilon} d\epsilon \rangle + \langle p, u_{min} \rangle \\ &= \int_{u_{min}}^{u_{max}} \langle p, \mathbb{1}_{u>\epsilon} \rangle d\epsilon. \end{aligned}$$

On the one hand, if $p \in \partial J(\mathbb{1}_{u>\epsilon})$ then $\langle p, \mathbb{1}_{u>\epsilon} \rangle = J(\mathbb{1}_{u>\epsilon})$. On the other hand, if there exists $\epsilon \in [u_{min}, u_{max}]$, such that $p \notin \partial J(\mathbb{1}_{u>\epsilon})$ we have $\langle p, \mathbb{1}_{u>\epsilon} \rangle < J(\mathbb{1}_{u>\epsilon})$. The last inequality stems from the fact that $J^*(p) = 0$ and using properties (iv) and (v) of subgradients as stated in Properties 1.1. Hence, by contradiction, we find that $p \in \partial J(\mathbb{1}_{u>\epsilon})$ for all $\epsilon \in [u_{min}, u_{max}]$. \square

Remark 3.2. *The Co-area property in Proposition 3.1 is satisfied for the non local TV functional $J = J_1$ of Property 3.1, i.e*

$$J_1(u) = \sum_{i,j=1}^N w_{ij} |u_i - u_j|. \quad (3.5)$$

In order to easily deal with the non differentiability of $\|u\|_1$ at $u(x) = 0$ we consider $\text{sign}(0) = -1$. Hence we take the specific bivalued subgradient

$$q(u_k) = \mathbb{1}_{u_k>0} - \mathbb{1}_{u_k\leq 0} \in \partial \|u_k\|_1 \quad (3.6)$$

such that

$$\bar{q}(u) = \frac{1}{N} (\|\mathbb{1}_{u>0}\|_1 - \|\mathbb{1}_{u\leq 0}\|_1).$$

Lemma 3.1. *Let u be the solution of the scheme (3.2), for q defined in (3.6), then $v = \mathbb{1}_{u>0} - N^{-1} \|\mathbb{1}_{u>0}\|_1$ is a bivalued eigenfunction of ∂J , i.e calibrable set of J .*

Proof. By definition we have that $\mathbb{1}_{u>0} + \mathbb{1}_{u\leq 0} = 1$ and $\|\mathbb{1}_{u>0}\|_1 + \|\mathbb{1}_{u\leq 0}\|_1 = |\Omega| = N$. From Proposition 2.3 we know that $u_k \rightarrow u$, so that

$$dt \left((R(u) (\mathbb{1}_{u>0} - \mathbb{1}_{u\leq 0}) - N^{-1} (\|\mathbb{1}_{u>0}\|_1 - \|\mathbb{1}_{u\leq 0}\|_1)) - p \right) = 0,$$

for $p \in \partial J(u)$. Hence $p \in \partial J(u)$ is a bivalued function such that

$$\begin{aligned} p &= R(u) (\mathbf{1}_{u>0} - \mathbf{1}_{u\leq 0} + N^{-1} (\|\mathbf{1}_{u\leq 0}\|_1 - \|\mathbf{1}_{u>0}\|_1)) \\ &= R(u) (2\mathbf{1}_{u>0} - 1 + N^{-1} (\|\mathbf{1}_{u\leq 0}\|_1 - \|\mathbf{1}_{u>0}\|_1)) \\ &= R(u) (2\mathbf{1}_{u>0} + N^{-1} (\|\mathbf{1}_{u\leq 0}\|_1 - \|\mathbf{1}_{u>0}\|_1 - N)) \\ &= 2R(u) (\mathbf{1}_{u>0} - N^{-1}\|\mathbf{1}_{u>0}\|_1) . \end{aligned}$$

Since u is of zero mean and $\|u\|_1 > 0$, then $0 \in [u_{min}, u_{max}]$. From Proposition 3.1, we thus have that $p \in \partial J(u) \Rightarrow p \in \partial J(v)$, for all $v = a\mathbf{1}_{u>0} + b$ with $a \geq 0$ and $b \in \mathbb{R}$. Taking $a = 1$ and $b = -N^{-1}\|\mathbf{1}_{u>0}\|_1$, we have

$$p = 2R(u)v \in \partial J(v).$$

□

The method in Lemma 3.1 for generating calibrable sets is illustrated in Fig 1.

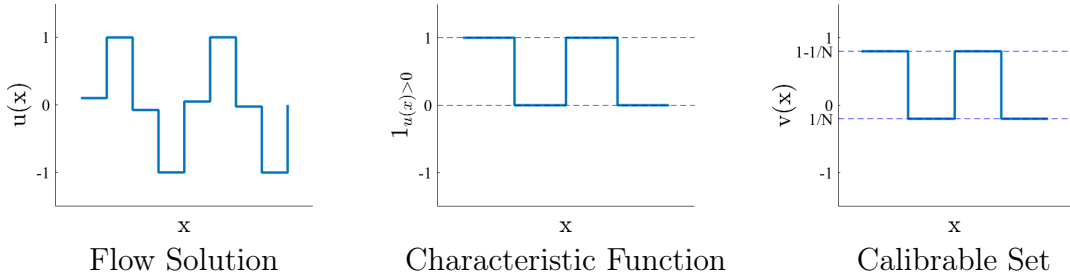


Figure 1: Generating calibrable sets of J using the solution of our flow 3.2.

4 Cheeger cut minimization

Given a graph $G = (V, W)$ where W is a similarity matrix with elements $(W)_{ij} = w_{ij}$, the Cheeger cut minimization problem is given by

$$\min_A CC(A, \bar{A}) = \frac{\sum_{i \in A, j \in \bar{A}} w_{ij}}{\min(|A|, |\bar{A}|)}, \quad s.t \quad A \sqcup \bar{A} = V \quad (4.1)$$

This problem is NP-hard. To approach this problem, the Cheeger cut is rewritten using the characteristic function of the set A , i.e $v = \mathbf{1}_A$ by

$$CC(A, \bar{A}) := \frac{\sum_{ij} w_{ij} |v_i - v_j|}{\|v - med(v)\|_1} \quad (4.2)$$

which motivates the continuous relaxation problem

$$\min_u \frac{\sum_{ij} w_{ij} |u_i - u_j|}{\|u - \text{med}(u)\|_1}, \quad \text{s.t. } u : V \rightarrow \mathbb{R} \text{ nonconstant} \quad (4.3)$$

This relaxation is the main focus in recent spectral clustering papers, e.g. (Hein & Bühler 2010), (Bresson et al. 2012).

In this work we offer a relaxed solution to 4.1 by minimizing an equivalent form of 4.3 given by

$$\min_{\text{med}(u)=0, \|u\|_2=1} R_1(u) := \frac{\sum_{ij} w_{ij} |u_i - u_j|}{\|u\|_1}, \quad \text{s.t. } u : V \rightarrow \mathbb{R} \quad (4.4)$$

We then use Lemma 4.3 to generate a characteristic eigenfunction of 1.4 that offers a solution to the exact Problem in 4.1. For solving 4.4 we present the following modified version of our main flow (1.7) and a second flow which performs a steepest descent algorithm with respect to the log of R_1 . This modification is done by the reduction of the median step, which as we will show, keeps monotonicity, convergence and the main theorem as the original main flow. We also note that both flows in Algorithm 2 are different only by a dynamic time rescaling, which is part of the update of $u_{k+1/2}$.

Algorithm 2: Modified flow for minimizing R_1

Data: positive time step and precision dt, ϵ .

Result: minimizer u of the Rayleigh quotient $R_1 = J_1/H_1$ with $\text{med}(u) = 0$.

Initialization: initial u_0 with $\text{med}(u) = 0$ and $\|u_0\|_2 = 1$.

(A) Modified main flow

(B) Logarithmic flow

while $\|R_1(u_{k+1}) - R_1(u_k)\| > \epsilon$ **do**

$$\left| \begin{array}{l} q_k \in \partial \bar{H}_1(u_k), \quad p_{k+1/2} \in \partial J_1(u_{k+1/2}) \\ u_{k+1/2} = u_k + dt (R_1(u_k)q_k - p_{k+1/2}) \\ \tilde{u}_{k+1/2} = u_{k+1/2} - \text{med}(u_{k+1/2}) \\ u_{k+1} = \frac{\tilde{u}_{k+1/2}}{\|\tilde{u}_{k+1/2}\|_2} \end{array} \right. \quad u_{k+1/2} = u_k + dt \left(\frac{q_k}{H(u_k)} - \frac{p_{k+1/2}}{J(u_k)} \right)$$

end while

In order to prevent u_k from converging to a constant solution, we use a different strategy than the one presented in the previous section. In a similar manner to (Hein & Bühler 2010), we consider a specific subset of subgradients of H_1 , defined as $\partial \bar{H}_1(u_k) = \{q \in \partial H_1(u_k) \mid \langle q, \mathbf{1} \rangle = 0\} \neq \emptyset$ for u_k of zero median. We now prove theoretical properties of the flow. These results are similar to the ones shown in Section 2, however they are adapted to the specific modifications introduced in Algorithm 2.A. The proofs for the properties of Algorithm 2.B are given in the appendix.

Proposition 4.1. *The flow in Algorithm 2.A admits the following properties:*

1. $\text{med}(u_k) = 0$ and $\|u_k\|_2 = 1$.
2. $0 < H(\tilde{u}_{k+1/2}) \leq L_H + dtL_H (R_1(u_0)L_H + L_J)$.

Proof. 1. This follows directly from the last two steps of the flow.

2. We have:

$$\tilde{u}_{k+1/2} = u_k + dt (R_1(u_k)q_k - p_{k+1/2}) - \text{med}(u_{k+1/2}).$$

Since $\langle \text{med}(u_{k+1/2}), q_{k+1} \rangle = 0$, the proof follows the one of Proposition 2.1. \square

Similarly to Section 2 the scheme is associated with the minimization of a convex functional, where here Equation 2.2 takes the form

$$u_{k+1/2} = \underset{u}{\text{argmin}} F_1(u) := \frac{1}{2dt} \|u - u_k\|_2^2 - R_1(u_k) \langle q_k, u \rangle + J_1(u). \quad (4.5)$$

Lemma 4.1 (Monotonicity). *The sequence u_k in Algorithm 2.A admits $R_1(u_{k+1}) \leq R_1(u_k)$.*

Proof. Recalling that $\langle q_k, \mathbf{1} \rangle = 0$ and so $\langle q_k, \tilde{u}_{k+1/2} \rangle = \langle q_k, u_{k+1/2} \rangle$, we get

$$\begin{aligned} F_1(u_{k+1/2}) &\leq F_1(u_k) \\ \frac{1}{2dt} \|u_{k+1/2} - u_k\|_2^2 - R_1(u_k) \langle q_k, u_{k+1/2} \rangle + J_1(u_{k+1/2}) &\leq -R_1(u_k) \langle q_k, u_k \rangle + J_1(u_k) \\ \frac{1}{2dt} \|u_{k+1/2} - u_k\|_2^2 - R_1(u_k) \langle q_k, \tilde{u}_{k+1/2} \rangle + J_1(\tilde{u}_{k+1/2}) &\leq 0 \\ \frac{1}{2dt} \|u_{k+1/2} - u_k\|_2^2 + J_1(\tilde{u}_{k+1/2}) &\leq R_1(u_k) \langle q_k, \tilde{u}_{k+1/2} \rangle \\ \frac{1}{2dt} \|u_{k+1/2} - u_k\|_2^2 + J_1(\tilde{u}_{k+1/2}) &\leq R_1(u_k) H_1(\tilde{u}_{k+1/2}) \\ \frac{1}{2dt H_1(\tilde{u}_{k+1/2})} \|u_{k+1/2} - u_k\|_2^2 + R_1(\tilde{u}_{k+1/2}) &\leq R_1(u_k) \\ \frac{1}{2dt H_1(\tilde{u}_{k+1/2})} \|u_{k+1/2} - u_k\|_2^2 + R_1(u_{k+1}) &\leq R_1(u_k). \end{aligned} \quad (4.6)$$

\square

The corresponding Proposition 4.1 and Lemma 4.1 for Algorithm 2.B are detailed and proved in B.

Lemma 4.2. *The sequence u_k in Algorithm 2 admits $\|u_{k+1} - u_k\|_2^2 \rightarrow 0$.*

Proof. Similarly to the proof of Lemma 2.3, we use (4.6) and point 2 in Proposition 4.1 to get a telescopic sum and conclude that $\|u_{k+1/2} - u_k\|_2^2 \rightarrow 0$. Then, from continuity of the median and the fact that $\text{med}(u_k) = 0$ we have that $\text{med}(u_{k+1/2}) \rightarrow 0$, so that $\|\tilde{u}_{k+1/2} - u_{k+1/2}\|_2 \rightarrow 0$, and by the triangle inequality $\|\tilde{u}_{k+1/2} - u_k\|_2^2 \rightarrow 0$. Finally, since $\|u_k\|_2 = 1$ we have that $\|\tilde{u}_{k+1/2}\|_2 \rightarrow 1$ and so $\|u_{k+1} - u_k\|_2^2 \rightarrow 0$. \square

We can now present the corresponding convergence theorem.

Theorem 4.1 (Convergence). *Let u_0 in X . Then there exist u, q and p in X such that up to a subsequence $u_k \rightarrow u, q_k \rightarrow q, p_{k+1/2} \rightarrow p, \|u\|_2 = 1$, and*

$$p = R_1(u)q, \quad q \in \partial H_1(u), \quad p \in \partial J_1(u). \quad (4.7)$$

Proof. Using Lemma 4.1,4.2 the proof is similar to the proof of Theorem 2.1. \square

In the following Lemma we show how we can generate characteristic eigenfunctions of the double-non linear eigenvalue problem (1.4) by thresholding the solution of our flow. These characteristic eigenfunctions represent our Cheeger cut candidates, of which the best candidate will be used for our final partition. We denoted these characteristic function by $v^>$ and $v^<$ and define them as follows.

Lemma 4.3. *Let u be the solution of Algorithm 2, then $v^> = \mathbf{1}_{u>\epsilon}$ with $\epsilon \in [0, u_{max}]$ and $v^< = \mathbf{1}_{u<\epsilon}$ with $\epsilon \in [u_{min}, 0]$ satisfy the double-non linear eigenvalue problem:*

$$p = \lambda q, \quad q \in \partial H_1(v), \quad p \in \partial J_1(v). \quad (4.8)$$

where the corresponding eigenvalue is given by $\lambda = R_1(u)$.

Proof. Recall that at convergence

$$0 \in p - R_1(u)\partial H_1(u), \quad p \in \partial J_1(u).$$

From Proposition 3.1 we have $p \in J_1(v^>)$. For $\epsilon \geq 0$ we have $v_i^> = 1 \rightarrow u_i > \epsilon \geq 0$ and so $\text{sign}(u_i) \subseteq \text{sign}(v_i^>)$, i.e $\partial H_1(u) \subseteq \partial H_1(v^>)$, Hence,

$$0 \in p - R_1(u)\partial H_1(v^>), \quad p \in \partial J_1(v^>),$$

so there exists $q \in \partial H_1(v^>)$ such that

$$p = R_1(u)q, \quad q \in \partial H_1(v^>), \quad p \in \partial J_1(v^>).$$

We can similarly show that $-v^< = -\mathbf{1}_{u<\epsilon}$ satisfies equation 4.8 and since p, q and R_1 are 0-homogeneous we conclude the proof for $v^<$. \square

Remark 4.1. *The characteristic eigenfunctions $v^>$ and $v^<$ has zero median.*

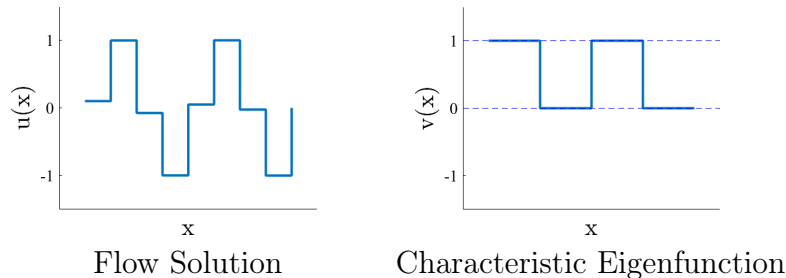


Figure 2: Generating characteristic eigenfunctions that represent a Cheeger cut.

4.1 Experiments

We used the settings in (Bresson et al. 2012) for both flows in Algorithm 2, with the corresponding dynamic time rescaling. We used (Chambolle & Pock 2011) for the minimization of F_1 step. We tested our flows on the known benchmark classification task of MNIST49 two digits partition, and compared the results with two leading methods for Cheeger cut minimization: steepest descent (SD) (Bresson et al. 2012) and inverse-power-method (IPM) (Hein & Bühler 2010). We evaluated the best time step for each of the algorithms, which resulted in $dt = 1, 1, 50$ for the SD algorithm, Algorithm 2.A and Algorithm 2.B, respectively. The IPM method does not include a time step parameter. Table 4.1 summarizes the averaged cut, error and speed over 10 experiments using different random initializations, where all algorithm have been given the same input. We observe that our two proposed algorithms outperform the IPM algorithm, and perform comparably with respect to the SD algorithm. Since the main difference between our two algorithms and SD algorithm is a dynamic time rescaling, we compared the performances of the algorithms over a range of time steps (where the IPM is invariant under this change). The results are shown in Figure 3. It can be seen that Algorithm 2.B is highly robust to time step changes. On the other hand, Algorithm 2.A and the SD algorithm are sensitive to larger time steps. This advantage of Algorithm 2.B also enables to accelerate speed by using larger time steps. Note that the SD Algorithm and Algorithm 2.A still outperform the IPM method in all experiments. This emphasizes the significance of using procedures based on time steps for such tasks. In future work, we intend to generalize our algorithms to a multi class settings and to use it for genome assembly and hashing tasks.

5 Continuous framework

In this section we analyze, in a time continuous setting, the properties of the flow for functions J satisfying Properties 3.1 and thus using the modification of \bar{q} presented

| | MNIST 49 | | |
|--------------------------|-----------|-----------|------------|
| Method: | Cut value | Error (%) | Time (sec) |
| IPM (Hein & Bühler 2010) | 0.24998 | 15.583 | 4.13 |
| SD (Bresson et al. 2012) | 0.11572 | 1.643 | 6.70 |
| Proposed flow 2.A | 0.11569 | 1.640 | 10.55 |
| Proposed flow 2.B | 0.11570 | 1.637 | 5.13 |

Table 1: MNIST49 comparison averaged over 10 experiments.

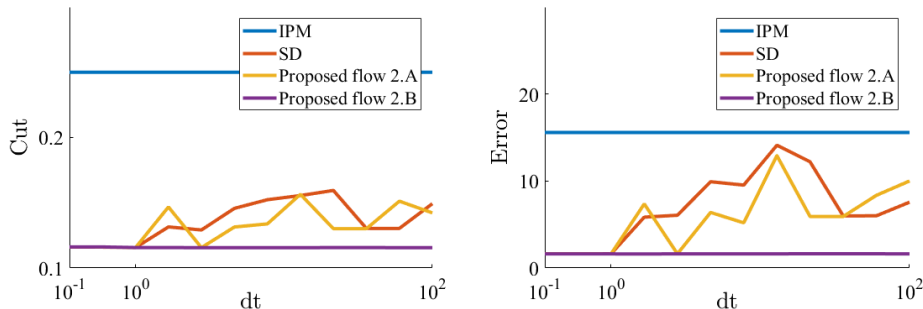


Figure 3: MNIST49 comparison using a range of time steps.

in section 3. In that case

$$u_t = R(u)(q - \bar{q}) - p, \quad q \in \partial H(u), p \in \partial J(u), \quad (5.1)$$

where $u(\cdot, 0) = u_0$, and $t \in (0, T)$ with $T > 0$.

The plan of the section is as follows: We first introduce some preliminary results on absolutely continuous functions in subsection 5.1, which is the proper mathematical setting to analyse the flow (5.1). We state several properties satisfied by solutions of the flow in subsection 5.2. We can then establish existence and uniqueness results for the solution of the flow (5.1) in subsection 5.3: we first state a general result in Theorem 5.1 with a generic assumption on the subgradient of H . We then give a particular instance of the result in the case when $H(u) = \|u\|_\alpha$ with $\alpha \geq 2$ in Corollary 1. Eventually, we resort to Moreau-Yosida regularization in subsection 5.4 to provide existence and uniqueness results in Corollary 2 without additional assumptions on H .

5.1 Preliminary results

In order to study the properties of the flow (5.1), we first give the following Lemma, which can be found in (Brézis 1973), Lemme 3.3 page 73 (see also Lemma 4.1 in (Apidopoulos, Aujol & Dossal 2017)). It allows us to use the “chain rule for differentiation” for the mapping $t \rightarrow J(u(t))$ and $t \rightarrow H(u(t))$.

Lemma 5.1. *Let $T > 0$ and F be a convex, lower semi-continuous, proper function and $v \in W^{1,2}((0, T); X)$. Let also $h \in L^2((0, T); X)$, such that $h \in \partial F(v(t))$ a.e. in $(0, T)$. Then the function $F \circ v : [0, T] \rightarrow \mathbb{R}$ is absolutely continuous in $[0, T]$ with*

$$\frac{d}{dt}(F(v(t))) = \langle z, v_t \rangle \quad \forall z \in \partial F(v(t)) \quad \text{a.e. in } (0, T). \quad (5.2)$$

We recall here the definition of an absolutely continuous function (e.g Example 1.13 in (Clarke 2013)).

Definition 3. *Let $[a, b]$ be an interval in $[0, +\infty)$. A function $G : [a, b] \rightarrow \mathbb{R}$ is said to be absolutely continuous if for every $\varepsilon > 0$, there exists $\delta > 0$ such that for every finite collection $\{[a_i, b_i]\}_{i \in I}$ of disjoint subintervals of $[a, b]$, we have*

$$\sum_{i \in I} (b_i - a_i) < \delta \quad \implies \quad \sum_{i \in I} |G(b_i) - G(a_i)| < \varepsilon. \quad (5.3)$$

Equivalently a function $G : [a, b] \rightarrow \mathbb{R}$ is absolutely continuous if there exists a function $v \in L^1(a, b)$, such that

$$G(t) = G(s) + \int_s^t v(\tau) d\tau \quad \forall a \leq s \leq t \leq b. \quad (5.4)$$

In that case we have that G is differentiable a.e. in (a, b) with $\dot{G}(t) = v(t)$ a.e. in (a, b) .

Remark 5.1. *From the definition of absolute continuity, e.g (5.4), it follows that an absolutely continuous function with non-positive derivative a.e. in (a, b) is non-increasing.*

A basic property of absolutely continuous function is that if F and G are absolutely continuous and bounded on $[a, b]$, then so is the product FG . If $0 < c < G < C$ and F bounded, then F/G is absolutely continuous. In view of Lemma 5.1, if u in $W^{1,2}((0, T); X)$, we get that $J(u(t))$ and $H(u(t))$ are absolutely continuous in $[0, T]$ with:

$$\frac{d}{dt}J(u(t)) = \langle p, u_t \rangle \quad \forall p \in \partial J(u(t)) \quad \text{a.e. in } (0, T), \quad (5.5)$$

and

$$\frac{d}{dt}H(u(t)) = \langle q, u_t \rangle \quad \forall q \in \partial H(u(t)) \quad \text{a.e. in } (0, T). \quad (5.6)$$

Moreover, if $0 < c \leq \frac{J}{H} \leq C < +\infty$ then J/H is absolutely continuous in $[0, T]$ with

$$\frac{d}{dt}R(u) = \frac{1}{H^2(u)} \langle H(u)p - J(u)q, u_t \rangle \quad \text{a.e. in } (0, T). \quad (5.7)$$

5.2 Properties of the flow

Proposition 5.1. *A solution u in $W^{1,2}((0, T), X)$ of the flow (5.1) admits*

1. *If u_0 is of zero mean then $u(t)$ is of zero mean.*
2. *$\frac{d}{dt}\|u(t)\|_2^2 = 0$, so that $H(u(t)) > 0$ from the equivalence of norms.*
3. *There exists c and $C > 0$ such that for all t in $[0, T)$ we have: $c \leq R(u(t)) \leq C$.*
4. *$\frac{d}{dt}R(u(t)) \leq 0$ for a.e. $t \in (0, T)$, and $t \rightarrow R(u(t))$ is absolutely continuous and non increasing.*
5. *If $u(t)$ converges to u then $p = R(u)(q - \bar{q})$ with $p \in \partial J(u)$ and $q \in \partial H(u)$.*

Proof. 1. This follows from the fact that $p \in \partial J(u)$, $q - \bar{q}$ and u_0 are of zero mean.

2. We have:

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \|u\|_2^2 &= \langle u, u_t \rangle \\ &= R(u) \langle u, q - \bar{q} \rangle - \langle u, p \rangle \\ &= J(u) - J(u) \\ &= 0. \end{aligned}$$

3. It is a straightforward consequence of points 1 and 2.

4. From point 3, using (5.7), we see that the mapping $u \mapsto R(u)$ is absolutely continuous on the trajectory, and that we have for almost every $t \in (0, T)$

$$\begin{aligned} \frac{d}{dt} R(u) &= \frac{1}{H^2(u)} \langle H(u)p - J(u)q, u_t \rangle \\ &= \frac{1}{H(u)} \langle p - R(u)(q - \bar{q}), u_t \rangle \\ &= -\frac{1}{H(u)} \|u_t\|_2^2 \\ &\leq 0. \end{aligned}$$

5. The relation is simply obtained with $u_t = 0$.

□

5.3 Existence and uniqueness

Our study follows the work of Brezis (Brézis 1973): in particular, it gives existence and uniqueness results for flows governed by a maximal monotone operator with possible Lipschitz perturbation. In order to show the existence of (5.1), we need to show that the mapping $R(u(t))(q - \bar{q})$ is locally Lipschitz continuous. To that end, we will decompose the problem into several simpler ones using the following result. The following lemma is standard and we state it without proof:

Lemma 5.2. *If $f : u \in \mathbb{R}^n \rightarrow \mathbb{R}_+$ is a bounded and locally Lipschitz continuous function on $u \in S$ and $u \in \mathbb{R}^n \mapsto q(u) \in \mathbb{R}^m$, $m \geq 1$ is a bounded, locally Lipschitz continuous mapping for $u \in S$, then $f q$ is a locally Lipschitz continuous mapping on $u \in S$.*

We can then look at the Rayleigh quotient $R = J/H$.

Proposition 5.2. *The function $R(u)$ is bounded and Lipschitz continuous on the trajectory of the flow.*

The proof is given in C.

Theorem 5.1 (Existence under Lipschitz assumption on q). *Let $\|u_0\|_2 = 1$ and $u \mapsto q(u) \in \partial H(u)$ be a locally Lipschitz continuous mapping. Then there exists a unique u in $W^{1,2}((0, T), X)$ solution of (5.1).*

Proof. We have

$$u_t = R(u(t))(q - \bar{q}) - p, \quad p \in \partial J(u(t)),$$

i.e.

$$u_t = A(u) - p, \quad p \in \partial J(u(t)),$$

where A is a Lipschitz function (thanks to the hypotheses of the Theorem 5.1 and the Proposition 5.2), and J a proper lower semi continuous convex function. Then the existence and uniqueness result is just a consequence of Theorem 3.17 page 105 of (Brézis 1973), based on the theory of maximal monotone operators. We recall below the proof of uniqueness, since it shows the role of the Lipschitz continuity of the mappings. Let u and v be two trajectories of the flow. We consider

$$\begin{aligned} & \frac{d}{dt} \frac{1}{2} \|u - v\|_2^2 \\ &= \langle u - v, u_t - v_t \rangle \\ &= \langle u - v, R(u)(q(u) - \bar{q}(u)) - R(v)(q(v) - \bar{q}(v)) \rangle - \langle u - v, p(u) - p(v) \rangle \\ &= \langle u - v, R(u)q(u) - R(v)q(v) \rangle - \langle u - v, p(u) - p(v) \rangle. \end{aligned} \tag{5.8}$$

and observe that

$$\begin{aligned} - \langle u - v, p(u) - p(v) \rangle &= -J(u) - J(v) + \langle u, p(v) \rangle + \langle v, p(u) \rangle \\ &\leq -J(u) - J(v) + J(u) + J(v) \\ &\leq 0. \end{aligned} \tag{5.9}$$

Since we assumed that q is a locally Lipschitz continuous mapping, from Lemma 5.2 and Proposition 5.2, there exists a constant D such that

$$\langle u - v, R(u)q(u) - R(v)q(v) \rangle \leq D\|u - v\|_2^2. \quad (5.10)$$

Combining relations (5.8), (5.9) and (5.10), we obtain

$$\frac{d}{dt} \frac{1}{2} \|u - v\|_2^2 \leq D\|u - v\|_2^2,$$

so that

$$\|u - v\|_2^2 \leq \|u_0 - v_0\|_2^2 \exp(D(t - t_0)),$$

which gives $u = v$ as long as $u_0 = v_0$. \square

The main issue is to show the local Lipschitz continuity of the mapping $u \rightarrow q(u) \in \partial H(u)$. To this end, we present the particular case where $H(u) = \|u\|_\alpha$ with $\alpha \geq 2$.

Proposition 5.3. *For $u \in X$, $H(u) = \|u\|_\alpha$ with $\alpha \in [2, +\infty)$, the mapping $q(u) \in \partial \|u\|_\alpha$ is locally Lipschitz continuous on the set $\|u\|_2 = 1$ and $\bar{u} = 0$.*

Proof. First notice that for $\|u\|_2 = 1$ and $\bar{u} = 0$, we necessarily have $|u_i| < 1$. From the equivalence of norms in finite dimension, there also exists a constant $c_\alpha > 0$ such that $c_\alpha \leq \|u\|_\alpha < 1$. Next, for $\alpha \geq 2$ we have

$$(q(u))_i = (\nabla(\|\cdot\|_\alpha)(u))_i = \frac{\|u\|_\alpha}{\|u\|_\alpha^\alpha} \text{sign}(u_i) |u_i|^{\alpha-1}. \quad (5.11)$$

It can then be shown that all the terms of (5.11) are bounded and locally Lipschitz on an open neighborhood V of u such that for all $v \in V$, $|v_i| < 1$. The proof can be concluded with Lemma 5.2. \square

As a consequence of Proposition 5.3 and Theorem 5.1, we get the following result.

Corollary 1. *Let $\|u_0\|_2 = 1$ and $H(u) = \|u\|_\alpha$ with $\alpha \in [2, +\infty)$. Then there exists a unique u in $W^{1,2}((0, T), X)$ solution of (5.1).*

5.4 Regularization of the flow

In order to deal with more general functions H , we now consider the Moreau-Yosida regularization of H defined, for $\epsilon > 0$ as

$$H^\epsilon(u) = \sup_q \langle u, q \rangle - H^*(q) - \frac{\epsilon}{2} \|q\|_2^2. \quad (5.12)$$

We denote as q^ϵ the unique q that reaches the previous supremum and remark that $q^\epsilon = \nabla H^\epsilon(u^\epsilon)$. Since $\|q^\epsilon\|_2 \leq L_H$, the function H^ϵ is Lipschitz continuous with

constant L_H . Observe that for $\epsilon \rightarrow 0$, $q^\epsilon \rightarrow q^*(u) = \operatorname{argmin}_{q \in \partial H(u)} \|q\|_2^2$, the subgradient of $H(u)$ of minimal norm. Since $\langle u, q \rangle - H^*(q) - \frac{\epsilon}{2} \|q\|_2^2 = 0$ for $q = 0 \in \partial H(u)$, we deduce from (5.12) that

$$0 \leq H^\epsilon(u) \leq \langle u, q^\epsilon \rangle - H^*(q^\epsilon) \leq \sup_q \langle u, q \rangle - H^*(q) = H(u). \quad (5.13)$$

5.4.1 Properties of regularized flows

For an initial point u_0^ϵ of zero mean, we consider the regularized flow

$$\begin{cases} u^\epsilon(0) &= u_0^\epsilon, \\ u_t^\epsilon &= \frac{J(u^\epsilon(t))}{H^\epsilon(u^\epsilon(t))} (q^\epsilon - \bar{q}^\epsilon) - p, \quad p \in \partial J(u^\epsilon(t)) \quad q^\epsilon = \nabla H^\epsilon(u^\epsilon(t)). \end{cases} \quad (5.14)$$

where $\bar{q}^\epsilon = \frac{1}{|\Omega|} \int_\Omega q^\epsilon \in \mathbb{R}$ so that $\langle \bar{q}^\epsilon, v \rangle = 0$ for all v of zero mean on Ω . We look for solutions of (5.14) that reside in $W^{1,2}((0, T), X)$. The trajectories of this regularized flow are similar to the ones of the general flow (5.1).

Proposition 5.4. *Assume that there exists a solution u^ϵ in $W^{1,2}((0, T), X)$ of the flow (5.14). Then the following properties hold*

1. $u^\epsilon(t)$ is of zero mean.
2. $\frac{d}{dt} \|u^\epsilon(t)\|_2^2 \geq 0$, so that $H^\epsilon(u^\epsilon(t)) > 0$ from the equivalence of norms.
3. There exists α and $\beta > 0$ such that for all t in $[0, T)$ we have $\alpha \leq \frac{J(u^\epsilon(t))}{H^\epsilon(u^\epsilon(t))} \leq \beta$.
4. $\frac{d}{dt} \frac{J(u^\epsilon(t))}{H^\epsilon(u^\epsilon(t))} \leq 0$ for almost every t in $(0, T)$, $t \mapsto \frac{J(u^\epsilon(t))}{H^\epsilon(u^\epsilon(t))}$ is absolutely continuous and non increasing.
5. If $u^\epsilon(t)$ converges to u_*^ϵ then $p = \frac{J(u_*^\epsilon)}{H^\epsilon(u_*^\epsilon)} (q_*^\epsilon - \bar{q}_*^\epsilon)$ with $p_* \in \partial J(u_*^\epsilon)$ and $q_*^\epsilon = \nabla H^\epsilon(u_*^\epsilon)$.

Proof. 1. First notice that if u_0^ϵ is of zero mean and as $\langle p^\epsilon, \mathbf{1} \rangle = 0$ (using Property 2(c) since $J^*(p^\epsilon) = 0$), we have that $u^\epsilon(t)$ is of zero mean, for all $t > 0$.

2. From (5.13), we have

$$\begin{aligned} \frac{d}{dt} \frac{1}{2} \|u^\epsilon\|_2^2 &= \langle u^\epsilon, u_t^\epsilon \rangle = \left\langle u^\epsilon, \frac{J(u^\epsilon)}{H^\epsilon(u^\epsilon)} (q^\epsilon - \bar{q}^\epsilon) - p \right\rangle \\ &= J(u^\epsilon) \left(\frac{\langle u^\epsilon, q^\epsilon \rangle}{H^\epsilon(u^\epsilon)} - 1 \right) \geq 0. \end{aligned}$$

3. From point 2, we know that if t in $(0, T]$, then $\|u^\epsilon(t)\|_2 \geq \|u^\epsilon(0)\|_2$. Hence if t in $[0, T]$, then $0 < \alpha \leq H^\epsilon(u^\epsilon)$. Moreover, from (5.13), we know that $H^\epsilon(u^\epsilon) \leq H(u^\epsilon)$. We thus deduce that the following inequality holds:

$$0 \leq \frac{J(u^\epsilon)}{H(u^\epsilon)} \leq \frac{J(u^\epsilon)}{H^\epsilon(u^\epsilon)} \leq \frac{J(u^\epsilon)}{\alpha} \quad (5.15)$$

Since J is continuous, we get that $\frac{J(u^\epsilon)}{\alpha}$ is bounded. Since J and H are absolutely one homogeneous, we see that $\frac{J(\lambda u^\epsilon)}{H(\lambda u^\epsilon)} = \frac{J(u^\epsilon)}{H(u^\epsilon)}$ for all $\lambda \neq 0$. Hence $\inf \frac{J(u^\epsilon)}{H(u^\epsilon)} \geq \min_{\{\|u^\epsilon\|_2=1\}} \frac{J(u^\epsilon)}{H(u^\epsilon)} > 0$.

4. Using (5.7), we know that the mapping $u \mapsto \frac{J(u^\epsilon)}{H^\epsilon(u^\epsilon)}$ is absolutely continuous along the trajectory, and that we have for almost every $t \in (0, T)$

$$\begin{aligned} \frac{d}{dt} \frac{J(u^\epsilon)}{H^\epsilon(u^\epsilon)} &= \frac{1}{(H^\epsilon(u^\epsilon))^2} \langle H^\epsilon(u^\epsilon)p - J(u^\epsilon)q^\epsilon, u_t^\epsilon \rangle \\ &= \frac{1}{H^\epsilon(u^\epsilon)} \left\langle p - \frac{J(u^\epsilon)}{H^\epsilon(u^\epsilon)}(q^\epsilon - \bar{q}^\epsilon), u_t^\epsilon \right\rangle \\ &= -\frac{1}{H^\epsilon(u^\epsilon)} \|u_t^\epsilon\|_2^2 \leq 0. \end{aligned}$$

5. The last point is a direct consequence of $u_t^\epsilon = 0$. □

5.4.2 Lipschitz continuity of regularized flows

We assume that the initial point of the trajectory satisfies $\|u_0^\epsilon\|_2 > 0$. From point 2 of Proposition 5.4, we know that the flow ensures that $\|u^\epsilon\|_2$ is non decreasing. Hence $\|u^\epsilon\|_2 > 0$.

Proposition 5.5. *The function $J(u^\epsilon)/H^\epsilon(u^\epsilon)$ is Lipschitz continuous on the trajectory of the regularized flow.*

Proof. Thanks to points (3) and (4) of Proposition 5.4, the flow (5.14) ensures that $J(u^\epsilon)/H^\epsilon(u^\epsilon)$ is decreasing and thus bounded by $J(u_0^\epsilon)/H^\epsilon(u_0^\epsilon)$. Notice also that the function H^ϵ is Lipschitz continuous with constant L_H , since the mapping $u^\epsilon \mapsto q^\epsilon = \nabla H^\epsilon(u^\epsilon)$ necessarily admits $q^\epsilon \in \partial H(0)$ so that $\|q^\epsilon\|_2 \leq L_H$. The arguments are then the same as the ones in the proof of Proposition 5.2. □

Proposition 5.6. *For $\epsilon > 0$, the mapping $u^\epsilon \mapsto q^\epsilon(u^\epsilon)$ is Lipschitz continuous with constant $1/\epsilon$.*

Proof. We first observe from (5.12) that

$$\begin{aligned} q^\epsilon(u^\epsilon) &= \operatorname{argmin}_q H^*(p) + \frac{\epsilon}{2} \|q\|_2^2 - \langle u^\epsilon, q \rangle \\ &= \operatorname{argmin}_q H^*(q) + \frac{\epsilon}{2} \|q - u^\epsilon/\epsilon\|_2^2. \end{aligned} \quad (5.16)$$

Recalling that $H^*(p) = \chi_{\partial H(0)}(q)$, this amounts to compute the orthogonal projection of u^ϵ/ϵ onto $\partial H(0)$. For all u^ϵ, v^ϵ , one thus has

$$\|q^\epsilon(u^\epsilon) - q^\epsilon(v^\epsilon)\|_2 = \|\operatorname{Proj}_{\partial H(0)}(u^\epsilon/\epsilon) - \operatorname{Proj}_{\partial H(0)}(v^\epsilon/\epsilon)\|_2 \leq \frac{1}{\epsilon} \|u^\epsilon - v^\epsilon\|_2. \quad (5.17)$$

□

The constant $1/\epsilon$ of Proposition 5.6 is the main issue to pass to the limit and obtain results on flows (5.1) involving general functions H , namely $H(u) = \|u\|_\alpha$, with $\alpha \in [1; 2]$.

To that end, it is indeed necessary to show that the mapping $\frac{J(u^\epsilon(t))}{H^\epsilon(u^\epsilon(t))} (q^\epsilon - \bar{q}^\epsilon) - p$ is locally Lipschitz continuous with a constant that does not explode with $\epsilon \rightarrow 0$. However, we can only show such a property for particular function such as $H(u) = \|u\|_2$, that has already been treated in the previous section without regularization.

Proposition 5.7. *For $H(u) = \|u\|_2$, $\|u_0^\epsilon\|_2 > 1$ and $\epsilon \leq 1$, the mapping $u^\epsilon \mapsto q^\epsilon$ is locally 1-Lipschitz continuous.*

Proof. For $H(u) = \|u\|_2$ then $\partial H(0) = \{q, \text{ s.t. } \|q\|_2 \leq 1\}$, in this case, for all $\|u^\epsilon\|_2 > L_H = 1$ and $\epsilon \leq 1$, we have $\operatorname{Proj}_{\partial H(0)}(u^\epsilon/\epsilon) = \operatorname{Proj}_{\partial H(0)}(u^\epsilon) = u^\epsilon/\|u^\epsilon\|_2$. As a consequence for all $\|u^\epsilon\|_2 > L_H$ and $\|v^\epsilon\|_2 > L_H$ in the neighborhood of u^ϵ we get

$$\|q^\epsilon(u^\epsilon) - q^\epsilon(v^\epsilon)\|_2 = \|\operatorname{Proj}_{\partial H(0)}(u^\epsilon) - \operatorname{Proj}_{\partial H(0)}(v^\epsilon)\|_2 \leq \|u^\epsilon - v^\epsilon\|_2. \quad (5.18)$$

□

We nevertheless have this general existence and uniqueness result which is a consequence of Theorem 5.1 and Propositions 5.5 and 5.6.

Corollary 2. *Let $\|u_0\|_2 = 1$. Then there exists a unique u^ϵ in $W^{1,2}((0, T), X)$ solution of (5.14).*

6 Conclusion

In this paper we addressed the general problem of minimizing Rayleigh quotients of two absolutely one-homogeneous functionals. A simple iterative algorithm is provided motivated by the theory of Brezis for flows with maximal monotone operators. We show that the iterative algorithm monotonically decreases the Rayleigh

quotient and prove convergence. The solution at convergence admits a nonlinear eigenvalue problem with respect to the subgradients of the two functionals. The time-continuous case is harder to analyze and requires additional regularity assumptions. Thus, for the continuous analysis we use Moreau-Yosida regularization. We further suggest an iterative scheme based on minimizing the log of the Rayleigh quotient. We show that both algorithms perform well for approximating the Cheeger cut problem.

Acknowledgments: This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 777826. The authors would like to thank Antonin Chambolle for many discussions related to this work. Jean-François Aujol is a member of Institut Universitaire de France (IUF). Guy Gilboa acknowledges support by the Israel Science Foundation (grant No. 718/15). This work was supported by the Technion Ollendorff Minerva Center. This study has also been carried out with financial support from the French State, managed by the French National Research Agency (ANR GOTMI) (ANR-16-CE33-0010-01). The authors would also like to thank the Isaac Newton Institute for Mathematical Sciences, Cambridge, for support and hospitality during the program "Variational methods and effective algorithms for imaging and vision" where work on this paper was undertaken. This work was supported by EPSRC grant no EP/K032208/1.

A Analysis of the main flow with \bar{q} modification

The results presented here are valid for functions J satisfying the properties 3.1.

For u_0 of zero mean and $\|u_0\|_2 = 1$, we consider the following flow scheme

$$\begin{cases} u_{k+1/2} &= u_k + dt \left(\frac{J(u_k)}{H(u_k)} (q_k - \bar{q}_k) - p_{k+1/2} \right), \quad q_k \in \partial H(u_k), \quad p_{k+1/2} \in \partial J(u_{k+1/2}), \\ u_{k+1} &= \frac{u_{k+1/2}}{\|u_{k+1/2}\|_2}, \end{cases} \quad (\text{A.1})$$

where \bar{q}_k is the mean value of q_k . We can thus observe that

$$u_{k+1/2} = \underset{u}{\operatorname{argmin}} F(u) := \frac{1}{2dt} \|u - u_k\|_2^2 - \langle R(u_k)(q_k - \bar{q}_k), u \rangle + J(u). \quad (\text{A.2})$$

Proposition A.1. *The flow (A.1) admits the following properties:*

1. $1 = \|u_k\|_2^2 \leq \langle u_{k+1/2}, u_k \rangle \leq \|u_{k+1/2}\|_2^2$
2. $\|u_{k+1} - u_k\|_2 \leq \|u_{k+1/2} - u_k\|_2$.
3. $0 < H(u_{k+1/2}) \leq C = 1 + dt L_H ((R(u_0)(1 + N^{-1}) L_H + L_J)$

Proof. 1. Relation (A.2) can be rewritten as

$$u_{k+1/2} = \operatorname{argmin}_u \frac{1}{2dt} \|u - (u_k + dt(R(u_k)(q_k - \bar{q}_k))\|_2^2 + J(u).$$

Since $J(u+c) = J(u)$ for every constant c , the mean value of the ℓ_2 data term should be conserved, so that

$$\langle u_{k+1/2}, \mathbf{1} \rangle = \langle u_k + dtR(u_k)(q_k - \bar{q}_k), \mathbf{1} \rangle = 0.$$

As u_{k+1} is a normalization of $u_{k+1/2}$, its norm is 1 and it is also of zero mean.

2. One has $u_{k+1/2} = u_k + dt(R(u_k)(q_k - \bar{q}_k) - p_{k+1/2})$. Taking the scalar product of $u_{k+1/2}$ with u_k gives

$$\begin{aligned} \langle u_{k+1/2}, u_k \rangle &= \|u_k\|_2^2 + dt(J(u_k)(1 - \langle \bar{q}_k, u_k \rangle) - \langle p_{k+1/2}, u_k \rangle) \\ &= \|u_k\|_2^2 + dt(J(u_k) - \langle p_{k+1/2}, u_k \rangle) \\ &\geq \|u_k\|_2^2. \end{aligned}$$

On the other hand, by cauchy-schwarz inequality

$$\langle u_{k+1/2}, u_k \rangle \leq \|u_{k+1/2}\|_2 \|u_k\|_2 \leq \|u_{k+1/2}\|_2^2 \quad (\text{A.3})$$

Observing that u_k and u_{k+1} are the orthogonal projections of u_k and $u_{k+1/2}$ of the convex ℓ_2 ball, we conclude that

$$\|u_{k+1} - u_k\|_2 \leq \|u_{k+1/2} - u_k\|_2. \quad (\text{A.4})$$

3. We have:

$$u_{k+1/2} = u_k + dt(R(u_k)(q_k - \bar{q}_k) - p_{k+1/2})$$

Hence:

$$\begin{aligned} H(u_{k+1/2}) &= \langle u_k, q_{k+1} \rangle + dt(R(u_k)(\langle q_k, q_{k+1} \rangle - \bar{q}_k \langle \mathbf{1}, q_{k+1} \rangle) - \langle p_{k+1/2}, q_{k+1} \rangle) \\ H(u_{k+1/2}) &\leq 1 + dt(R(u_k)(\langle q_k, q_{k+1} \rangle - N^{-1} \langle \mathbf{1}, q_k \rangle \langle \mathbf{1}, q_{k+1} \rangle) - \langle p_{k+1/2}, q_{k+1} \rangle) \end{aligned}$$

From 1.11 there exists $L_J, L_H < \infty$ such that $\|p\|_2 \leq L_J$ and $\|q\|_2 \leq L_H$, thus

$$\begin{aligned} H(u_{k+1/2}) &\leq 1 + dtL_H(R(u_k)(1 + N^{-1})L_H + L_J) \\ &\leq 1 + dtL_H((R(u_0)(1 + N^{-1})L_H + L_J). \end{aligned}$$

where the last inequality involves the following Lemma. □

Lemma A.1 (Monotonicity). *The sequence u_k of the flow (A.1) admits $R(u_{k+1}) \leq R(u_k)$.*

Proof. We follow Lemma 2.2 and remark that $u_{k+1/2}$ now minimizes $\frac{1}{2dt}\|u - u_k\|_2^2 - R(u) \langle q_k - \bar{q}_k, u \rangle + J(u)$. Recalling that $\langle u_k, \bar{q}_k \rangle = \langle u_{k+1/2}, \bar{q}_k \rangle = 0$, we get

$$\begin{aligned} \frac{1}{2dt}\|u_{k+1/2} - u_k\|_2^2 - R(u_k) \langle q_k, u_{k+1/2} \rangle + J(u_{k+1/2}) &\leq -R(u_k) \langle q_k, u_k \rangle + J(u_k) \\ \frac{1}{2dt}\|u_{k+1/2} - u_k\|_2^2 + J(u_{k+1/2}) &\leq R(u_k)H(u_{k+1/2}) \\ \frac{1}{2dtH(u_{k+1/2})}\|u_{k+1/2} - u_k\|_2^2 + R(u_{k+1}) &\leq R(u_k). \end{aligned} \tag{A.5}$$

□

Lemma A.2 (Convergence). *The sequence u_k of the flow (A.1) admits $\|u_{k+1} - u_k\|_2^2 \rightarrow 0$.*

Proof. From (A.5) and point 3 of Proposition 2.2, we have

$$\begin{aligned} \frac{1}{2dtH(u_{k+1/2})}\|u_{k+1/2} - u_k\|_2^2 + R(u_{k+1}) &\leq R(u_k) \\ \frac{1}{2dt}\|u_{k+1/2} - u_k\|_2^2 &\leq C(R(u_k) - R(u_{k+1})) \end{aligned}$$

Summing the previous relation from $k = 0$ to K we get:

$$\sum_{k=0}^K \|u_{k+1/2} - u_k\|_2^2 \leq 2dtC(R(u_0) - R(u_{K+1})).$$

which means that $\sum \|u_{k+1/2} - u_k\|_2^2$ converges, hence $\|u_{k+1/2} - u_k\|_2^2 \rightarrow 0$. Then, using inequality (A.4) we conclude that $\|u_{k+1} - u_k\|_2^2 \rightarrow 0$. □

Theorem A.1. *Let u_0 in X with $J(u_0) < +\infty$. Then there exist u , p and q in X such that up to a subsequence $u_k \rightarrow u$, $p_{k+1/2} \rightarrow p$, $q_k \rightarrow q$, $\|u\|_2 = 1$, and*

$$J(u)(q - \bar{q}) - p = 0, \quad q \in \partial H(u), \quad p \in \partial J(u), \tag{A.6}$$

with \bar{q} the mean of q .

Proof. The goal is to pass to the limit $k \rightarrow +\infty$ in (A.1) that we recall here

$$\begin{cases} u_{k+1/2} &= u_k + dt (R(u_k)(q_k - \bar{q}_k) - p_{k+1/2}), \quad q_k \in \partial H(u_k), \quad p_{k+1/2} \in \partial J(u_{k+1/2}), \\ u_{k+1} &= \frac{u_{k+1/2}}{\|u_{k+1/2}\|_2}. \end{cases} \tag{A.7}$$

From the previous results, we know that u_k , q_k and $p_{k+1/2}$ are bounded in X . Hence there exist u , q and p in X such that up to a subsequence $u_k \rightarrow u$, $q_k \rightarrow q$, $p_{k+1/2} \rightarrow$

p . Since H and J are continuous functions we also have $H(u_{k+1}) \rightarrow H(u)$ and $J(u_k) \rightarrow J(u)$. Then, using Proposition A.2, we see that we can pass to the limit in (A.1) to get

$$p = R(u)(q - \bar{q}), \quad q \in \partial H(u), \quad p \in \partial J(u) \quad (\text{A.8})$$

We have $q_k \in \partial H(u_k)$. This means that for all v , we have: $H(v) - H(u_k) \geq \langle q_k, v - u_k \rangle$. Passing to the limit, we get $H(v) - H(u) \geq \langle q, v - u \rangle$, i.e. $q \in \partial H(u)$. Similarly, we get that $p \in \partial J(u)$, which concludes the proof. \square

B Logarithmic flow corresponding proofs

Proposition B.1. *The log flow in Algorithm 2 admits the following properties:*

1. $\text{med}(u_k) = 0$ and $\|u_k\|_2 = 1$.
2. $0 < \frac{H_1(\tilde{u}_{k+1/2})}{J_1(u_k)} \leq C$

Proof. 1. This follows directly from the last two steps of the flow.

2. We have:

$$\tilde{u}_{k+1/2} = u_k + dt \left(\frac{q_k}{H_1(u_k)} - \frac{p_{k+1/2}}{J_1(u_k)} \right) - \text{med}(u_{k+1/2})$$

Hence:

$$\begin{aligned} H_1(\tilde{u}_{k+1/2}) &= \langle u_k, q_{k+1} \rangle + dt \left(\frac{\langle q_k, q_{k+1} \rangle}{H_1(u_k)} - \frac{\langle p_{k+1/2}, q_{k+1} \rangle}{J_1(u_k)} \right) \\ H_1(\tilde{u}_{k+1/2}) &= 1 + \frac{dt}{J_1(u_k)} (R_1(u_k) \langle q_k, q_{k+1} \rangle - \langle p_{k+1/2}, q_{k+1} \rangle) \\ \frac{H_1(\tilde{u}_{k+1/2})}{J_1(u_k)} &\leq \frac{1}{J_1(u_k)} + \frac{dt}{J_1^2(u_k)} (R_1(u_k) \langle q_k, q_{k+1} \rangle - \langle p_{k+1/2}, q_{k+1} \rangle) \end{aligned}$$

From 1.11 there exists $L_J, L_H < \infty$ such that $\|p\|_2 \leq L_J$ and $\|q\|_2 \leq L_H$. Furthermore, since $\text{med}(u_k) = 0$ and $\|u_k\|_2 = 1$ there exist $c > 0$ such that $J_1(u_k) \geq c$, thus

$$\begin{aligned} \frac{H_1(\tilde{u}_{k+1/2})}{J_1(u_k)} &\leq c^{-1} + c^{-2} dt L_H (R_1(u_k) L_H + L_J) \\ &\leq c^{-1} + c^{-2} dt L_H (R_1(u_0) L_H + L_J). \end{aligned}$$

where the last inequality involves the following Lemma. \square

Lemma B.1 (Monotonicity). *The sequence u_k in Algorithm 2.B admits $R(u_{k+1}) \leq R(u_k)$.*

Proof. We follow Lemma 2.2 and remark that $u_{k+1/2}$ now minimizes $\frac{1}{2dt}\|u - u_k\|_2^2 - \frac{\langle q_k, u \rangle}{H_1(u_k)} + \frac{J_1(u)}{J_1(u_k)}$. Recalling that $\langle q_k, \mathbf{1} \rangle = 0$ and so $\langle q_k, \tilde{u}_{k+1/2} \rangle = \langle q_k, u_{k+1/2} \rangle$, we get

$$\begin{aligned}
\frac{1}{2dt}\|u_{k+1/2} - u_k\|_2^2 - \frac{\langle q_k, u_{k+1/2} \rangle}{H_1(u_k)} + \frac{J_1(u_{k+1/2})}{J_1(u_k)} &\leq -\frac{\langle q_k, u_k \rangle}{H_1(u_k)} + \frac{J_1(u_k)}{J_1(u_k)} \\
\frac{1}{2dt}\|u_{k+1/2} - u_k\|_2^2 + \frac{J_1(\tilde{u}_{k+1/2})}{J_1(u_k)} &\leq \frac{\langle q_k, \tilde{u}_{k+1/2} \rangle}{H_1(u_k)} \\
\frac{1}{2dt}\|u_{k+1/2} - u_k\|_2^2 + \frac{J_1(\tilde{u}_{k+1/2})}{J_1(u_k)} &\leq \frac{\langle \tilde{q}_{k+1/2}, \tilde{u}_{k+1/2} \rangle}{H_1(u_k)} \\
\frac{1}{2dt}\|u_{k+1/2} - u_k\|_2^2 + \frac{J_1(\tilde{u}_{k+1/2})}{J_1(u_k)} &\leq \frac{H_1(\tilde{u}_{k+1/2})}{H_1(u_k)} \\
\frac{J_1(u_k)}{2dtH_1(\tilde{u}_{k+1/2})}\|u_{k+1/2} - u_k\|_2^2 + R_1(\tilde{u}_{k+1/2}) &\leq R_1(u_k) \\
\frac{J_1(u_k)}{2dtH_1(\tilde{u}_{k+1/2})}\|u_{k+1/2} - u_k\|_2^2 + R_1(u_{k+1}) &\leq R_1(u_k).
\end{aligned} \tag{B.1}$$

□

C Proof of Proposition 5.2

Proof. We denote $R(u) = J(u)/H(u)$. Thanks to point 3 of Proposition 5.1, the flow (5.1) ensures that $R(u) < +\infty$. We recall that L_J (resp. L_H) is the upper bound of the norm of the elements $p \in \partial J(0)$ (resp. $q \in \partial H(0)$). Let u and v be two trajectories of the flow, we have

$$\begin{aligned}
\left\| \frac{J(u)}{H(u)} - \frac{J(v)}{H(v)} \right\|_2 &\leq \left\| \frac{J(u) - J(v)}{H(u)} \right\|_2 + \left\| \left(\frac{1}{H(u)} - \frac{1}{H(v)} \right) J(v) \right\|_2 \\
&\leq \frac{1}{H(u)} \|J(u) - J(v)\|_2 + \frac{J(v)}{H(u)H(v)} \|H(u) - H(v)\|_2 \\
&\leq \frac{1}{H(u)} (J(u) - v) + R(v_0)L_H \|u - v\|_2 \\
&\leq \frac{R(u)}{J(u)} (L_J + R(v_0)L_H) \|u - v\|_2 \\
&\leq \frac{R(u_0)\kappa_J}{\|u\|_2} (L_J + R(v_0)L_H) \|u - v\|_2 \\
&\leq \frac{R(u_0)\kappa_J}{\|u_0\|_2} (L_J + R(v_0)L_H) \|u - v\|_2,
\end{aligned} \tag{C.1}$$

since, from Proposition 5.1 and relation (1.13), we know that there exists $0 < \kappa_J < +\infty$ such that $1 = \|u_0\|_2 = \|u\|_2 \leq \kappa_J J(u)$. \square

References

- Andreu, F., Ballester, C., Caselles, V. & Mazón, J. M. (2001). Minimizing total variation flow, *Differential and Integral Equations* **14**(3): 321–360.
- Apidopoulos, V., Aujol, J.-F. & Dossal, C. (2017). On a second order differential inclusion modeling the FISTA algorithm, *SIAM Journal on Optimization* (in press).
- Aujol, J., Gilboa, G. & Papadakis, N. (2018). Theoretical analysis of flows estimating eigenfunctions of one-homogeneous functionals, *SIAM Journal on Imaging Sciences* **11**(2): 1416–1440.
- Bacák, M., Bergmann, R., Steidl, G. & Weinmann, A. (2016). A second order non-smooth variational model for restoring manifold-valued images, *SIAM Journal on Scientific Computing* **38**(1): A567–A597.
- Bellettini, G., Caselles, V. & Novaga, M. (2002). The total variation flow in R^N , *Journal of Differential Equations* **184**(2): 475–525.
- Bresson, X., Laurent, T., Uminsky, D. & Brecht, J. V. (2012). Convergence and energy landscape for Cheeger Cut clustering, *Advances in Neural Information Processing Systems 25*, (NIPS '12), pp. 1385–1393.
- Brézis, H. (1973). *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert*, North Holland.
- Bühler, T. & Hein, M. (2009). Spectral clustering based on the graph p-laplacian, *International Conference on Machine Learning*, (ICML '09), pp. 81–88.
- Burger, M., Gilboa, G., Moeller, M., Eckardt, L. & Cremers, D. (2016). Spectral decompositions using one-homogeneous functionals, *SIAM Journal on Imaging Sciences* **9**(3): 1374–1408.
- Burger, M., Gilboa, G., Möller, M., Eckardt, L. & Cremers, D. (2016). Spectral decompositions using one-homogeneous functionals, *SIAM Journal on Imaging Sciences* **9**(3): 1374—1408.
- Chambolle, A. & Pock, T. (2011). A first-order primal-dual algorithm for convex problems with applications to imaging, *J. Math. Imaging Vis.* **40**(1): 120–145.
URL: <http://dx.doi.org/10.1007/s10851-010-0251-1>

- Clarke, F. (2013). *Functional analysis, calculus of variations and optimal control*, Vol. 264, Springer Science & Business Media.
- Giaquinta, M., Modica, G. & Souček, J. (1989). Cartesian currents and variational problems for mappings into spheres, *Annali della Scuola Normale Superiore di Pisa-Classe di Scienze* **16**(3): 393–485.
- Güttel, S. & Tisseur, F. (2017). The nonlinear eigenvalue problem, *Acta Numerica* **26**: 1–94.
- Hein, M. & Bühler, T. (2010). An inverse power method for nonlinear eigenproblems with applications in 1-spectral clustering and sparse pca, *Advances in Neural Information Processing Systems*, (NIPS '10), pp. 847–855.
- Lellmann, J., Strelakovsky, E., Koetter, S. & Cremers, D. (2013). Total variation regularization for functions with values in a manifold, *IEEE International Conference on Computer Vision*, (ICCV '13), pp. 2944–2951.
- Meyer, Y. (2001). Oscillating patterns in image processing and in some nonlinear evolution equations. The 15th Dean Jacqueline B. Lewis Memorial Lectures.
- Nossek, R. Z. & Gilboa, G. (2018). Flows generating nonlinear eigenfunctions, *Journal of Scientific Computing* **75**(2): 859–888.
- Rivière, T. (1995). Everywhere discontinuous harmonic maps into spheres, *Acta Mathematica* **175**(2): 197–226.
- Schwetlick, H. & Schreiber, K. (2012). Nonlinear rayleigh functionals, *Linear Algebra and Its Applications* **436**: 3991–4016.
- Sun, J., Qu, Q. & Wright, J. (2017). Complete dictionary recovery over the sphere i: Overview and the geometric picture, *IEEE Transactions on Information Theory* **63**(2): 853–884.