Convergence and Energy Landscape for Cheeger Cut Clustering

Xavier Bresson
City University of Hong Kong
Hong Kong
xbresson@cityu.edu.hk

Thomas Laurent
University of California, Riversize
Riverside, CA 92521
laurent@math.ucr.edu

David Uminsky
University of San Francisco
San Francisco, CA 94117
duminsky@usfca.edu

James H. von Brecht
University of California, Los Angeles
Los Angeles, CA 90095
jub@math.ucla.edu

Abstract

This paper provides both theoretical and algorithmic results for the ℓ_1-relaxation of the Cheeger cut problem. The ℓ_2-relaxation, known as spectral clustering, only loosely relates to the Cheeger cut; however, it is convex and leads to a simple optimization problem. The ℓ_1-relaxation, in contrast, is non-convex but is provably equivalent to the original problem. The ℓ_1-relaxation therefore trades convexity for exactness, yielding improved clustering results at the cost of a more challenging optimization. The first challenge is understanding convergence of algorithms. This paper provides the first complete proof of convergence for algorithms that minimize the ℓ_1-relaxation. The second challenge entails comprehending the ℓ_1-energy landscape, i.e. the set of possible points to which an algorithm might converge. We show that ℓ_1-algorithms can get trapped in local minima that are not globally optimal and we provide a classification theorem to interpret these local minima. This classification gives meaning to these suboptimal solutions and helps to explain, in terms of graph structure, when the ℓ_1-relaxation provides the solution of the original Cheeger cut problem.

1 Introduction

Partitioning data points into sensible groups is a fundamental problem in machine learning. Given a set of data points $V = \{x_1, \cdots, x_n\}$ and similarity weights $\{w_{i,j}\}_{1 \leq i,j \leq n}$, we consider the balance Cheeger cut problem [4]:

$$\text{Minimize} \quad \mathcal{C}(S) = \frac{\sum_{x_i \in S} \sum_{x_j \in S^c} w_{i,j}}{\min(|S|, |S^c|)} \qquad \text{over all subsets } S \subsetneq V. \tag{1}$$

Here $|S|$ denotes the number of data points in S and S^c is the complementary set of S in V. While this problem is NP-hard, it has the following *exact continuous* ℓ_1-relaxation:

$$\text{Minimize} \quad E(f) = \frac{\frac{1}{2}\sum_{i,j} w_{i,j}|f_i - f_j|}{\sum_i |f_i - \text{med}(f)|} \qquad \text{over all non-constant functions } f : V \to \mathbb{R}. \tag{2}$$

Here $\text{med}(f)$ denotes the median of $f \in \mathbb{R}^n$ and $f_i \equiv f(x_i)$. Recently, various algorithms have been proposed [13, 6, 7, 1, 10, 5] to minimize ℓ_1-relaxations of the Cheeger cut (1) and of other related problems. Typically these ℓ_1-algorithms provide excellent unsupervised clustering results

and improve upon the standard ℓ_2 (spectral clustering) method [11, 14] in terms of both Cheeger energy and classification error. However, complete theoretical guarantees of convergence for such algorithms do not exist. This paper provides the first proofs of convergence for ℓ_1-algorithms that attempt to minimize (2).

In this work we consider two algorithms for minimizing (2). We present a new steepest descent (SD) algorithm and also consider a slight modification of the inverse power method (IPM) from [6]. We provide convergence results for both algorithms and also analyze the energy landscape. Specifically, we give a complete classification of local minima. This understanding of the energy landscape provides intuition for when and how the algorithms get trapped in local minima. Our numerical experiments show that the two algorithms perform equally well with respect to the quality of the achieved cut. Both algorithms produce state of the art unsupervised clustering results. Finally, we remark that the SD algorithm has a better theoretical guarantee of convergence. This arises from the fact that the distance between two successive iterates necessarily converges to zero. In contrast, we cannot guarantee this holds for the IPM without further assumptions on the energy landscape. The simpler mathematical structure of the SD algorithm also provides better control of the energy descent.

Both algorithms take the form of a fixed point iteration $f^{k+1} \in \mathcal{A}(f^k)$, where $f \in \mathcal{A}(f)$ implies that f is a critical point. To prove convergence towards a fix point typically requires three key ingredients: the first is monotonicity of $\mathcal{A}$, that is $E(z) \leq E(f)$ for all $z \in \mathcal{A}(f)$; the second is some estimate that guarantees the successive iterates remain in a compact domain on which E is continuous; lastly, some type of continuity of the set-valued map $\mathcal{A}$ is required. For set valued maps, closedness provides the correct notion of continuity [8]. Monotonicity of the IPM algorithm was proven in [6]. This property alone is not enough to obtain convergence, and the closedness property proves the most challenging ingredient to establish for the algorithms we consider. Section 2 elucidates the form these properties take for the SD and IPM algorithms. In Section 3 we show that that if the iterates of either algorithm approach a neighborhood of a strict local minimum then both algorithms will converge to this minimum. We refer to this property as local convergence. When the energy is non-degenerate, section 4 extends this local convergence to global convergence toward critical points for the SD algorithm by using the additional structure afforded by the gradient flow. In Section 5 we develop an understanding of the energy landscape of the continuous relaxation problem. For non-convex problems an understanding of local minima is crucial. We therefore provide a complete classification of the local minima of (2) in terms of the combinatorial local minima of (1) by means of an explicit formula. As a consequence of this formula, the problem of finding local minima of the combinatorial problem is equivalent to finding local minima of the continuous relaxation. The last section is devoted to numerical experiments.

We now present the SD algorithm. Rewrite the Cheeger functional (2) as $E(f) = T(f)/B(f)$, where the numerator $T(f)$ is the total variation term and the denominator $B(f)$ is the balance term. If T and B were differentiable, a mixed explicit-implicit gradient flow of the energy would take the form $(f^{k+1} - f^k)/\tau^k = -(\nabla T(f^{k+1}) - E(f^k)\nabla B(f^k))/(B(f^k))$, where $\{\tau^k\}$ denotes a sequence of time steps. As T and B are not differentiable, particularly at the binary solutions of paramount interest, we must consider instead their subgradients

$$\partial T(f) := \{v \in \mathbb{R}^n : T(g) - T(f) \geq \langle v, g - f\rangle \ \forall g \in \mathbb{R}^n\}, \tag{3}$$

$$\partial_0 B(f) := \{v \in \mathbb{R}^n : B(g) - B(f) \geq \langle v, g - f\rangle \ \forall g \in \mathbb{R}^n \text{ and } \langle \mathbf{1}, v\rangle = 0\}. \tag{4}$$

Here $\mathbf{1} \in \mathbb{R}^n$ denotes the constant vector of ones. Also note that if f has zero median then $B(f) = ||f||_1$ and $\partial_0 B(f) = \{v \in \text{ sign}(f), \text{s.t. mean}(v) = 0\}$. After an appropriate choice of time steps we arrive to the SD Algorithm summarized in table 1(on left), i.e. a non-smooth variation of steepest descent. A key property of the the SD algorithm's iterates is that $\|f^{k+1} - f^k\|_2 \to 0$. This property allows us to conclude global convergence of the SD algorithm in cases where we can not conclude convergence for the IPM algorithm. We also summarize the IPM algorithm from [6] in Table 1 (on right). Compared to the original algorithm from [6], we have added the extra step to project onto the sphere $\mathcal{S}^{n-1}$, that is $f^{k+1} = h^k/||h^k||_2$. While we do not think that this extra step is essential, it simplifies the proof of convergence.

The successive iterates of both algorithms belong to the space

$$\mathcal{S}_0^{n-1} := \{f \in \mathbb{R}^n : ||f||_2 = 1 \quad \text{and} \quad \text{med}(f) = 0\}. \tag{5}$$

Table 1: $\mathcal{A}_{\mathrm{SD}}$: SD Algorithm. $\mathcal{A}_{\mathrm{IPM}}$: Modifed IPM Algorithm [6].

f^0 nonzero function with $\mathrm{med}(f) = 0$.	f^0 nonzero function with $\mathrm{med}(f) = 0$.
c positive constant.	**while** $E(f^k) - E(f^{k+1}) \geq \mathrm{TOL}$ **do**
while $E(f^k) - E(f^{k+1}) \geq \mathrm{TOL}$ **do**	$v^k \in \partial_0 B(f^k)$
$v^k \in \partial_0 B(f^k)$	$D^k = \min_{\|\|u\|\|_2 \leq 1} T(u) - E(f^k)\langle u, v^k\rangle$
$g^k = f^k + c\, v^k$	$g^k = \underset{\|\|u\|\|_2 \leq 1}{\arg\min}\, T(u) - E(f^k)\langle u, v^k\rangle$ if $D^k < 0$
$\hat{h}^k = \underset{u \in \mathbb{R}^n}{\arg\min}\, T(u) + \frac{E(f^k)}{2c}\|\|u - g^k\|\|_2^2$	$g^k = f^k$ if $D^k = 0$
$h^k = \hat{h}^k - \mathrm{med}(\hat{h}^k)\mathbf{1}$	$h^k = g^k - \mathrm{med}(g^k)\mathbf{1}$
$f^{k+1} = \frac{h^k}{\|h^k\|_2}$	$f^{k+1} = \frac{h^k}{\|\|h^k\|\|_2}$
end while	**end while**

As the successive iterates have zero median, $\partial_0 B(f^k)$ is never empty. For example, we can take $v^k \in \mathbb{R}^n$ so that $v^k(x_i) = 1$ if $f(x_i) > 0$, $v^k(x_i) = -1$ if $f(x_i) < 0$ and $v^k(x_i) = (n^- - n^+)/(n_0)$ if $f(x_i) = 0$ where n^+, n^- and n^0 denote the cardinalities of the sets $\{x_i : f(x_i) > 0\}$, $\{x_i : f(x_i) > 0\}$ and $\{x_i : f(x_i) = 0\}$, respectively. Other possible choices also exist, so that v^k is not uniquely defined. This idea, i.e. choosing an element from the subdifferential with mean zero, was introduced in [6] and proves indispensable when dealing with median zero functions. As v^k is not uniquely defined in either algorithm, we must introduce the concepts of a *set-valued map* and a *closed map*, which is the proper notion of continuity in this context:

Definiton 1 (Set-valued Map, Closed Maps)**.** *Let X and Y be two subsets of $\mathbb{R}^n$. If for each $x \in X$ there is a corresponding set $F(x) \subset Y$ then F is called a **set-valued map** from X to Y . We denote this by $F : X \rightrightarrows Y$. The graph of F, denoted Graph(F) is defined by*

$$\mathit{Graph}(F) = \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^n : x \in X,\ y \in F(x)\}.$$

*A set-valued map F is called **closed** if $\mathit{Graph}(F)$ is a closed subset of $\mathbb{R}^n \times \mathbb{R}^n$. In other words, if $y^k \in F(x^k)$, $x^k \to x^*$ and $y^k \to y^*$ then $x^* \in F(y^*)$.*

With these notations in hand we can write $f^{k+1} \in \mathcal{A}_{\mathrm{SD}}(f^k)$ (SD algorithm) and $f^{k+1} \in \mathcal{A}_{\mathrm{IPM}}(f^k)$ (IPM algorithm) where $\mathcal{A}_{\mathrm{SD}}, \mathcal{A}_{\mathrm{IPM}} : \mathcal{S}_0^{n-1} \rightrightarrows \mathcal{S}_0^{n-1}$ are the appropriate set-valued maps. The notion of a closed map proves useful when analyzing the step $\hat{h}^k \in \mathcal{H}(f^k)$ in the SD algorithm. Particularly,

Lemma 1 (Closedness of $\mathcal{H}(f)$)**.** *The set-valued map $\mathcal{H} : \mathcal{S}_0^{n-1} \rightrightarrows \mathbb{R}^n$*

$$\mathcal{H}(f) := \underset{u}{\arg\min} \left\{ T(u) + \frac{E(f)}{2c}\|\|u - (f + c\partial_0 B(f))\|\|_2^2 \right\}$$

is closed.

Currently, we can only show that lemma 1 holds at strict local minima for the analogous step, g^k, of the IPM algorithm. That lemma 1 holds without this further restriction on $f \in \mathcal{S}_0^{n-1}$ will allow us to demonstrate stronger global convergence results for the SD algorithm. We pause briefly to state closedness of the set-valued map $\partial_0 B(f) : \mathcal{S}_0^{n-1} \rightrightarrows \mathbb{R}^n$, as we need this result in many of the proofs that follow.

Lemma 2 (Closedness of $\partial_0 B(f)$)**.** *The set-valued map $\partial_0 B : \mathcal{S}_0^{n-1} \rightrightarrows \mathbb{R}^n$*

$$\partial_0 B(f) := \{v \in \mathbb{R}^n : B(g) - B(f) \geq \langle v, g - f\rangle\ \forall g \in \mathbb{R}^n \text{ and } \langle \mathbf{1}, v\rangle = 0\}$$

is closed.

Proof. See appendix A. □

2 Properties of $\mathcal{A}_{\mathrm{SD}}$ and $\mathcal{A}_{\mathrm{IPM}}$

This section establishes the required properties of the of the set-valued maps $\mathcal{A}_{\mathrm{SD}}$ and $\mathcal{A}_{\mathrm{IPM}}$ mentioned in the introduction. In section 2.1 we first elucidate the monotonicity and compactness of

$\mathcal{A}_{\mathrm{SD}}$ and $\mathcal{A}_{\mathrm{IPM}}$. Section 2.2 demonstrates that a local notion of closedness holds for each algorithm. This form of closedness suffices to show *local* convergence toward isolated local minima (c.f. Section 3). In particular, this more difficult and technical section is necessary as monotonicity alone does not guarantee this type of convergence.

2.1 Monotonicity and Compactness

We provide the monotonicity and compactness results for each algorithm in turn. Lemmas 3 and 4 establish monotonicity and compactness for $\mathcal{A}_{\mathrm{SD}}$ while Lemmas 5 and 6 establish monotonicity and compactness for $\mathcal{A}_{\mathrm{IPM}}$.

Lemma 3 (Monotonicity of $\mathcal{A}_{\mathrm{SD}}$)**.** *Let $f \in \mathcal{S}_0^{n-1}$ and define $v, g, \hat{h}$ and h according to the SD algorithm. Then neither $\hat{h}$ nor h is a constant vector. Moreover, the energy inequality*

$$E(f) \geq E(h) + \frac{E(f)}{B(h)} \frac{\|\hat{h} - f\|_2^2}{c} \tag{6}$$

holds. As a consequence, if $z \in \mathcal{A}_{\mathrm{SD}}(f)$ then $E(z) = E(h) < E(f)$ unless $z = f$.

Proof. The definition of $\hat{h}$ implies that $E(f)\left(\frac{\hat{h}-g}{c}\right) \in -\partial T(\hat{h})$. The definition of ∂T, the invariance of T under addition of a constant and the fact that $\langle v, \mathbf{1}\rangle = 0$ combine to imply

$$T(f) \geq T(\hat{h}) + \frac{E(f)}{c}\langle g - \hat{h}, f - \hat{h}\rangle = T(h) + \frac{E(f)}{c}||f - \hat{h}||_2^2 - E(f)\langle v, \hat{h} - f\rangle \tag{7}$$

$$= T(h) + \frac{E(f)}{c}||f - \hat{h}||_2^2 - E(f)\langle v, h - f\rangle. \tag{8}$$

As also $v \in \partial_0 B(f)$ we have $E(f)B(h) \geq E(f)B(f) + E(f)\langle v, h - f\rangle$. Adding these two last inequalities yields

$$T(f) + E(f)B(h) \geq T(h) + E(f)B(f) + \frac{E(f)}{c}\|\hat{h} - f\|_2^2.$$

In other words,

$$E(f)B(h) \geq T(h) + \frac{E(f)}{c}\|\hat{h} - f\|_2^2.$$

Note that $E(f) > 0$ as $f \in \mathcal{S}_0^{n-1}$. Therefore, if h were constant, then $B(h) = 0$ and $\hat{h} = h = f$. This is a contradiction since $f \in \mathcal{S}_0^{n-1}$ and is thus not constant. Consequently $B(h) > 0$, so we may divide in the last expression to obtain (6). The last statement then follows as E is invariant under scalings. □

Lemma 4 (Compactness of $\mathcal{A}_{\mathrm{SD}}$)**.** *Let $f^0 \in \mathcal{S}_0^{n-1}$ and define a sequence of iterates $(g^k, \hat{h}^k, h^k, f^{k+1})$ according to the SD algorithm. Then for any such sequence*

$$\|\hat{h}^k\|_2 \leq \|g^k\|_2, \quad 1 \leq ||g^k||_2 \leq 1 + c\sqrt{n} \quad \textit{and} \quad 0 < ||h^k||_2 \leq (1 + \sqrt{n})||\hat{h}^k||_2. \tag{9}$$

Moreover, we have

$$||\hat{h}^k - f^k||_2 \to 0, \qquad \mathrm{med}(\hat{h}^k) \to 0, \qquad \|f^k - f^{k+1}\|_2 \to 0. \tag{10}$$

Therefore $\mathcal{S}_0^{n-1}$ attracts the sequences $\{\hat{h}^k\}$ and $\{h^k\}$.

Proof. To prove that $\|\hat{h}\|_2 \leq \|g\|_2$, note

$$\hat{h} = \mathrm{prox}_\Phi(g) := \arg\min_u \left\{ \Phi(u) + \frac{\|u - g\|_2^2}{2} \right\} \quad \text{where} \quad \Phi(u) = \frac{c}{E(f)} T(u).$$

As proximal mappings are Lipshitz continuous with constant one and $\mathrm{prox}_\Phi(0) = 0$, we have

$$\|\hat{h}\|_2 = \|\mathrm{prox}_\Phi(g) - \mathrm{prox}_\Phi(0)\|_2 \leq \|g\|_2. \tag{11}$$

As $B(f)$ is one-homogeneous, $\langle f, v\rangle = B(f) > 0$, so that directly computing $||g||_2^2$ directly shows

$$||g||_2^2 = 1 + 2c\langle f, v\rangle + c^2||v||_2^2 > 1.$$

The inequality $\|g\|_2 \leq 1 + c\sqrt{n}$ follows from the fact that $||v||_2 \leq \sqrt{n}\|v\|_\infty \leq \sqrt{n}$ and the triangle inequality. The bound $0 < ||h||_2$ follows since $\hat{h}$ is not constant, and the upper bound $||h||_2 \leq (1+\sqrt{n})||\hat{h}||_2$ again follows from the triangle inequality.

For the second statement, as $f^k \in \mathcal{S}_0^{n-1}$ it follows that $E(f^k) \geq \alpha > 0$. From (6), then,

$$\|\hat{h}^k - f^k\|_2^2 \leq \frac{c}{\alpha} B(h^k)(E(f^k) - E(f^{k+1})). \tag{12}$$

From (9) we have $B(h^k) = \|h^k\|_1 \leq \sqrt{n}\|h^k\|_2 \leq (1+\sqrt{n})(\sqrt{n} + nc)$, and therefore

$$\|\hat{h}^k - f^k\|_2^2 \leq \frac{c}{\alpha}(1+\sqrt{n})(\sqrt{n}+nc)(E(f^k) - E(f^{k+1})) \to 0.$$

The last line follows as $E(f^k)$ is decreasing and bounded from below, and therefore converges. By continuity of the median and the fact that $\mathrm{med}(f^k) = 0$, any limit point of the $\{f^k\}$ must have median zero. As $\|\hat{h}^k - f^k\|_2^2 \to 0$, any limit point of the $\{\hat{h}^k\}$ must also have median zero, which implies that $\mathrm{med}(\hat{h}^k) \to 0$ as well. The triangle inequality then implies $||h^k - f^k|| \to 0$, so that $||h^k|| \to 1$ and $||f^{k+1} - f^k||_2 \to 0$ as desired. □

By the monotonicity result of Hein and Bühler [6] we have

Lemma 5 (Monotonicity of $\mathcal{A}_{\mathrm{IPM}}$)**.** *Let $f \in \mathcal{S}_0^{n-1}$. If $z \in \mathcal{A}_{\mathrm{IPM}}(f)$ then $E(z) < E(f)$ unless $z = f$.*

To prove convergence for $\mathcal{A}_{\mathrm{IPM}}$ using our techniques, we must also maintain control over the iterates after subtracting the median. This control is provided by the following lemma.

Lemma 6 (Compactness of $\mathcal{A}_{\mathrm{IPM}}$)**.** *Let $f \in \mathcal{S}_0^{n-1}$ and define v, D, g and h according to the IPM.*

1. *The minimizer is unique when $D < 0$, i.e. $g \in \mathcal{S}^{n-1}$ is a single point.*
2. *$1 \leq ||h||_2 \leq 1 + \sqrt{n}$. In particular, $\mathcal{A}_{\mathrm{IPM}}(f)$ is always well-defined for a given choice of $v \in \partial_0 B(f)$.*

Proof.

(1.) Let $D < 0$, and suppose there existed two distinct minimizers g_1 and g_2 that lie on the boundary of the unit ball. For any $0 < \theta < 1$ define $g_\theta = \theta g_1 + (1-\theta)g_2$ and note that $||g_\theta||_2 < 1$. By convexity of T and linearity of the inner product,

$$T(g_\theta) - E(f)\langle g_\theta, \partial_0 B(f)\rangle \leq \theta D + (1-\theta)D = D.$$

By one-homogeneity of T and the inner-product, and the fact that D is the global minimum it follows that

$$||g_\theta||_2 D \leq ||g_\theta||_2 \left[T\left(\frac{g_\theta}{||g_\theta||_2}\right) - E(f)\left\langle \frac{g_\theta}{||g_\theta||_2}, \partial_0 B(f)\right\rangle\right] \leq D.$$

This cannot happen as $D < 0$ and $||g_\theta||_2 < 1$.

(2.) If $D = 0$ then the statement holds trivially. Otherwise $D < 0$, so that if $||h||_2 < 1$ then

$$||h||_2 D \leq ||h||_2 \left[T\left(\frac{h}{||h||_2}\right) - E(f)\left\langle \frac{h}{||h||_2}, \partial_0 B(f)\right\rangle\right] = T(h) - E(f)\langle h, \partial_0 B(f)\rangle = D.$$

The last inequality follows since, due to the choice of subdifferential $\partial_0 B$, we may add a constant to the global minimizer g without changing the value of the expression. If $||h||_2 < 1$ we therefore arrive at a contradiction. From the triangle inequality it follows that also $||h||_2 \leq 1 + \sqrt{n}$.

(3.) The one-homogeneity of B and the definition of the subgradient combine to show $\langle h, \partial_0 B(f)\rangle \leq B(h)$. When $D < 0$ we have

$$T(h) - E(f)\langle h, \partial_0 B(f)\rangle = T(g) - E(f)\langle g, \partial_0 B(f)\rangle < 0$$

so that $T(h) < E(f)B(h)$. As $||h||_2 \geq 1$ and $\mathrm{med}(h) = 0$ we know h is non-constant, so we can divide by $B(h)$ to obtain $E(S(f)) = E(h) < E(f)$ as desired. If $D = 0$ then $\mathcal{A}_{\mathrm{IPM}}(f) = f$, so the claim follows. □

2.2 Closedness Properties

The final ingredient to prove local convergence is some form of closedness. We require closedness of the set valued maps $\mathcal{A}$ at strict local minima of the energy. As the energy (2) is invariant under constant shifts and scalings, the usual notion of a strict local minimum on $\mathbb{R}^n$ does not apply. We must therefore remove the effects of these invariances when referring to a local minimum as strict. To this end, define the spherical and annular neighborhoods on $\mathcal{S}_0^{n-1}$ by

$$\mathcal{B}_\epsilon(f^\infty) := \{||f - f^\infty||_2 \leq \epsilon\} \cap \mathcal{S}_0^{n-1} \quad \mathcal{A}_{\delta,\epsilon}(f^\infty) := \{\delta \leq ||f - f^\infty||_2 \leq \epsilon\} \cap \mathcal{S}_0^{n-1}.$$

With these in hand we introduce the proper definition of a strict local minimum.

Definiton 2 (Strict Local Minima)**.** *Let $f^\infty \in \mathcal{S}_0^{n-1}$. We say f^∞ is a **strict local minimum** of the energy if there exists $\epsilon > 0$ so that $f \in \mathcal{B}_\epsilon(f^\infty)$ and $f \neq f^\infty$ imply $E(f) > E(f^\infty)$.*

This definition then allows us to formally define closedness at a strict local minimum in Definition 3. For the IPM algorithm this is the only form of closedness we are able to establish. Closedness at an arbitrary $f \in \mathcal{S}_0^{n-1}$ (c.f. lemma 1) does in fact hold for the SD algorithm. Once again, this fact manifests itself in the stronger global convergence results for the SD algorithm in section 4.

Definiton 3 (CLM/CSLM Mappings)**.** *Let $\mathcal{A}(f) : \mathcal{S}_0^{n-1} \rightrightarrows \mathcal{S}_0^{n-1}$ denote a set-valued mapping. We say $\mathcal{A}(f)$ is **closed at local minima** (CLM) if $z^k \in \mathcal{A}(f^k)$ and $f^k \to f^\infty$ imply $z^k \to f^\infty$ whenever f^∞ is a local minimum of the energy. If $z^k \to f^\infty$ holds only when f^∞ is a strict local minimum then we say $\mathcal{A}(f)$ is **closed at strict local minima** (CSLM).*

The CLM property for the SD algorithm, provided by lemma 7, follows as a straight forward consequence of lemma 1. The CSLM property for the IPM algorithm provided by lemma 8 requires the additional hypothesis that the local minimum is strict.

Lemma 7 (CLM Property for $\mathcal{A}_{\mathrm{SD}}$)**.** *For $f \in \mathcal{S}_0^{n-1}$ define $g, \hat{h}$ and h according to the SD algorithm. Then $\mathcal{A}_{\mathrm{SD}}(f)$ defines a CLM mapping.*

Proof. Given $f^k \to f^\infty$ and $z^k \in \mathcal{A}(f^k)$, let $\hat{h}^k \in \mathcal{H}(f^k)$ be such that $h^k = \hat{h}^k - m(\hat{h}^k)\mathbf{1}$ and $z^k = h^k||h^k||_2^{-1}$. As $\{\hat{h}^k\}$ lies in a compact set, any subsequence of $\{\hat{h}^k\}$ has a further convergent subsequence $\hat{h}^{k_i} \to \hat{h}^\infty$. As $f^{k_i} \to f^\infty$ and $\mathcal{H}$ is closed, $\hat{h}^\infty \in \mathcal{H}(f^\infty)$. Thus, there exists $v^\infty \in \partial_0 B(f^\infty)$ so that

$$\hat{h}^\infty = \arg\min_u \left\{ T(u) + \frac{E(f^\infty)}{2c}||u - f^\infty - cv^\infty||_2^2 \right\}.$$

Note this happens if and only if $0 \in \partial T(\hat{h}^\infty) + \frac{E(f^\infty)}{c}(\hat{h}^\infty - f^\infty - cv^\infty)$. From the fact that $\partial T(f^\infty) - E(f^\infty)v^\infty = \partial T(f^\infty) + \frac{E(f^\infty)}{c}(f^\infty - f^\infty - cv^\infty)$ and the uniqueness of minimizers, it follows that $\hat{h}^\infty = f^\infty$ provided $0 \in \partial T(f^\infty) - E(f^\infty)v^\infty$. In this case, both h^{k_i} and z^{k_i} must then converge to f^∞ as well. As any subsequence of $\{z^k\}$ has a further subsequence that converges to f^∞, in fact the whole sequence converges to f^∞ as desired.

It remains only to show that $0 \in \partial T(f^\infty) - E(f^\infty)v^\infty$ whenever $v^\infty \in \partial_0 B(f^\infty)$ and f^∞ is a local minimum of the energy. Take $\epsilon > 0$ so that $f \in \mathcal{B}_\epsilon(f^\infty)$ implies $E(f) \geq E(f^\infty)$, and suppose that $E(f^\infty)v^\infty \notin \partial T(f^\infty)$. By definition, there then exists a $g \in \mathbb{R}^n$ so that

$$T(g) - T(f^\infty) < E(f^\infty)\langle v^\infty, g - f^\infty\rangle = E(f^\infty)\langle v^\infty, g\rangle - T(f^\infty).$$

For $0 < \theta < 1$ set $g_\theta := (1-\theta)f^\infty + \theta g$, then compute

$$T(g) < E(f^\infty)\langle v^\infty, g\rangle = \frac{1}{\theta}E(f^\infty)\langle v^\infty, g_\theta - (1-\theta)f^\infty\rangle$$
$$T(g_\theta) \leq (1-\theta)T(f^\infty) + \theta T(g) < E(f^\infty)\langle v^\infty, g_\theta\rangle$$

by using the fact that $B(f^\infty) = \langle v^\infty, f^\infty\rangle$ and the fact that T is convex. By definition of $\partial_0 B(f^\infty)$ it follows that $\langle v^\infty, g_\theta\rangle \leq B(g_\theta)$, which yields

$$T(g_\theta) < E(f^\infty)B(g_\theta)$$

whenever $0 < \theta < 1$. This implies $B(g_\theta) > 0$, so that $E(g_\theta) < E(f^\infty)$ for all $0 < \theta < 1$ as well. Put $\hat{g}_\theta = g_\theta - \mathrm{med}(g_\theta)\mathbf{1}$ and note that $\hat{g}_\theta \to f^\infty$ as $\theta \to 0$ since f^∞ has zero median. But $E(||\hat{g}_\theta||_2^{-1}\hat{g}_\theta) = E(\hat{g}_\theta) = E(g_\theta) < E(f^\infty)$ and $||\hat{g}_\theta||_2^{-1}\hat{g}_\theta \in \mathcal{B}_\epsilon(f^\infty)$ for all θ sufficiently close to zero, which contradicts the assumption that f^∞ is a local minimizer. □

Lemma 8 (CSLM Property for $\mathcal{A}_{\mathrm{IPM}}$)**.** *For $f \in \mathcal{S}_0^{n-1}$ define v, D, g, h according to the IPM. Then $\mathcal{A}_{\mathrm{IPM}}(f)$ defines a CSLM mapping.*

Proof. Consider a sequence of points $f^k \in \mathcal{B}_\epsilon$ with $f^k \to f^\infty$. Let $z^k = S(f^k)$ and also let D^k, g^k, h^k denote the intermediate steps in the algorithm above. We will show any subsequence of $\{z^k\}$ has a further subsequence that converges to f^∞.

Define

$$\mathcal{K} := \left\{k \in \mathbb{N} : D^k = 0\right\},$$

and consider an arbitrary subsequence of the z^k. If the subsequence has only finitely elements in $\mathcal{K}^c$, then $z^k = f^k$ for all but finitely many elements of the subsequence. Since then $z^k = f^k$ for all but finitely many k and $f^k \to f^\infty$ by hypothesis, the whole subsequence converges to f^∞. Otherwise, an infinite number of terms lie in $\mathcal{K}^c$. By restricting to only those elements of the subsequence that lie in $\mathcal{K}^c$, and by extracting enough convergent subsequences of (f^k, g^k, h^k, z^k) we may assume that

$$f^k \to f^\infty,\ g^k \to g^*,\ h^k \to h^* = g^* - \mathrm{med}(g^*)\mathbf{1},\ z^k \to z^* = \frac{h^*}{||h^*||_2} \in \mathcal{S}_0^{n-1}.$$

Since the subdifferential $\partial_0 B(f^k)$ is closed, we may assume (by extracting yet another subsequence) that $\partial_0 B(f^k) \to v^* \in \partial_0 B(f^\infty)$. Define

$$D^* = \min_{||u||_2 \leq 1} T(u) - E(f^\infty)\langle u, v^*\rangle$$

and assume for the sake of contradiction that

$$D^* < T(g^*) - E(f^\infty)\langle g^*, v^*\rangle,$$

i.e. that g^* does not attain the minimum. If this were the case, then there exists a q^* with $||q^*||_2 \leq 1$ with the property that

$$T(q^*) - E(f^\infty)\langle q^*, v^*\rangle < T(g^*) - E(f^\infty)\langle g^*, v^*\rangle.$$

But as

$$T(q^*) - E(f^\infty)\langle q^*, v^*\rangle = \lim_{k\to\infty} T(q^*) - E(f^k)\langle q^*, \partial_0 B(f^k)\rangle$$
$$T(g^*) - E(f^\infty)\langle g^*, v^*\rangle = \lim_{k\to\infty} T(g^k) - E(f^k)\langle g^k, \partial_0 B(f^k)\rangle$$

we see that $T(q^*) - E(f^k)\langle q^*, \partial_0 B(f^k)\rangle < T(g^k) - E(f^k)\langle g^k, \partial_0 B(f^k)\rangle$ for all k sufficiently large, which contradicts the definition of g^k as the global minimizer.

Suppose now that $z^* \neq f^\infty$, and recall that $D^* \leq 0$. Then from the preceeding argument, we know

$$D^* = T(g^*) - E(f^\infty)\langle g^*, v^*\rangle \leq 0.$$

From the fact that $\langle v^*, \mathbf{1}\rangle = 0$ we have

$$T(h^*) - E(f^\infty)\langle h^*, v^*\rangle = T(g^*) - E(f^\infty)\langle g^*, v^*\rangle \leq 0.$$

By using one-homogeneity of T it then follows that

$$P^* := T(z^*) - E(f^\infty)\langle z^*, v^*\rangle \leq 0$$

as well. Define $\hat{z}_\theta = \theta z^* + (1-\theta)f^\infty$ and also

$$z_\theta := \hat{z}_\theta - \mathrm{med}(\hat{z}_\theta)\mathbf{1}.$$

Again as $\langle v^*, \mathbf{1}\rangle = 0$, the convexity of T and linearity of the inner product imply

$$T(z_\theta) - E(f^\infty)\langle z_\theta, v^*\rangle = T(\hat{z}_\theta) - E(f^\infty)\langle \hat{z}_\theta, v^*\rangle \leq \theta P^* \leq 0.$$

The fact that $B(z_\theta) \geq \langle z_\theta, v^*\rangle$ then implies the inequality

$$E(z_\theta) \leq E(f^\infty)$$

holds for all $0 < \theta < 1$. As $\mathrm{med}(\hat{z}_\theta) \to 0$ as $\theta \to 0$, by the reverse triangle inequality it follows that for all θ small

$$||z_\theta||_2 \geq 1 - 2\theta - \mathrm{med}(\hat{z}_\theta)\sqrt{n} \geq 1/4.$$

From scale invariance of the energy, for all such θ we have that both

$$E\left(\frac{z_\theta}{||z_\theta||_2}\right) = E(z_\theta) \leq E(f^\infty)$$

and

$$\left\|\frac{z_\theta}{||z_\theta||_2} - f^\infty\right\|_2 \leq 4||z_\theta - ||z_\theta||_2 f^\infty||_2 \to 0$$

hold as $\theta \to 0$. Since $z_\theta/||z_\theta||_2 \in \mathcal{S}_0^{n-1}$, this contradicts the fact that f^∞ is a strict local minimium in $\mathcal{B}_\epsilon$. Thus, we must have $z^* = f^\infty$. Therefore any subsequence of $\{z^k\}$ has a further subsequence that converges to f^∞. This implies that in fact the whole sequence z^k converges to f^∞ as desired. □

3 Local Convergence of $\mathcal{A}_{\mathrm{SD}}$ and $\mathcal{A}_{\mathrm{IPM}}$ at Strict Local Minima

Due to the lack of convexity of the energy (2) , at best we can only hope to obtain convergence to a local minimum of the energy. An analogue of Lyapunov's method from differential equations allows us to show that such convergence does occur provided the iterates reach a neighborhood of an isolated local minimum. To apply the lemmas from section 2 we must assume that $f^\infty \in \mathcal{S}_0^{n-1}$ is a local minimum of the energy. We will assume further that f^∞ is an isolated critical point of the energy according to the following definition.

Definiton 4 (Isolated Critical Points)**.** *Let $f \in \mathcal{S}_0^{n-1}$. We say that f is a **critical point** of the energy $E(f)$ if there exist $w \in \partial T(f)$ and $v \in \partial_0 B(f)$ so that*

$$0 = w - E(f)v.$$

This generalizes the usual quotient rule

$$0 = \nabla T(f) - E(f)\nabla B(f).$$

*If there exists $\epsilon > 0$ so that f is the only critical point in $\mathcal{B}_\epsilon(f^\infty)$ we say f is an **isolated critical point** of the energy.*

Note that as any local minimum is a critical point of the energy, if f^∞ is an isolated critical point and a local minimum then it is necessarily a strict local minimum. The CSLM property therefore applies.

Finally, to show convergence, the set-valued map $\mathcal{A}$ must possess one further property, i.e. the critical point property.

Definiton 5 (Critical Point Property)**.** *Let $\mathcal{A}(f) : \mathcal{S}_0^{n-1} \rightrightarrows \mathcal{S}_0^{n-1}$ denote a set-valued mapping. We say that $\mathcal{A}(f)$ satisfies the **critical point property** (CP property) if, given any sequence satisfying $f^{k+1} \in \mathcal{A}(f^k)$, all limit points of $\{f^k\}$ are critical points of the energy.*

Analogously to the CLM property, for the SD algorithm the CP property follows as a direct consequence of lemma 1. For the proof, see the first statement in theorem 2. We establish this for the IPM algorithm in the following lemma.

Lemma 9 (CP Property for the IPM Algorithm)**.** *The set-valued mapping $\mathcal{A}_{\text{IPM}}(f) : \mathcal{S}_0^{n-1} \rightrightarrows \mathcal{S}_0^{n-1}$ satisfies the critical point property.*

Proof. Let $f^{k_j} \to f^* \in \mathcal{S}_0^{n-1}$ denote a convergent subsequence. Define $v^{k_j}, D^{k_j}, g^{k_j}$ and h^{k_j} according to the IPM algorithm. By compactness, we can extract enough further subsequences (still denoted f^{k_j}) to find

$$f^{k_j} \to f^* \quad g^{k_j} \to g^* \quad v^{k_j} \to v^* \in \partial_0 B(f^*).$$

The fact that $v^* \in \partial_0 B(f^*)$ follows from the closedness established in lemma 2. Define

$$D^* := \min_{||u||_2 \leq 1} T(u) - E(f^*)\langle u, v^* \rangle.$$

As in the proof of the CSLM property we know g^* must attain the minimum, i.e. $D^* = T(g^*) - E(f^*)\langle g^*, v^* \rangle$. Suppose that $D^* < 0$. Then as

$$D^* = \lim_{j \to \infty} T(g^{k_j}) - E(f^*)\langle g^{k_j}, v^* \rangle,$$

there exists J sufficiently large so that $j \geq J$ implies

$$T(h^{k_j}) - E(f^*)\langle h^{k_j}, v^* \rangle = T(g^{k_j}) - E(f^*)\langle g^{k_j}, v^* \rangle < 0.$$

But this implies

$$E(f^{k_j+1}) = E(h^{k_j}) < E(f^*)$$

for all j sufficiently large, a contradiction. Thus $D^* = 0$ and f^* must be the minimizer of

$$\min_{u \in \mathbb{R}^n} T(u) - E(f^*)\langle u, v^* \rangle.$$

This implies $0 \in \partial T(f^*) - E(f^*)v^*$ so f^* is a critical point as desired. □

The proof of local convergence utilizes a version of Lyapunov's direct method for set-valued maps, and we adapt this technique from the strategy outlined in [8]. We first demonstrate that if any iterate f^k lies in a sufficiently small neighborhood $\mathcal{B}_\gamma(f^\infty)$ of the strict local minimum then all subsequent iterates remain in the neighborhood $\mathcal{B}_\epsilon(f^\infty)$ in which f^∞ is an isolated critical point. By compactness and the CP property, any subsequence of $\{f^k\}$ must have a further subsequence that converges to the only critical point in $\mathcal{B}_\epsilon(f^\infty)$, i.e. f^∞. This implies that the whole sequence must converge to f^∞ as well. We formalize this argument in lemma 10 and its corollary theorem 1.

Lemma 10 (Lyapunov Stability at Strict Local Minima)**.** *Suppose $\mathcal{A}(f)$ is a monotonic, CSLM mapping. Fix $f^0 \in \mathcal{S}_0^{n-1}$ and let $\{f^k\}$ denote any sequence satisfying $f^{k+1} \in \mathcal{A}(f^k)$. If f^∞ is a strict local minimum of the energy, then for any $\epsilon > 0$ there exists a $\gamma > 0$ so that if $f^0 \in \mathcal{B}_\gamma(f^\infty)$ then $\{f^k\} \subset \mathcal{B}_\epsilon(f^\infty)$.*

Proof. The proof follows [8]. By taking ϵ smaller if necessary, we can assume that f^∞ is a strict local minimum on $\mathcal{B}_\epsilon(f^\infty)$. From the CSLM property, we can choose $0 < \delta < \epsilon$ small enough to guarantee

$$f \in \mathcal{B}_\delta \quad \text{implies} \quad \mathcal{A}(f) \subset \mathcal{B}_\epsilon.$$

For such a choice of δ, define

$$\mu := \min_{f \in \mathcal{A}_{\delta,\epsilon}(f^\infty)} E(f) - E(f_\infty) > 0.$$

By continuity of E on $\mathcal{S}_0^{n-1}$, we can then choose $0 < \gamma < \delta$ small enough so that $f \in \mathcal{B}_\gamma$ implies $E(f) - E(f_\infty) < \frac{\mu}{2}$. Take any initial point $f^0 \in \mathcal{B}_\gamma$. Let K be the first integer so that $||f^K - f_\infty||_2 \geq \delta$. By assumption, since $f^{K-1} \in \mathcal{B}_\delta$ we must have $f^K \in \mathcal{A}_{\delta,\epsilon}(f^\infty)$. But then

$$E(f^K) - E(f_\infty) \geq \mu$$

by definition as well. However, since E always decreases we must have

$$\frac{\mu}{2} \geq E(f^0) - E(f^\infty) \geq E(f^K) - E(f_\infty) \geq \mu,$$

which is a contradiction. Thus, the whole sequence $\{f^k\} \subset \mathcal{B}_\delta \subset \mathcal{B}_\epsilon$. □

Theorem 1 (Local Convergence at Isolated Critical Points)**.** *Let $\mathcal{A}(f) : \mathcal{S}_0^{n-1} \rightrightarrows \mathcal{S}_0^{n-1}$ denote a monotonic, CSLM, CPP mapping. Let $f^0 \in \mathcal{S}_0^{n-1}$ and suppose $\{f^k\}$ is any sequence satisfying $f^{k+1} \in \mathcal{A}(f^k)$. Let f^∞ denote a local minimum that is an isolated critical point of the energy. If $f^0 \in \mathcal{B}_\gamma(f^\infty)$ for $\gamma > 0$ sufficiently small then $f^k \to f^\infty$.*

Proof. Choose $\epsilon > 0$ so that f^∞ is the only critical point of the energy in $\mathcal{B}_\epsilon$. By stability of CSLM mappings, we can choose $\gamma > 0$ so that $f^0 \in \mathcal{B}_\gamma$ implies $\{f^k\} \subset \mathcal{B}_\epsilon$. By compactness of $\{f^k\}$ and the critical point property, any subsequence has a further subsequence that converges to a critical point of the energy that lies in $\mathcal{B}_\epsilon$. As f^∞ is the only such critical point, we find any subsequence of $\{f^k\}$ has a further subsequence that converges to f^∞, so the whole sequence converges as desired. □

Note that both algorithms satisfy the hypothesis of theorem 1, and therefore possess identical local convergence properties. A slight modification of the proof of theorem 1 yields the following corollary that also applies to both algorithms.

Corollary 1. *Let $f^0 \in \mathcal{S}_0^{n-1}$ be arbitrary, and define $f^{k+1} \in \mathcal{A}(f^k)$ according to either algorithm. If any accumulation point f^* of the sequence $\{f^k\}$ is both an isolated critical point of the energy and a local minimum, then the whole sequence $f^k \to f^*$.*

4 Global Convergence for $\mathcal{A}_{\mathrm{SD}}$

To this point the convergence properties of both algorithms appear identical. However, we have yet to take full advantage of the superior mathematical structure afforded by the SD algorithm. In particular, from lemma 4 we know that $||f^{k+1} - f^k||_2 \to 0$ without any further assumptions regarding the initialization of the algorithm or the energy landscape. This fact combines with the fact that lemma 1 also holds globally for $f \in \mathcal{S}_0^{n-1}$ to yield theorem 2. Once again, we arrive at this conclusion by adapting the proof from [8].

Theorem 2 (Convergence of the SD Algorithm)**.** *Take $f^0 \in \mathcal{S}_0^{n-1}$ and fix a constant $c > 0$. Let $\{f^k\}$ denote any sequence satisfying $f^{k+1} \in \mathcal{A}_{\mathrm{SD}}(f^k)$. Then*

1. *Any accumulation point f^* of the sequence is a critical point of the energy.*
2. *Either the sequence converges, or the set of accumulation points form a continuum in $\mathcal{S}_0^{n-1}$.*

Proof. (1.) The proof is inspired by [8]. Let $f^{k_i} \to f^*$ denote a convergent subsequence. As $\{f^{k_i+1}\} \subset \mathcal{S}_0^{n-1}$, we may assume (after extracting a further subsequence if necessary) that there exits $f' \in \mathcal{S}_0^{n-1}$ so that, as $i \to \infty$,

$$f^{k_i} \to f^* \tag{13}$$

$$f^{k_i+1} \to f'. \tag{14}$$

However, because of (10) we have

$$f' = f^* = \lim_{i\to\infty} \hat{h}^{k_i} \in \mathcal{H}(f^{k_i}). \tag{15}$$

Therefore, as $f^{k_i} \to f^*$ and $\mathcal{H}$ is closed we have $f^* \in \mathcal{H}(f^*)$. By definition of $\mathcal{H}(f^*)$, if $f^* \in \mathcal{H}(f^*)$ then there exists $y^* \in \mathcal{Y}^c(f^*)$ so that

$$f^* = \arg\min_u \left\{ T(u) + E(f^*) \frac{\|u - y^*\|_2^2}{2c} \right\}.$$

Therefore there exists $w^* \in \partial T(f^*)$ so that $0 = cw^* + E(f^*)(f^* - y^*)$. By definition of $\mathcal{Y}^c(f^*)$ there exists $v^* \in \partial_0 B(f^*)$ so that

$$0 = cw^* + E(f^*)(f^* - (f^* + cv^*)) = c(w^* - E(f^*)v^*).$$

Thus f^* is a critical point of the energy according to definition 5.

(2.) For any sequence generated by the algorithm, $||f^{k+1} - f^k||_2 \to 0$ according to (10). Moreover, they lie in the bounded set $\mathcal{S}^{n-1} \subset \mathbb{R}^n$. The hypotheses of Theorem 26.1 of [9] are therefore satisfied, giving the desired conclusion. □

We might hope to rule out the second possibility in statement 2 by showing that E can nevery have an uncountable number of critical points. Unfortunately, we can exhibit (c.f. section 5.3) simple examples to show that a continuum of local or global minima can in fact happen. This degeneracy of a continuum of critical points arises from a lack of uniqueness in the underlying combinatorial problem. We explore this aspect of convergence further in section 5.

By assuming additional structure in the energy landscape we can generalize the local convergence result, theorem 1, to yield global convergence of both algorithms. This is the content of corollary 2 for the SD algorithm and the content of corollary 3 for the IPM algorithm. The hypotheses required for each corollary clearly demonstrate the benefit of knowing apriori that $||f^{k+1} - f^k||_2 \to 0$ occurs for the SD algorithm. For the IPM algorithm, we can only deduce this aposteriori from the fact that the iterates converge.

Corollary 2. *Let $f^0 \in \mathcal{S}_0^{n-1}$ be arbitrary and define $f^{k+1} \in \mathcal{A}_{\mathrm{SD}}(f^k)$. If the energy has only countably many critical points in $\mathcal{S}_0^{n-1}$ then $\{f^k\}$ converges.*

Corollary 3. *Let $f^0 \in \mathcal{S}_0^{n-1}$ be arbitrary and define $f^{k+1} \in \mathcal{A}_{\mathrm{IPM}}(f^k)$. Suppose all critical points of the energy are isolated in $\mathcal{S}_0^{n-1}$ and are either local maxima or local minima. Then $\{f^k\}$ converges.*

Proof. Let $\{f^k\} \subset \mathcal{S}_0^{n-1}$ denote any sequence satisfying $f^{k+1} \in \mathcal{A}_{\mathrm{IPM}}(f^k)$. Assume first that $0 \notin \partial T(f^k) - E(f^k)\partial_0 B(f^k)$ for infinitely many k. Then there exists a subsequence f^{k_j} with the property that $E(f^{k_{j+1}}) < E(f^{k_j})$ for all j. We can extract a further subsequence (still denoted $\{f^{k_j}\}$) and a point f^* so that $f^{k_j} \to f^*$. By the CP property it follows that f^* is a critical point, hence either a local maximum or a local minimum. However, as $E(f^{k_j}) > E(f^*)$ for all j and $||f^{k_j} - f^*||_2 \to 0$ we conclude that f^* cannot be a local maximum. Thus, as all critical points are isolated we know f^* is actually a strict local minimum, so $f^k \to f^*$ by corollary 1.

Otherwise, there exists K sufficiently large so that $0 \in \partial T(f^k) - E(f^k)\partial_0 B(f^k)$ for all $k \geq K$. But then $D^k = 0$ for all $k \geq K$, which implies that $f^k = f^K$ for all $k \geq K$ by definition of the iterates, so the algorithm converges. □

While at first glance corollary 3 provides hope that global convergence holds for the IPM algorithm, our simple examples (c.f. section 5.3) demonstrate that even benign graphs with well-defined cuts have critical points of the energy that are neither local maxima nor local minima.

5 Energy Landscape of the Cheeger Functional

This section demonstrates that the continuous problem (2) provides an exact relaxation of the combinatorial problem (1). Specifically, we provide an explicit formula that gives an exact correspondence between the global minimizers of the continuous problem and the global minimizers of the combinatorial problem. This extends previous work [13, 12, 10] on the relationship between the global minima of (1) and (2). We also completely classifiy the local minima of the continuous problem by introducing a notion of local minimum for the combinatorial problem. Any local minimum of the combinatorial problem then determines a local minimum of the combinatorial problem by means of an explicit formula, and vice-versa. Theorem 4 provides this formula, which also gives a sharp condition for when a global minimum of the continuous problem is two-valued (binary), three-valued (trinary), or k-valued in the general case. This provides an understanding the energy landscape, which is essential due to the lack of convexity present in the continuous problem. Most importantly, we can classify the types of local minima encountered and when they form a continuum. This is germane to the global convergence results of the previous sections. The proofs in this section follow closely the ideas from [13, 12].

5.1 Local and Global Minima

We first introduce the two fundamental definitions of this section. The first definition introduces the concept of when a set $S \subset V$ of vertices is compatible with an increasing sequence $S_1 \subsetneq S_2 \subsetneq \cdots \subsetneq S_k$ of vertex subsets. Loosely speaking, a set S is compatible with $S_1 \subsetneq S_2 \subsetneq \cdots \subsetneq S_k$ whenever the cut defined by the pair (S, S^c) neither intersects nor crosses any of the cuts (S_i, S_i^c). Definition 6 formalizes this notion.

Definiton 6 (Compatible Vertex Set)**.** *A vertex set S is* **compatible** *with an increasing sequence $S_1 \subsetneq S_2 \subsetneq \cdots \subsetneq S_k$ if $S \subseteq S_1,\ S_k \subseteq S$ or*

$$S_1 \subsetneq S_2 \subsetneq \cdots \subsetneq S_i \subseteq S \subseteq S_{i+1} \subsetneq \cdots \subsetneq S_k \quad \textit{for some } 1 \leq i \leq k-1,$$

The concept of compatible cuts then allows us to introduce our notion of a local minimum of the combinatorial problem, i.e. definition 7.

Definiton 7 (Combinatorial k-Local Minima)**.** *An increasing collection of nontrivial sets $S_1 \subsetneq S_2 \subsetneq \cdots \subsetneq S_k$ is called a* **k-*local minimum*** *of the combinatorial problem if $\mathcal{C}(S_1) = \mathcal{C}(S_2) = \cdots = \mathcal{C}(S_k) \leq \mathcal{C}(S)$ for all S compatible with $S_1 \subsetneq S_2 \subsetneq \cdots \subsetneq S_k$.*

Pursuing the previous analogy, a collection of cuts $(S_1, S_1^c), \cdots, (S_k, S_k^c)$ forms a k-local minimum of the combinatorial problem precisely when they do not intersect, have the same energy and all other non-intersecting cuts (S, S^c) have higher energy. The case of a 1-local minimum is paramount. A cut (S_1, S_1^c) defines a 1-local minimum if and only if it has lower energy than all cuts that do not intersect it. As a consequence, if a 1-local minimum is not a global minimum then the cut (S_1, S_1^c) necessarily intersects all of the cuts defined by the global minimizers. This is a fundamental characteristic of local minima: they are never "parallel" to global minima.

For the continuous problem, combinatorial k-local minima naturally correspond to vertex functions $f \in \mathbb{R}^n$ that take $(k+1)$ distinct values. We therefore define the concept of a $(k+1)$-valued local minimum of the continuous problem.

Definiton 8 (Continuous $(k+1)$-valued Local Minima)**.** *We call a vertex function $f \in \mathbb{R}^n$ a* **$(\mathbf{k+1})$-valued local minimum** *of the continuous problem if f is a local minimum of E and if its range contains exactly $k+1$ distinct values.*

Theorem 3 provides the intuitive picture connecting these two concepts of minima, and it follows as a corollary of the more technical and explicit theorem 4.

Theorem 3. *The continuous problem has a $(k+1)$-valued local minimum if and only if the combinatorial problem has a k-local minimum.*

For example, if the continuous problem has a trinary local minimum in the usual sense then the combinatorial problem must have a 2-local minimum in the sense of definition 7. As the cuts (S_1, S_1^c) and (S_2, S_2^c) defining a 2-local minimum do not intersect, a 2-local minimum separates the vertices of the graph into three disjoint domains. A trinary function therefore makes intuitive sense. We make this intuition precise in theorem 4. Before stating it we require two further definitions.

Definiton 9 (Characteristic Functions)**.** *Given $\emptyset \neq S \subset V$, define its* **characteristic function** *f_S as*

$$f_S = \textit{Cut}(S, S^c)^{-1}\chi_S \quad \textit{if } |S| \leq n/2 \qquad \textit{and} \qquad f_S = -\textit{Cut}(S, S^c)^{-1}\chi_{S^c} \quad \textit{if } |S| > n/2. \tag{16}$$

Note that f_S has median zero and TV-norm equal to 1.

Definiton 10 (Strict Convex Hull)**.** *Given k functions $f_1, \cdots, f_k$, their* **strict convex hull** *is the set*

$$\textit{sch}\{f_1, \cdots, f_k\} = \{\theta_1 f_1 + \cdots + \theta_k f_k : \theta_i > 0 \ \textit{for}\ 1 \leq i \leq k \ \textit{and}\ \theta_1 + \cdots + \theta_k = 1\} \tag{17}$$

Theorem 4 (Explicit Correspondence of Local Minima)**.**

1. *Suppose $S_1 \subsetneq S_2 \subsetneq \cdots \subsetneq S_k$ is a k-local minimum of the combinatorial problem and let $f \in \textit{sch}\{f_{S_1}, \cdots, f_{S_k}\}$. Then any function of the form $g = \alpha f + \beta \mathbf{1}$ defines a $(k+1)$-valued local minimum of the continuous problem and with $E(g) = \mathcal{C}(S_1)$.*
2. *Suppose that f is a $(k+1)$-valued local minimum and let $c_1 > c_2 > \cdots > c_{k+1}$ denote its range. For $1 \leq i \leq k$ set $\Omega_i = \{f = c_i\}$. Then the increasing collection of sets $S_1 \subsetneq \cdots \subsetneq S_k$ given by*

$$S_1 = \Omega_1, \quad S_2 = \Omega_1 \cup \Omega_2 \qquad \cdots \qquad S_k = \Omega_1 \cup \cdots \cup \Omega_k$$

is a k-local minimum of the combinatorial problem with $\mathcal{C}(S_i) = E(f)$.

Remark 1 (Isolated vs Continuum of Local Minima)**.** *If a set S_1 is a 1-local min then the strict convext hull* (17) *of its characteristic function reduces to the single binary function f_{S_1}. Thus every*

1-local minimum generates exactly one local minimum of the continuous problem in $\mathcal{S}_0^{n-1}$, and this local minimum is binary. On the other hand, if $k \geq 2$ then every k-local minimum of the combinatorial problem generates a continuum (in $\mathcal{S}_0^{n-1}$) of non-binary local minima of the continuous problem. As a consequence, the hypotheses of theorem 1, corollary 2 or corollary 3 can hold only if no such higher order k-local minima exist. When these theorems do apply the algorithms therefore converge to a binary function.

As a final consequence, we summarize the fact that theorem 4 implies that the continuous relaxation of the Cheeger cut problem is exact. In other words,

Theorem 5. *Given $\{f \in \arg\min E\}$ there exists an explicit formula to construct the set $\{S \in \arg\min \mathcal{C}\}$, and vice-versa.*

5.2 Proofs of Lemmas and Theorems

The proof closely follows the arguments from [13]. Define the median of $f \in \mathbb{R}^n$ as

$$\text{med}(f) = \min\{c \in \text{range}(f) \text{ satisfying } |\{f \leq c\}| \geq n/2\} \tag{18}$$

By this definition, the median of f is the $n/2$ smallest entry when n is even.

We define the TV-sphere $\mathfrak{X}$ by:

$$\mathfrak{X} = \{f \in \mathbb{R}^n : T(f) = 1 \text{ and } \text{med}(f) = 0\}.$$

Definiton 11 (Local Minima on the TV-sphere)**.** *$f \in \mathfrak{X}$ is a* **local minimum on the TV-sphere** *if there exists $\epsilon > 0$ such that $E(f) \leq E(g)$ for all $g \in \mathfrak{X}$ satisfying $\|g - f\|_2 \leq \epsilon$.*

The following lemma states that it is enough to consider local minima of E on the TV-sphere.

Lemma 11. *A non constant function $f \in \mathbb{R}^n$ is a local minimum of E in the usual sense if and only if $\tilde{f} = (f - \text{med}(f))/T(f - \text{med}(f))$ is a local minimum of E on the TV-sphere.*

Proof. Suppose that

$$\tilde{f} = \text{Proj}_{\mathfrak{X}}(f) = \frac{f - \text{med}(f)}{T(f - \text{med}(f))} \tag{19}$$

is a local minimum of E on $\mathfrak{X}$ but f is not a local minimum of E on $\mathbb{R}^n_{\text{non-cst}}$. Then there exists $f_n \to f$ with $E(f_n) < E(f)$. By continuity of $\text{Proj}_{\mathfrak{X}}$, $\tilde{f}_n \to \tilde{f}$ and since E is invariant under $\text{Proj}_{\mathfrak{X}}$, we have $E(\tilde{f}_n) < E(\tilde{f})$ which is a contradiction. Suppose now that f is a local min of E in $\mathbb{R}^n_{\text{non-cst}}$ but $\tilde{f}$ is not a local min of E on $\mathfrak{X}$. Then there exists $\tilde{f}_n \to \tilde{f}$ with $E(\tilde{f}_n) < E(\tilde{f})$. Since there exists $\alpha \neq 0$ and β such that $f = \alpha\tilde{f} + \beta\mathbf{1}$ it is clear that $\alpha\tilde{f}_n + \beta\mathbf{1} \to f$ and $E(\alpha\tilde{f}_n + \beta\mathbf{1}) < E(f)$ which is a contradiction. □

Recall that a polyedron is a set defined by a finite number of linear equalities and inequalities, and that it is necessarily convex. Given a permutation $\sigma \in \mathfrak{S}_n$ the polyedron

$$\mathcal{P}_\sigma = \{f \in \mathbb{R}^n : f_{\sigma(1)} \geq f_{\sigma(2)} \geq \cdots \geq f_{\sigma(n)}\}.$$

represents one possible ordering of the function $f \in \mathbb{R}^n$. We then define the face $\mathfrak{F}_\sigma$ of the TV-sphere by

$$\mathfrak{F}_\sigma = \{f \in \mathbb{R}^n : f \in \mathcal{P}_\sigma, \|f\|_{TV} = 1 \text{ and } \text{med}(f) = 0\}.$$

As the median and the total variation are linear functions on $\mathcal{P}_\sigma$, we have simply added two linear constraints so that $\mathfrak{F}_\sigma$ is also a polyhedron . Obviously we have

$$\mathfrak{X} = \cup_\sigma \mathfrak{F}_\sigma$$

where the union is taken over all possible permutations. Using the same arguments from [13, Lemma 2.1] yields:

Lemma 12 ([13])**.** *Suppose $f \in \mathfrak{F}_\sigma$. Then f is a binary function if and only if f is an extreme point of $\mathfrak{F}_\sigma$.*

The next lemma then gives an explicit description of the face $\mathfrak{F}_\sigma$.

Lemma 13. *$\mathfrak{F}_\sigma$ is the $n-2$ dimensional simplex*

$$\mathfrak{F}_\sigma = ch\{f_{S_1}, f_{S_2} \cdots, f_{S_{n-1}}\}. \tag{20}$$

Here, $ch\{f_{S_1}, f_{S_2} \cdots, f_{S_{n-1}}\}$ denotes the convex hull of the characteristic functions f_{S_i} of the increasing sequence of sets $S_1 \subsetneq S_2 \subsetneq \cdots \subsetneq S_{n-1}$ defined by $S_i = \{x_{\sigma(1)}, \cdots, x_{\sigma(i)}\}, 1 \leq i \leq n-1$. Moreover the functions $f_{S_1}, f_{S_2} \cdots, f_{S_{n-1}}$ are linearly independent.

Proof. The fact that $f_{S_1}, f_{S_2}, \cdots, f_{S_{n-1}}$ are linearly independent (and therefore affinely independent) can be directly read from definition (16) of f_S. Also from this same definition it is clear that $f_{S_1}, f_{S_2}, \cdots, f_{S_{n-1}}$ are binary functions that belongs to $\mathcal{F}_\sigma$, and that these are the only such binary functions. The conclusion then comes from the fact that a compact convex set is the convex hull of its extreme points. □

Proposition 1 (Decomposition in Binary Functions)**.** *Let $f \in \mathfrak{X}$ be a function whose range contains exactly k distinct values. Then there exists a unique increasing collection of nontrivial sets $S_1 \subsetneq S_2 \subsetneq \cdots \subsetneq S_k$ and a unique vector $\theta = (\theta_1, \cdots, \theta_k) \succ 0$, $\theta \cdot \mathbf{1} = 1$, so that*

$$f = \sum_{i=1}^{k} \theta_i f_{S_i}. \tag{21}$$

We will refer to (21) *as the* **decomposition of f in binary functions**.

Proof. Since $f \in \mathfrak{F}_\sigma$ for some permutation σ the existence of such a decomposition is clear from Lemma 13. Suppose now that (21) is such a decomposition. To show that the decomposition is unique, let i_0 be such that $|S_{i_0}| \leq n/2$ and $|S_{i_0+1}| > n/2$. Also let $\alpha_i = 1/\text{Cut}(S_i, S_i^c) > 0$. Then combining (16) and (21) we find that

$$f(x) = \begin{cases} \alpha_1\theta_1 + \alpha_2\theta_2 + \alpha_3\theta_3 + \cdots + \alpha_{i_0}\theta_{i_0} & \text{if } x \in S_1 \\ \alpha_2\theta_2 + \alpha_3\theta_3 + \cdots + \alpha_{i_0}\theta_{i_0} & \text{if } x \in S_2 \backslash S_1 \\ \alpha_3\theta_3 + \cdots + \alpha_{i_0}\theta_{i_0} & \text{if } x \in S_3 \backslash S_2 \\ \vdots & \vdots \\ \alpha_{i_0}\theta_{i_0} & \text{if } x \in S_{i_0} \backslash S_{i_0-1} \\ 0 & \text{if } x \in S_{i_0+1} \backslash S_{i_0} \\ -\alpha_{i_0+1}\theta_{i_0+1} & \text{if } x \in S_{i_0+2} \backslash S_{i_0+1} \\ -\alpha_{i_0+1}\theta_{i_0+1} - \alpha_{i_0+2}\theta_{i_0+2} & \text{if } x \in S_{i_0+3} \backslash S_{i_0+2} \\ \vdots & \vdots \\ -\alpha_{i_0+1}\theta_{i_0+1} - \alpha_{i_0+2}\theta_{i_0+2} - \cdots - \alpha_k\theta_k & \text{if } x \in V \backslash S_k \end{cases}$$

Therefore f takes its greatest value on S_1, its second greatest value on $S_2 \backslash S_1$, etc. As a consequence the sets S_i are uniquely determined by f, and since the f_{S_i} are linearly independent there is a unique possible choice for the θ_i. □

As a direct corollary of the previous proof, the decomposition of a function f in binary functions can easily be recovered.

Corollary 4. *Suppose $f \in \mathfrak{X}$ and range$(f) = \{c_1, \cdots, c_k\}$ where $c_1 > c_2 > \cdots > c_k$. Let $f = \sum_{i=1}^k \theta_i f_{S_i}$ be its unique decomposition in binary functions. Then*

$$S_i = \bigcup_{j=1}^{i} \{f = c_j\}, \quad i = 1, \cdots, k-1.$$

Lemma 14. *Let $f \in \mathfrak{X}$ and let $\sum_{i=1}^k \theta_i f_{S_i}$ be its unique decomposition in binary functions. Also let S be a nontrivial set. Then f and f_S belong to a common face of the TV-sphere if and only if S is compatible with $S_1 \subsetneq \cdots \subsetneq S_k$.*

Proof. Suppose S is not compatible with with $S_1 \subsetneq \cdots \subsetneq S_k$. Then there exists S_i such that $S \not\subseteq S_i$ and $S_i \not\subseteq S$. Then there exists $x_{\text{in}} \in S \backslash S_i$ and $x_{\text{out}} \in S_i \backslash S$. Since $x_{\text{in}} \in S$ and $x_{\text{out}} \notin S$ it is clear from (16) that $f_S(x_{\text{in}}) > f_S(x_{\text{out}})$. On the other hand, since $x_{\text{out}} \in S_i$ and $x_{\text{in}} \notin S_i$ we have by definition of the binary decomposition that $f(x_{\text{out}}) > f(x_{\text{in}})$. This follows as the values that f takes on S_i are greater or equal than the values that it takes outside of S_i. Thus f_S and f have a different ordering and therefore cannot belong to a common face of the TV-sphere.

Similarly, if f and f_S have a different ordering then there exist two points x_{in} and x_{out} such that $f_S(x_{\text{in}}) > f_S(x_{\text{out}})$ and $f(x_{\text{in}}) < f(x_{\text{out}})$. Clearly $x_{\text{in}} \in S$ and $x_{\text{out}} \notin S$. On the other hand there must exist an S_i such that $x_{\text{in}} \notin S_i$ and $x_{\text{out}} \in S_i$. This implies that $S \not\subseteq S_i$ and $S_i \not\subseteq S$. Therefore S is not compatible with $S_1 \subsetneq \cdots \subsetneq S_k$. □

We are now ready to prove theorem 4.

Proof of Theorem 4. Given a function $f \in \mathfrak{X}$, define its binary neighbors on the TV-sphere by

$$\mathcal{N}_{\text{bin}}(f) = \{g \in \mathfrak{X} : g \text{ is binary and } f \text{ and } g \text{ belong to a common face of the TV-sphere}\}.$$

A function $f \in \mathfrak{X}$ is a local minimum of E on the TV-sphere if and only if f is a local max of the ℓ^1-norm on the TV-sphere. As we restrict to functions with zero median, the ℓ^1-norm is a linear function on each face $\mathfrak{F}_\sigma$. Therefore a function f is a local maximum of the ℓ^1-norm if and only if

$$\|f\|_1 \geq \|g\|_1 \qquad \text{for all } g \in \mathcal{N}_{\text{bin}}(f). \tag{22}$$

Indeed, if f has a binary neighbor g with strictly greater ℓ^1 norm then any function of the form $\theta f + (1-\theta)g, \theta \in (0,1)$ has strictly greater ℓ^1-norm than f. Therefore f is not a local maximum. On the other hand assume that (22) holds and let $\mathfrak{F}_\sigma$ be a face to which f belongs. Then all the extreme points of $\mathfrak{F}_\sigma$ belong to $\mathcal{N}_{\text{bin}}(f)$ and therefore f has ℓ^1-norm greater than or equal to that of the extreme points. Therefore f has ℓ^1-norm greater or equal than all the functions in $\mathfrak{F}_\sigma$. As the face $\mathfrak{F}_\sigma$ to which f belonged was arbitrary, f must be a local maximum.

To prove the first statement of the theorem, suppose $S_1 \subsetneq S_2 \subsetneq \cdots \subsetneq S_k$ is a k-local minimum of the combinatorial problem and that $f \in \text{sch}\{f_{S_1}, \cdots, f_{S_k}\}$. That is, $f = \sum_{i=1}^k \theta_i f_{S_i}$ where each $\theta_i > 0$ and sum to 1. Using Lemma 14 we see that

$$\mathcal{N}_{\text{bin}}(f) = \{f_S : S \text{ is compatible with } S_1 \subsetneq \cdots \subsetneq S_k\}. \tag{23}$$

As $S_1 \subsetneq \cdots \subsetneq S_k$ is a combinatorial k-local minimum by assumption, inequality (22) holds and f is a local minimum of the energy on the TV-sphere.

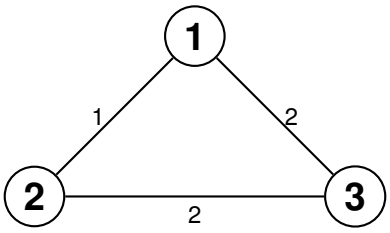
To prove the second statement of the theorem, suppose that f is a local minimum and let $f = \sum_{i=1}^k \theta_i f_{S_i}$ be its decomposition in binary functions. As the functions $f_{S_1}, \cdots, f_{S_k}$ all belong to the same face of the TV-sphere, we must have $E(f) = E(f_{S_1}) = \cdots = E(f_{S_k})$. This, in turn, implies $\mathcal{C}(S_1) = \cdots = \mathcal{C}(S_k)$. The binary neighbors of f are again defined by (23) and therefore, because of (22), we must have $E(f) \leq E(f_S)$ for all S compatible with $S_1 \subsetneq \cdots \subsetneq S_k$. This implies that $S_1 \subsetneq S_2 \subsetneq \cdots \subsetneq S_k$ is a k-local minimum of the combinatorial problem. □

5.3 Critical points

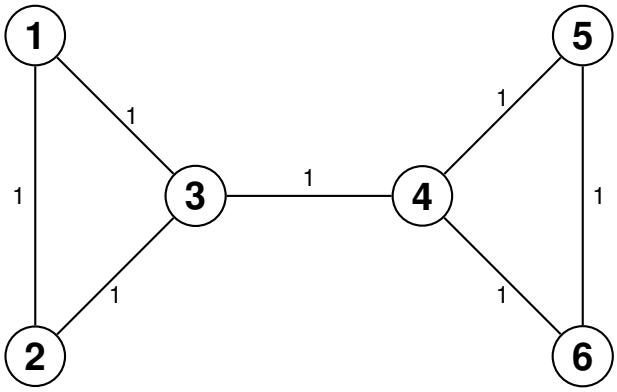
To conclude, we provide a few simple examples that illustrate the previous theorems and demonstrate the distinction between local minima and critical points (definition 5). Consider first the graph on three vertices $V = \{x_1, x_2, x_3\}$ with symmetric edge weights $(w_{12}, w_{13}, w_{23}) = (1, 2, 2)$, i.e. an isoceles triangle (see figure 5.3 (a)). To see that a continuum in $\mathcal{S}_0^2$ of global minima may occur, define $p_\alpha = (\alpha, \alpha - 1, 0)$ for $\alpha \in [0,1]$. Then $\text{med}(p_\alpha) = 0$, $||p_\alpha||_1 = 1$ and $T(p_\alpha) = 3$ for all $\alpha \in [0,1]$. Thus,

$$E(p_\alpha) = 3 = \min_{f \in \mathbb{R}^3} E(f).$$

If we then set $f_\alpha = p_\alpha / ||p_\alpha||_2 \in \mathcal{S}_0^2$, we have that $E(f_\alpha) = 3$ for all $\alpha \in [0,1]$. As each f_α attains the global minimum of E on $\mathcal{S}_0^2$, it follows that $0 \in \partial T(f_\alpha) - E(f_\alpha)\partial B(f_\alpha)$ for each $\alpha \in [0,1]$ as well. We therefore have a continuum of critical points that are also global minima. This corresponds to the fact that the sets $S_1 = \{x_1\}$ and $S_2 = \{x_1, x_3\}$ define a 2-local minimum of the combinatorial problem according to definition 7.



(a) Isoceles Triangle



(b) Bowtie

Figure 1: Small Graph Examples

We next examine the graph on six vertices $V = \{x_1, x_2, x_3, x_4, x_5, x_6\}$ that has symmetric edge weights with non-zero entries $w_{12} = w_{13} = w_{23} = w_{34} = w_{45} = w_{46} = w_{56} = 1$. We call this graph the bowtie (see figure 5.3 (b)). Consider the cut defined by the binary function $f = (1, 1, 0, 0, 0, 0)^T$ that has energy $E(f) = 1$. According to definition 5, f defines a critical point of the energy if there exist $w \in \partial T(f)$ and $v \in \partial_0 B(f)$ so that $w = v$. By taking $v = (1, 1, -1, -1, 0, 0)$ and computing the subdifferential of T explicitly, we see this occurs if there exist $s_{ij} \in [-1, 1]$ satisfying

$$\begin{pmatrix} s_{12} + 1 \\ -s_{12} + 1 \\ -2 + s_{34} \\ -s_{34} + s_{45} + s_{46} \\ -s_{45} + s_{46} \\ -s_{46} - s_{56} \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \\ 0 \\ 0 \end{pmatrix}.$$

This requires $s_{12} = 0$ and $s_{34} = 1$, which then yields a convenient choice $s_{45} = s_{46} = 0$. Thus f defines a critical point of the energy. However, direct computation shows that $f^\theta := (1, \theta, 0, 0, 0, 0)^T$ has strictly greater energy for any $0 < \theta < 1$ and $f^\kappa := (1, 1, \kappa, 0, 0, 0)^T$ has strictly lesser energy for any $0 < \kappa < 1$. Thus we have a critical point that is neither a local maximum nor a local minimum. In particular, corollary 3 does not apply even for this simple example.

6 Experiments

In all experiments, we take the constant $c = 1$ in the SD algorithm. We use the method from [3] to solve the minimization problem in the SD algorithm and the method from [7] to solve the minimization problem in the IPM algorithm. We terminate each minimization when either a stopping tolerance of $\varepsilon = 10^{-10}$ (i.e. $\|u^{j+1} - u^j\|_1 \leq \varepsilon$) or $2,000$ iterations is reached. All experiments that follow use a symmetric k-nearest neighbor graph combined with the weight similarity function $w_{i,j} = \exp(-r_{i,j}^2/\sigma^2)$. Here, $r_{i,j} = \|x_i - x_j\|_2$ and the scale parameter $\sigma^2 = 3d_k^2$, where d_k denotes the mean distance of the k^{th} nearest neighbor. We use the two-moon, MNIST, USPS and COIL datasets. The two-moon dataset [2] uses the same setting as in [13]. We take $k = 5$ nearest

neighbors to construct the graph. We preprocessed the MNIST, USPS and COIL data by projecting onto the first 50 principal components, and take $k = 10$ nearest neighbors for the MNIST and USPS datasets and $k = 5$ nearest neighbors for the COIL dataset. The first set of experiments considers the two-moon dataset and pairs of image digits extracted from the MNIST dataset. The first table summarizes the results of these tests. It shows the mean Cheeger energy value (2), the mean error of classification (% of misclassified data) and the mean computational time for both algorithms over 10 experiments with the same random initialization for both algorithms in each of the individual experiments.

	SD Algorithm			Modified IPM Algorithm [7]		
	Energy	Error (%)	Time (sec.)	Energy	Error (%)	Time (sec.)
2 moons	0.126	8.69	2.06	0.145	14.12	1.98
4's and 9's	0.115	1.65	52.4	0.185	20.53	57.3
3's and 8's	0.086	1.217	49.2	0.086	1.219	48.1

Our second set of experiments applies both algorithms to multi-class clustering problems using a standard, recursive bi-partitioning method. The table below presents the mean Cheeger energy, classification error and time over 10 experiments as before.

	SD Algorithm			Modified IPM Algorithm [7]		
	Energy	Err. (%)	Time (min.)	Energy	Err. (%)	Time (min.)
MNIST (10 classes)	1.30	11.78	45.01	1.29	11.75	42.83
USPS (10 classes)	2.37	4.11	5.15	2.37	4.13	4.81
COIL (20 classes)	0.19	1.58	4.31	0.18	2.52	4.20

Overall, the results show that both algorithms perform equivalently for both two-class and multi-class clustering problems. As our interest here lies in the theoretical properties of both algorithms, we will study practical implementation details for the SD algorithm in future work. For instance, as Hein and Bühler remark [6], solving the minimization problem for the IPM algorithm precisely is unnecessary. Analogously for the SD Algorithm, we only need to lower the energy sufficiently before proceeding to the next iteration of the algorithm. It proves convenient to stop the minimization when a weaker form of the energy inequality (6) holds, such as

$$E(f) \geq E(h) + \theta \left(\frac{E(f)}{B(h)} \frac{||\hat{h} - f||_2^2}{c} \right)$$

for some constant $0 < \theta < 1$. This condition provably holds in a finite number of iterations and still guarantees that $||f^{k+1} - f^k||_2 \to 0$. The concrete decay estimate provided by SD algorithm therefore allows us to give precise meaning to "sufficiently lowers the energy." We investigate these aspects of the algorithm and prove convergence for this practical implementation in future work.

Reproducible research: The code is available at http://www.cs.cityu.edu.hk/∼xbresson/codes.html

Acknowledgements: This work supported by AFOSR MURI grant FA9550-10-1-0569, NSF grant DMS-0902792, and Hong Kong GRF grant #110311.

A Closedness of $\mathcal{H}^c$

Define the annulus

$$K_0 = \{u \in \mathbb{R}^n : 1 \leq \|u\|_2 \leq 1 + c\sqrt{n}\} \tag{24}$$

along with the set-valued map $\mathcal{Y}^c : \mathcal{S}_0^{n-1} \rightrightarrows K_0$

$$\mathcal{Y}^c(f) := f + c\partial_0 B(f).$$

That the range of $\mathcal{Y}^c$ lies in K_0 follows from (9).

Lemma 15. *The set-valued map $\mathcal{Y}^c$ is closed.*

Proof. We first show the set-valued map $\partial_0 B : \mathcal{S}_0^{n-1} \rightrightarrows K_0$ is closed. To this end, given any

$$f^k \to f^* \quad \text{with} \quad f^k, f^* \in \mathcal{S}_0^{n-1} \tag{25}$$

$$z^k \in \partial_0 B(f^k) \quad \text{with} \quad z^k \to z^*, \tag{26}$$

we must to show that $z^* \in \partial_0 B(f^*)$. As $B(g) \geq B(f^k) + \langle z^k, g - f^k \rangle$ for all $g \in \mathbb{R}^n$ by definition, by continuity of B on $\mathcal{S}_0^{n-1}$ we have $B(g) \geq B(f^*) + \langle z^*, g - f^* \rangle$ as well. Moreover, $\langle z^*, \mathbf{1} \rangle = \lim \langle z^k, \mathbf{1} \rangle = 0$ and $z^* \in \partial_0 B(f^*)$ as desired. To show that $\mathcal{Y}^c$ is closed, assume

$$f^k \to f^* \quad \text{with} \quad f^k, f^* \in \mathcal{S}_0^{n-1} \tag{27}$$

$$g^k \in \mathcal{Y}^c(f^k) = f^k + cz^k \to g^* \tag{28}$$

for some $z^k \in \partial_0 B(f^k)$. As $\{z^k\}$ lies in a compact set and $\partial_0 B$ is closed, there exists a subsequence with $f^{k_i} \to f^*$ and $z^{k_i} \to z^* \in \partial_0 B(f^*)$. Therefore

$$g^* = \lim g^{k_i} = f^* + cz^* \in \mathcal{Y}^c(f^*)$$

by the definition of $\mathcal{Y}^c(f^*)$ as desired. □

Define the function $\Psi^c : \mathcal{S}_0^{n-1} \times K_0 \to \mathbb{R}^d$ by

$$\Psi^c(f, g) = \arg\min_u \left\{ T(u) + E(f) \frac{\|u - g\|_2^2}{2c} \right\}$$

Lemma 16. *The function Ψ^c is continuous on $\mathcal{S}_0^{n-1} \times K_0$.*

Proof. Let $h = \Psi^c(f, g)$ and $h' = \Psi^c(f', g')$. Then we have $E(f)\frac{h-g}{c} \in -\partial T(h)$ and $E(f')\frac{h'-g'}{c} \in -\partial T(h')$ so

$$T(h') \geq T(h) - \left\langle E(f)\frac{h-g}{c}, h' - h \right\rangle$$

$$T(h) \geq T(h') - \left\langle E(f')\frac{h'-g'}{c}, h - h' \right\rangle.$$

By adding these two inequalities,

$$\langle E(f)(h-g) - E(f')(h'-g'), h - h' \rangle \leq 0.$$

Adding and subtracting we get

$$\left\langle E(f)(h-g) - E(f)(h'-g'), h - h' \right\rangle + \left\langle (E(f) - E(f'))(h'-g'), h - h' \right\rangle \leq 0$$

$$E(f)\left\langle (h-h') - (g-g'), h - h' \right\rangle + (E(f) - E(f'))\left\langle h'-g', h - h' \right\rangle \leq 0$$

$$E(f)\left(\|h-h'\|_2^2 - \left\langle g-g', h-h' \right\rangle \right) + (E(f) - E(f'))\left\langle h'-g', h-h' \right\rangle \leq 0$$

$$\|h-h'\|_2^2 \leq \left\langle g-g', h-h' \right\rangle - \frac{(E(f) - E(f'))}{E(f)}\left\langle h'-g', h-h' \right\rangle$$

From Cauchy-Schwarz we have

$$\|h' - h\|_2 \leq \|g' - g\|_2 + \frac{|E(f') - E(f)|}{E(f)} \|h' - g'\|_2 \leq \|g' - g\|_2 + \frac{|E(f') - E(f)|}{E(f)} 2\|g'\|_2$$

The last inequality follows from (11). We then easily conclude that if $(f', g') \to (f, g)$ then $h' \to h$, due to the continuity of E on $\mathcal{S}_0^{n-1}$. □

Next, define the set-valued map $\mathcal{H} : \mathcal{S}_0^{n-1} \rightrightarrows \mathbb{R}^n$

$$\mathcal{H}(f) = \Psi^c(f, \mathcal{Y}^c(f)).$$

Note this definition coincides with the definition in lemma 1.

Lemma 17. *The set-valued map $\mathcal{H}$ is closed.*

Proof. Suppose that

$$f^k \to f^* \tag{29}$$

$$h^k \in \mathcal{H}(f^k) = \Psi^c(f^k, \mathcal{Y}^c(f^k)) \to h^*. \tag{30}$$

We must show that $h^* \in \mathcal{H}(f^*)$. Clearly there exist $g^k \in \mathcal{Y}^c(f^k)$ such that

$$h^k = \Psi^c(f^k, g^k).$$

As the sequence g^k is in the compact set K_0, there exists $g^* \in K_0$ and a subsequence $g^{k_i} \to g^*$. Consequently

$$f^{k_i} \to f^* \tag{31}$$

$$g^{k_i} \in \mathcal{Y}^c(f^{k_i}) \to g^*, \tag{32}$$

from which we may conclude $g^* \in \mathcal{Y}^c(f^*)$ because $\mathcal{Y}^c$ is closed. Now since Ψ^c is continuous we have

$$h^{k_i} = \Psi^c(f^{k_i}, g^{k_i}) \to \Psi^c(f^*, g^*) \in \Psi^c(f^*, \mathcal{Y}^c(f^*)) = \mathcal{H}(f^*).$$

But $h^{k_i} \to h^*$, so we may conclude $h^* \in \mathcal{H}(f^*)$ as desired. □

References

[1] X. Bresson, X.-C. Tai, T.F. Chan, and A. Szlam. Multi-Class Transductive Learning based on ℓ^1 Relaxations of Cheeger Cut and Mumford-Shah-Potts Model. *UCLA CAM Report*, 2012.

[2] T. Bühler and M. Hein. Spectral Clustering Based on the Graph p-Laplacian. In *International Conference on Machine Learning*, pages 81–88, 2009.

[3] A. Chambolle and T. Pock. A First-Order Primal-Dual Algorithm for Convex Problems with Applications to Imaging. *Journal of Mathematical Imaging and Vision*, 40(1):120–145, 2011.

[4] J. Cheeger. A Lower Bound for the Smallest Eigenvalue of the Laplacian. *Problems in Analysis*, pages 195–199, 1970.

[5] F. R. K. Chung. *Spectral Graph Theory*, volume 92 of *CBMS Regional Conference Series in Mathematics*. Published for the Conference Board of the Mathematical Sciences, Washington, DC, 1997.

[6] M. Hein and T. Bühler. An Inverse Power Method for Nonlinear Eigenproblems with Applications in 1-Spectral Clustering and Sparse PCA. In *In Advances in Neural Information Processing Systems (NIPS)*, pages 847–855, 2010.

[7] M. Hein and S. Setzer. Beyond Spectral Clustering - Tight Relaxations of Balanced Graph Cuts. In *In Advances in Neural Information Processing Systems (NIPS)*, 2011.

[8] R.R. Meyer. Sufficient conditions for the convergence of monotonic mathematical programming algorithms. *Journal of Computer and System Sciences*, 12(1):108 – 121, 1976.

[9] A. M. Ostrowski. *Solution of Equations in Euclidean and Banach Spaces*. Academic Press, New York, 1973.

[10] S. Rangapuram and M. Hein. Constrained 1-Spectral Clustering. In *International conference on Artificial Intelligence and Statistics (AISTATS)*, pages 1143–1151, 2012.

[11] J. Shi and J. Malik. Normalized Cuts and Image Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 22(8):888–905, 2000.

[12] G. Strang. Maximal Flow Through A Domain. *Mathematical Programming*, 26:123–143, 1983.

[13] A. Szlam and X. Bresson. Total variation and cheeger cuts. In *Proceedings of the 27th International Conference on Machine Learning*, pages 1039–1046, 2010.

[14] L. Zelnik-Manor and P. Perona. Self-tuning Spectral Clustering. In *In Advances in Neural Information Processing Systems (NIPS)*, 2004.