COMMUN. MATH. SCI.
Vol. 19, No. 3, pp. 761–786

LOVÁSZ EXTENSION AND GRAPH CUT*

KUNG-CHING CHANG†, SIHONG SHAO‡, DONG ZHANG§, AND WEIXI ZHANG¶

Abstract. A set-pair Lovász extension is established to construct equivalent continuous optimization problems for graph k-cut problems.

Keywords. Lovász extension; Submodular; Dual Cheeger cut; Maxcut; Graph cut.

AMS subject classifications. 90C35; 90C27; 05C85; 58E05.

1. Introduction

Motivated by the need of practical application in e.g., machine learning and big data, it is not only natural but also imperative for applied mathematicians to plug into valuable subjects that have emerged from well-established mathematics such as analytic techniques, topological views and algebraic structures. As many problems in network science and combinatorial optimization can be translated into graph partitioning problems which are usually NP-hard, well-established methods in continuous optimization should be helpful in searching for approximate solutions from a practical viewpoint. Along this direction, there have been various ways to solve combinatorial optimization problems by means of continuous optimization methods, including continuous reformulations [6,16,19] and continuous relaxations [2,12,17].

As a well-known continuous reformulation, the Lovász extension [16] provides a both explicit and equivalent continuous optimization problem for a discrete optimization problem. However, the original Lovász extension deals with set-functions which admit only one input set and thus correspond to so-called 2-cut problems, for instance, the Cheeger cut problem [5]. Therefore it is not straightforward to apply the original Lovász extension into general k-cut problems, such as the dual Cheeger cut [4,21]. Accordingly, we ask

QUESTION 1.1. *How to write down a both explicit and equivalent continuous optimization problem for a graph k-cut problem?*

Let us introduce some notations first. $G=(V,E)$ is an unweighted and undirected graph with vertex set $V=\{1,2,\cdots,n\}$ and edge set E, and w_{ij} the weight of the edge $i\sim j$. For two disjoint subsets A and B of V, let $E(A,B)$ denote the set of edges that cross A and B. For $S\subset V$, let $S^c=V\setminus S$ be the complement of S. The edge boundary of S is $\partial S=\partial S^c=E(S,S^c)$. The amount of edge set $E(A,B)$ is denoted by $|E(A,B)|=\sum_{i\in A}\sum_{j\in B}w_{ij}$, and the volume of S is defined to be $\mathrm{vol}(S)=\sum_{i\in S}d_i$, where $d_i=\sum_{j=1}^n w_{ij}$ is the degree of the vertex i.

DEFINITION 1.1 (dual Cheeger cut [4,21]). *The dual Cheeger problem is devoted to*

*Received: February 08, 2020; Accepted (in revised form): November 01, 2020. Communicated by Shi Jin.

†Laboratory of Mathematics and Its Applications (LMAM) and School of Mathematical Sciences, Peking University, Beijing 100871, China (kcchang@math.pku.edu.cn).

‡LMAM and School of Mathematical Sciences, Peking University, Beijing 100871, China (sihong@math.pku.edu.cn).

§LMAM and School of Mathematical Sciences, Peking University, Beijing 100871, China (dongzhang@pku.edu.cn).

¶LMAM and School of Mathematical Sciences, Peking University, Beijing 100871, China (zhangweixi@pku.edu.cn).

solving

$$h^+(G) = \max_{S_1\cap S_2=\varnothing, S_1\cup S_2\neq\varnothing} \frac{2|E(S_1,S_2)|}{\mathrm{vol}(S_1\cup S_2)}, \tag{1.1}$$

and we call $h^+(G)$ the dual Cheeger constant.

To say the least, before we discuss Question 1.1, the following specific question needs to be solved in the first place.

QUESTION 1.2. *Is there an explicit and equivalent continuous optimization for a graph 3-cut problem like the dual Cheeger cut (1.1)?*

The convex extension and some other continuous representations for solving integer programming [6, 19] may provide some answers to both Questions 1.1 and 1.2. In this work, we propose a set-pair Lovász extension which not only provides a complete answer to Question 1.2 (even works for a series of graph 3-cut problems), but also enlarges the feasible region of resulting equivalent continuous optimization problems from the first quadrant $\mathbb{R}^n_+\setminus\{\mathbf{0}\}$ (see Theorem 2.1) to the entire space $\mathbb{R}^n\setminus\{\mathbf{0}\}$ (see Theorem 1.3) for graph 2-cut problems like the Cheeger cut (see Theorem 1.9). This enlarged feasible region may have some advantages in designing solution algorithms. Indeed, without additional boundary constraints on $\mathbb{R}^n\setminus\{\mathbf{0}\}$, the Dinkelbach-type scheme like the inverse power method [20] can be applied directly with a good performance (see e.g. [9, 10, 13] for details).

DEFINITION 1.2 (set-pair Lovász extension). *Let $V=\{1,\ldots,n\}$. For $\boldsymbol{x}\in\mathbb{R}^n$, let $\sigma: V\cup\{0\}\to V\cup\{0\}$ be a bijection such that $|x_{\sigma(1)}|\le|x_{\sigma(2)}|\le\cdots\le|x_{\sigma(n)}|$ and $\sigma(0)=0$, where $x_0:=0$. One defines the sets*

$$V^{\pm}_{\sigma(i)} := \{j\in V : \pm x_j > |x_{\sigma(i)}|\},\quad i=0,1,\ldots,n-1. \tag{1.2}$$

Let

$$\mathcal{P}_2(V) = \{(A,B): A,B\subset V \text{ with } A\cap B=\varnothing\}. \tag{1.3}$$

Given $f:\mathcal{P}_2(V)\to[0,+\infty)$, the set-pair Lovász extension of f is a mapping from $\mathbb{R}^n$ to $\mathbb{R}$ defined by

$$f^L(\boldsymbol{x}) = \sum_{i=0}^{n-1}(|x_{\sigma(i+1)}|-|x_{\sigma(i)}|) f(V^+_{\sigma(i)}, V^-_{\sigma(i)}). \tag{1.4}$$

THEOREM 1.3. *Assume that $f,g:\mathcal{P}_2(V)\to[0,+\infty)$ are two set-pair functions with $g(A,B)>0$ whenever $A\cup B\neq\varnothing$. Then there hold both*

$$\min_{(A,B)\in\mathcal{P}_2(V)\setminus\{(\varnothing,\varnothing)\}} \frac{f(A,B)}{g(A,B)} = \inf_{\boldsymbol{x}\neq\mathbf{0}} \frac{f^L(\boldsymbol{x})}{g^L(\boldsymbol{x})}, \tag{1.5}$$

and

$$\max_{(A,B)\in\mathcal{P}_2(V)\setminus\{(\varnothing,\varnothing)\}} \frac{f(A,B)}{g(A,B)} = \sup_{\boldsymbol{x}\neq\mathbf{0}} \frac{f^L(\boldsymbol{x})}{g^L(\boldsymbol{x})}. \tag{1.6}$$

A similar deduction to that for Theorem 1.3 leads to

PROPOSITION 1.1. *Assume that $f,g:\mathcal{P}_2(V)\to[0,+\infty)$ are two set-pair functions satisfying $f(\varnothing,V)=f(V,\varnothing)=0$ and $g(A,B)>0$ whenever $(A,B)\notin\{(\varnothing,\varnothing),(\varnothing,V),(V,\varnothing)\}$, then there hold both*

$$\min_{(A,B)\in\mathcal{P}_2(V)\setminus\{(\varnothing,\varnothing),(\varnothing,V),(V,\varnothing)\}}\frac{f(A,B)}{g(A,B)}=\inf_{\boldsymbol{x}\ \text{nonconstant}}\frac{f^L(\boldsymbol{x})}{g^L(\boldsymbol{x})},\tag{1.7}$$

and

$$\max_{(A,B)\in\mathcal{P}_2(V)\setminus\{(\varnothing,\varnothing),(\varnothing,V),(V,\varnothing)\}}\frac{f(A,B)}{g(A,B)}=\sup_{\boldsymbol{x}\ \text{nonconstant}}\frac{f^L(\boldsymbol{x})}{g^L(\boldsymbol{x})}.\tag{1.8}$$

Theorem 1.3, Proposition 1.1 and their applications listed below show a natural answer to Question 1.2.

THEOREM 1.4.

$$1-h^+(G)=\inf_{\boldsymbol{x}\neq\mathbf{0}}\frac{I^+(\boldsymbol{x})}{\|\boldsymbol{x}\|},\tag{1.9}$$

where

$$\|\boldsymbol{x}\|=\sum_{i=1}^n d_i|x_i|,\tag{1.10}$$

$$I^+(\boldsymbol{x})=\sum_{i<j}w_{ij}|x_i+x_j|.\tag{1.11}$$

DEFINITION 1.3 (max 3-cut [11]). *The max 3-cut problem is to determine a graph 3-cut by solving*

$$h_{\max,3}(G)=\max_{A,B,C}\frac{2(|E(A,B)|+|E(B,C)|+|E(C,A)|)}{\mathrm{vol}(V)},\tag{1.12}$$

and the associate (A,B,C) is called a max 3-cut, where the subsets A,B,C satisfy $A\cap B=B\cap C=C\cap A=\varnothing$ and $A\cup B\cup C=V$.

THEOREM 1.5.

$$h_{\max,3}(G)=\sup_{\boldsymbol{x}\neq\mathbf{0}}\frac{I(\boldsymbol{x})+\hat{I}(\boldsymbol{x})}{\mathrm{vol}(V)\|\boldsymbol{x}\|_\infty},\tag{1.13}$$

where

$$I(\boldsymbol{x})=\sum_{i<j}w_{ij}|x_i-x_j|,\tag{1.14}$$

$$\hat{I}(\boldsymbol{x})=\sum_{i<j}w_{ij}\,||x_i|-|x_j||,\tag{1.15}$$

$$\|\boldsymbol{x}\|_\infty=\max\{|x_1|,\dots,|x_n|\}.\tag{1.16}$$

DEFINITION 1.4 (ratio max 3-cut I). *The first ratio max 3-cut problem is to determine a graph 3-cut by solving*

$$h_{\max,3,I}(G)=\max_{A,B,C}\frac{2(|E(A,B)|+|E(B,C)|+|E(C,A)|)}{\mathrm{vol}(A)+\mathrm{vol}(B)},\tag{1.17}$$

where $A\cap B=B\cap C=C\cap A=\varnothing$ *and* $A\cup B\cup C=V$.

THEOREM 1.6.

$$1-h_{\max,3,I}(G)=\inf_{\boldsymbol{x}\neq\mathbf{0}}\frac{I^+(\boldsymbol{x})-2\hat{I}(\boldsymbol{x})}{\|\boldsymbol{x}\|}.$$

DEFINITION 1.5 (ratio max 3-cut II). *The second ratio max 3-cut problem is to determine a graph 3-cut by solving*

$$h_{\max,3,II}(G)=\max_{A,B,C}\frac{2(|E(A,B)|+|E(B,C)|+|E(C,A)|)}{\max\{\mathrm{vol}(A\cup B),\mathrm{vol}(C)\}},\tag{1.18}$$

where $A\cap B=B\cap C=C\cap A=\varnothing$ *and* $A\cup B\cup C=V$.

THEOREM 1.7.

$$h_{\max,3,II}(G)=\sup_{\boldsymbol{x}\neq\mathbf{0}}\frac{2I(\boldsymbol{x})-\|\boldsymbol{x}\|+I^+(\boldsymbol{x})}{\mathrm{vol}(V)\|\boldsymbol{x}\|_\infty-\min\limits_{\alpha\in\mathbb{R}}\sum_{i=1}^n d_i\,||x_i|-\alpha|}.\tag{1.19}$$

In order to give a complete answer to Question 1.1, we propose an isomorphism to translate a k-cut ($k>3$) problem to a 3-cut one on a graph of larger size, and then still utilize Theorem 1.3 to derive the corresponding continuous problem. We refer to Section 3.2 for this method on establishing the equivalent continuous optimization for graph k-cut problems, by which the (relaxed) Dinkelbach iteration, as we have discussed, also applies. On the other hand, Theorem 1.3 also works for graph 2-cut problems, although it produces a different form from that by the original Lovaśz extension, during which Lemma 3.1 plays a key role and translates a graph 2-cut with symmetric form into a graph 3-cut. The main difference lies in the feasible region.

DEFINITION 1.6 (maxcut [15]). *The maxcut problem is to determine a graph cut by solving*

$$h_{\max}(G)=\max_{S\subset V}\frac{2|\partial S|}{\mathrm{vol}(V)}.\tag{1.20}$$

DEFINITION 1.7 (Cheeger cut [5]). *The Cheeger problem is to determine a graph cut by solving*

$$h(G)=\min_{S\subset V,S\notin\{\varnothing,V\}}\frac{|\partial S|}{\min\{\mathrm{vol}(S),\mathrm{vol}(S^c)\}}.$$

DEFINITION 1.8 (anti-Cheeger cut [22]). The anti-Cheeger constant $h_{\mathrm{anti}}(G)$ is defined as

$$h_{\mathrm{anti}}(G)=\max_{S\subset V}\frac{|\partial S|}{\max\{\mathrm{vol}(S),\mathrm{vol}(S^c)\}}.\tag{1.21}$$

The original Lovász extension (2.2) (vide post) yields the following equivalent continuous optimization problems:

$$h_{\max}(G) = \sup_{\boldsymbol{x}\in\mathbb{R}^n_+\setminus\{\mathbf{0}\}} \frac{I(\boldsymbol{x})}{\mathrm{vol}(V)\max\limits_i x_i}, \tag{1.22}$$

$$h(G) = \inf_{\boldsymbol{x} \text{ nonconstant in } \mathbb{R}^n_+} \sup_{c\in\mathbb{R}} \frac{I(\boldsymbol{x})}{\sum_{i=1}^n d_i|x_i - c|}, \tag{1.23}$$

$$h_{\mathrm{anti}}(G) = \sup_{\boldsymbol{x}\in\mathbb{R}^n_+\setminus\{\mathbf{0}\}} \frac{I(\boldsymbol{x})}{2\mathrm{vol}(V)\max\limits_i x_i - \min\limits_{c\in\mathbb{R}}\sum_{i=1}^n d_i|x_i-c|}. \tag{1.24}$$

In contrast, the proposed set-pair Lovász extension is capable of enlarging the feasible region from the first quadrant $\mathbb{R}^n_+\setminus\{\mathbf{0}\}$ in Equations (1.22)-(1.24) to the entire space $\mathbb{R}^n\setminus\{\mathbf{0}\}$ in Equations (1.25)-(1.27).

THEOREM 1.8.

$$h_{\max}(G) = \sup_{\boldsymbol{x}\neq\mathbf{0}} \frac{I(\boldsymbol{x})}{\mathrm{vol}(V)\|\boldsymbol{x}\|_\infty}. \tag{1.25}$$

THEOREM 1.9.

$$h(G) = \inf_{\boldsymbol{x} \text{ nonconstant}} \sup_{c\in\mathbb{R}} \frac{I(\boldsymbol{x})}{\sum_{i=1}^n d_i|x_i-c|}. \tag{1.26}$$

Using the fact that $\boldsymbol{y}\in\mathrm{median}(\boldsymbol{x})$ if and only if $\boldsymbol{y} = \arg\min_{c\in\mathbb{R}}\sum_{i=1}^n d_i|x_i-c|$ (see Lemma 2.3 in [8]), we are able to rewrite Equation (1.26) as $h(G) = \min_{\boldsymbol{x}\in\pi} I(\boldsymbol{x})$ with which Chang proved that the Cheeger constant happens to be the second eigenvalue of the graph 1-Laplacian [7], where $\pi = \{\boldsymbol{x}\in\mathbb{R}^n : 0\in\mathrm{median}(\boldsymbol{x}), \|\boldsymbol{x}\| = 1\}$. That is, the set-pair Lovász extension of the Cheeger cut problem produces a continuous reformulation corresponding to the graph 1-Laplacian.

THEOREM 1.10.

$$h_{\mathrm{anti}}(G) = \sup_{\boldsymbol{x}\neq\mathbf{0}} \frac{I(\boldsymbol{x})}{2\mathrm{vol}(V)\|\boldsymbol{x}\|_\infty - \min\limits_{\alpha\in\mathbb{R}}\|\boldsymbol{x}-\alpha\mathbf{1}\|}. \tag{1.27}$$

REMARK 1.1. Comparing (1.13) to (1.25), the continuous objective function for max 3-cut happens to only add a nonnegative term $\hat{I}(\boldsymbol{x})$ to the numerator. Such slight formal discrepancy may imply some deep connections between maxcut and max 3-cut which deserves more efforts to explore.

The fractional form of the equivalent continuous optimizations in Equations (1.5) and (1.6) implies that we can directly adopt the Dinkelbach iteration in Fractional Programming [20] to solve them. Moreover, both the numerators and denominators could be rewritten as the differences of two convex functions (see e.g. Equations (1.13) and (1.27)), implying that a simple and efficient algorithm can be obtained via further relaxing the Dinkelbach iteration by techniques in DC programming [14]. We refer the interested readers to [8, 9] and [10] for our preliminary attempts on the Cheeger cut, dual Cheeger cut and maxcut problems, respectively. It is noteworthy that a simple iterative algorithm based on the continuous reformulation by the set-pair Lovász extension provides the best cut values for maxcut on G-set among all existing continuous algorithms [10].

The rest of the paper is organized as follows. Section 2 collects basic properties of the Lovász extension including both continuity and convexity, and shows that the set-pair Lovász extension may be superior over the original one. Such superiority is further demonstrated in Section 3 by applying it into typical graph k-cut problems.

2. Set-pair Lovász extension

DEFINITION 2.1 (Lovász extension [16]). Let $V=\{1,\ldots,n\}\subset\mathbb{N}$. For $\boldsymbol{x}\in\mathbb{R}^n$, let $\sigma: V\cup\{0\}\to V\cup\{0\}$ be a bijection such that $x_{\sigma(1)}\le x_{\sigma(2)}\le\cdots\le x_{\sigma(n)}$ and $\sigma(0)=0$, where $x_0:=0$. One defines the sets

$$V_{\sigma(i)}:=\{j\in V: x_j>x_{\sigma(i)}\},\ \ i=1,\ldots,n-1, V_0=V.$$

Let

$$\mathcal{P}(V)=\{A: A\subset V\}. \tag{2.1}$$

Given $f:\mathcal{P}(V)\to[0,+\infty)$, the Lovász extension of f is a mapping from $\mathbb{R}^n$ to $\mathbb{R}$ defined by

$$f_o^L(\boldsymbol{x})=\sum_{i=0}^{n-1}(x_{\sigma(i+1)}-x_{\sigma(i)})f(V_{\sigma(i)}). \tag{2.2}$$

THEOREM 2.1 ([5], Theorem 1). *Assume that $f,g:\mathcal{P}(V)\to[0,+\infty)$ are two functions with $g(A)>0$ whenever $A\neq\varnothing$, then there hold both*

$$\min_{A\in\mathcal{P}(V)\setminus\{\varnothing\}}\frac{f(A)}{g(A)}=\inf_{\boldsymbol{x}\in\mathbb{R}_+^n\setminus\{\boldsymbol{0}\}}\frac{f_o^L(\boldsymbol{x})}{g_o^L(\boldsymbol{x})}, \tag{2.3}$$

and

$$\max_{A\in\mathcal{P}(V)\setminus\{\varnothing\}}\frac{f(A)}{g(A)}=\sup_{\boldsymbol{x}\in\mathbb{R}_+^n\setminus\{\boldsymbol{0}\}}\frac{f_o^L(\boldsymbol{x})}{g_o^L(\boldsymbol{x})}. \tag{2.4}$$

REMARK 2.1. The proof of Theorem 2.1 (see [5]) heavily depends on the non-negativity of the terms $x_{\sigma(1)}f(V_{\sigma(1)})$ and $(x_{\sigma(i+1)}-x_{\sigma(i)})f(V_{\sigma(i)})$ in the summation form (2.2), $i=1,\ldots,n-1$, thereby indicating that one needs the constraint $x_{\sigma(1)}\ge 0$, i.e., $\boldsymbol{x}\in\mathbb{R}_+^n$. The integral form (2.5) in Proposition 2.1 also manifests clearly such dependence through the last term. Indeed, the minor change of Equation (1.22):

$$\begin{aligned}-\infty=\inf_{\boldsymbol{x}\neq\boldsymbol{0}}\frac{I(\boldsymbol{x})}{\mathrm{vol}(V)\max\limits_i x_i}&<\min_{S\subset V}\frac{2|\partial S|}{\mathrm{vol}(V)}\\ &\le\max_{S\subset V}\frac{2|\partial S|}{\mathrm{vol}(V)}<\sup_{\boldsymbol{x}\neq\boldsymbol{0}}\frac{I(\boldsymbol{x})}{\mathrm{vol}(V)\max\limits_i x_i}=+\infty,\end{aligned}$$

shows an example in which Theorem 2.1 fails if we naively replace $\mathbb{R}_+^n\setminus\{\boldsymbol{0}\}$ by $\mathbb{R}^n\setminus\{\boldsymbol{0}\}$ in Equations (2.3) and (2.4). Fortunately, as the fruitful results and discussions in this section, we can enlarge the feasible region $\mathbb{R}_+^n\setminus\{\boldsymbol{0}\}$ to $\mathbb{R}^n\setminus\{\boldsymbol{0}\}$ using the proposed set-pair analog of Lovász extension.

PROPOSITION 2.1 ([1], Definition 3.1).

$$f_o^L(\boldsymbol{x})=\int_{\min_{1\le i\le n}x_i}^{\max_{1\le i\le n}x_i} f(V_t)dt+f(V)\min_{1\le i\le n} x_i, \tag{2.5}$$

where $V_t(\boldsymbol{x})=\{i\in V: x_i>t\}$.

DEFINITION 2.2. *A set-function* $f:\mathcal{P}(V)\to\mathbb{R}$ *is symmetric if* $f(A)=f(A^c)$ *for any subset* $A\subset V$. *A set-pair-function* $f:\mathcal{P}_2(V)\to\mathbb{R}$ *is symmetric if* $f(A,B)=f(B,A)$ *for any* $(A,B)\in\mathcal{P}_2(V)$.

PROPOSITION 2.2 ([1], Proposition 3.1). *For* $f_o^L(\boldsymbol{x})$ *by the Lovász extension, we have*

(1) $f_o^L(\boldsymbol{x}+\alpha\mathbf{1})=f_o^L(\boldsymbol{x})+\alpha f(V)$ *for any* $\alpha\in\mathbb{R}$.
(2) $f_o^L(\boldsymbol{x})$ *is one-homogeneous.*
(3) $(f+g)_o^L=f_o^L+g_o^L$, $(\lambda f)_o^L=\lambda f_o^L$, $\forall\lambda\ge 0$.
(4) $f_o^L(\boldsymbol{x})$ *is even if and only if* f *is symmetric.*

For $A\subset V$, $\mathbf{1}_A$ is the characteristic function of A. For the set-pair case, we denote

$$\mathbf{1}_{A,B}=\mathbf{1}_A-\mathbf{1}_B,\ \forall (A,B)\in\mathcal{P}_2(V). \tag{2.6}$$

Accordingly, for any set-pair function $f:\mathcal{P}_2(V)\to[0,+\infty)$, the following fact can be readily verified by Definition 1.2

$$f^L(\mathbf{1}_{A,B})=f(A,B),\ \forall (A,B)\in\mathcal{P}_2(V)\setminus\{(\varnothing,\varnothing)\}. \tag{2.7}$$

Particularly, the above equality is always true for any $(A,B)\in\mathcal{P}_2(V)$ if $f(\varnothing,\varnothing)=0$.

Similarly, we can derive an integral form of the set-pair Lovász extension.

PROPOSITION 2.3.

$$f^L(\boldsymbol{x})=\int_0^{\|x\|_\infty} f(V_t^+(\boldsymbol{x}),V_t^-(\boldsymbol{x}))dt, \tag{2.8}$$

where $V_t^{\pm}(\boldsymbol{x})=\{i\in V:\pm x_i>t\}$.

Proof. Let σ be a permutation defined in Definition 1.2. It is easy to check that if $|x_{\sigma(i)}|\le t<|x_{\sigma(i+1)}|$ then

$$V_t^{\pm}(\boldsymbol{x})=\{i\in V:\pm x_i>t\}=V_{\sigma(i)}^{\pm}.$$

Therefore,

$$\begin{aligned} f^L(\boldsymbol{x})&=\sum_{i=0}^{n-1}(|x_{\sigma(i+1)}|-|x_{\sigma(i)}|)f(V_{\sigma(i)}^+,V_{\sigma(i)}^-)\\ &=\sum_{i=0}^{n-1}\int_{|x_{\sigma(i)}|}^{|x_{\sigma(i+1)}|} f(V_t^+,V_t^-)dt\\ &=\int_0^{\|x\|_\infty} f(V_t^+,V_t^-)dt. \end{aligned}$$

∎

For simplicity, we denote by $(A,B)\subset(C,D)$ if $A\subset C$ and $B\subset D$ for $(A,B),(C,D)\in\mathcal{P}_2(V)$.

Proposition 2.4.

$$f^L(\boldsymbol{x})=\sum_{i=1}^{p}\lambda_i f(V_i^+,V_i^-) \tag{2.9}$$

whenever $(V_p^+,V_p^-)\subset\cdots\subset(V_0^+,V_0^-)$ $(p\in\mathbb{N}^+)$ *is a chain satisfying* $V_0^+\cap V_0^-=\varnothing$, $\sum_{i=0}^{p}\lambda_i\mathbf{1}_{V_i^+,V_i^-}=\boldsymbol{x}$ *and* $\sum_{i=0}^{p}\lambda_i=\|\boldsymbol{x}\|_\infty$, $\lambda_i\geq 0$.

Proof. Setting $t_i=\sum_{j=0}^{i-1}\lambda_j$, $i=0,\ldots,p$, $t_{p+1}=\|\boldsymbol{x}\|_\infty$. Now we verify

$$\lambda_i f(V_i^+,V_i^-)=\int_{t_{i-1}}^{t_i} f(V_t^+(\boldsymbol{x}),V_t^-(\boldsymbol{x}))dt,\quad i=0,\ldots,p, \tag{2.10}$$

where $V_t^\pm(\boldsymbol{x})=\{i\in V\,|\,\pm x_i>t\}$. It obviously holds for the case of $\lambda_i=0$, so we can assume that $\lambda_i\neq 0$. Since

$$\sum_{i=0}^{p}\lambda_i\mathbf{1}_{V_i^+,V_i^-}=\boldsymbol{x} \tag{2.11}$$

and

$$(V_p^+,V_p^-)\subset\cdots\subset(V_0^+,V_0^-),$$

we can assume that $j\in V_i^\pm$ and thus $j\notin V_i^\mp$ for any $0\leq i\leq p$. Consider the j-th component of (2.11) on both sides,

$$\pm\sum_{V_i^\pm\ni j}\lambda_i=\pm\sum_{i=0}^{p_j}\lambda_i=\pm t_{p_j+1}=x_j, \tag{2.12}$$

where p_j is the largest integer such that $j\in V_{p_j}^\pm$. Then

$$V_t^\pm(\boldsymbol{x})=V_i^\pm,\ \forall t\in[t_{i-1},t_i),$$

and

$$\int_{t_{i-1}}^{t_i} f(V_t^+(\boldsymbol{x}),V_t^-(\boldsymbol{x}))dt=\int_{t_{i-1}}^{t_i} f(V_i^+,V_i^-)dt=\lambda_i f(V_i^+,V_i^-).$$

Therefore

$$\begin{aligned}\sum_{i=0}^{p}\lambda_i f(V_i^+,V_i^-)&=\sum_{i=0}^{p}\int_{t_{i-1}}^{t_i} f(V_t^+(\boldsymbol{x}),V_t^-(\boldsymbol{x}))dt\\ &=\int_0^{t_p} f(V_t^+(\boldsymbol{x}),V_t^-(\boldsymbol{x}))dt=f^L(\boldsymbol{x}).\end{aligned}$$

In particular, let $p=n-1$, $V_i^\pm=V_{\sigma(i)}^\pm$ and $\lambda_i=|x_{\sigma(i+1)}|-|x_{\sigma(i)}|$, $i=0,1,\ldots,n-1$, then (2.9) returns to (1.4). ∎

Remark 2.2. A more detailed check of the proof of Proposition 2.4 shows that, for given $\boldsymbol{x}\neq\mathbf{0}$, if we assume every $\lambda_i>0$, then the chain $(V_p^+,V_p^-)\subset\cdots\subset(V_0^+,V_0^-)$ $(p\in\mathbb{N}^+)$ and $\{\lambda_i\}_{i=0}^p$ in Proposition 2.4 are uniquely determined by $\boldsymbol{x}$ and thus independent of f.

That is, if there are two chains $(V_p^+,V_p^-)\subset\cdots\subset(V_0^+,V_0^-)$ $(p\in\mathbb{N}^+)$, and $(\widetilde{V}_q^+,\widetilde{V}_q^-)\subset\cdots\subset(\widetilde{V}_0^+,\widetilde{V}_0^-)$ $(q\in\mathbb{N}^+)$, as well as two sequences of positive numbers $\{\lambda_i\}_{i=0}^p$ and $\{\widetilde{\lambda}_i\}_{i=0}^q$, such that $V_0^+\cap V_0^-=\widetilde{V}_0^+\cap\widetilde{V}_0^-=\varnothing$, $\sum_{i=0}^p\lambda_i\mathbf{1}_{V_i^+,V_i^-}=\sum_{i=0}^q\widetilde{\lambda}_i\mathbf{1}_{\widetilde{V}_i^+,\widetilde{V}_i^-}=\boldsymbol{x}$ and $\sum_{i=0}^p\lambda_i=\sum_{i=0}^q\widetilde{\lambda}_i=\|\boldsymbol{x}\|_\infty$, then $q=p$, $(\widetilde{V}_i^+,\widetilde{V}_i^-)=(V_i^+,V_i^-)$ and $\widetilde{\lambda}_i=\lambda_i$, $i=0,\ldots,p$.

Propositions 2.3 and 2.4 provide repectively the integral (continuous) form and chain (combinatorial) form of the set-pair Lovász extension, both of which are very helpful.

Remark 2.3. After the submission of this paper, we became aware of the submodular analysis involving the set-pair Lovasz extension investigated by Qi [18], in which the chain form (Proposition 2.4) was mentioned. But our integral form (2.8) (which didn't appear before the present paper) is much more convenient to obtain a closed formula of the equivalent continuous optimization for a graph cut problem. Moreover, Qi's paper and this paper focus on different aspects except the submodularity theorem (Theorem 2.4), and his proof of Theorem 2.4 is very different from ours.

The set-pair version of Proposition 2.2 reads as follows:

Proposition 2.5. *For $f^L(\boldsymbol{x})$ by the set-pair Lovász extension, we have*

(1) $f^L(\boldsymbol{x}+\alpha\operatorname{sign}(\boldsymbol{x}))=f^L(\boldsymbol{x})+\alpha f(V_0^+,V_0^-)$ *for any* $\alpha\ge 0$.

(2) $f^L(\boldsymbol{x})$ *is one-homogeneous.*

(3) $(f+g)^L=f^L+g^L$, $(\lambda f)^L=\lambda f^L$, $\forall\lambda\ge 0$.

(4) $f^L(\boldsymbol{x})$ *is even if and only if* f *is symmetric.*

Proof. We will give the proof in turn.

(1) Let $\tilde{\boldsymbol{x}}=\boldsymbol{x}+\alpha\operatorname{sign}(\boldsymbol{x})$. Then $\|\tilde{\boldsymbol{x}}\|_\infty=\|\boldsymbol{x}\|_\infty+\alpha$ and

$$\begin{cases} V_t^\pm(\tilde{\boldsymbol{x}})=V_{t-\alpha}^\pm(\boldsymbol{x}), & \text{if } t\ge\alpha,\\ V_t^\pm(\tilde{\boldsymbol{x}})=V_0^\pm(\boldsymbol{x}), & \text{if } t\in[0,\alpha).\end{cases}$$

According to Proposition 2.3, we have

$$\begin{aligned} f^L(\tilde{\boldsymbol{x}})&=\int_0^{\|\tilde{\boldsymbol{x}}\|_\infty}f(V_t^+(\tilde{\boldsymbol{x}}),V_t^-(\tilde{\boldsymbol{x}}))dt\\ &=\int_0^\alpha f(V_0^+,V_0^-)dt+\int_\alpha^{\|x\|_\infty+\alpha}f(V_t^+(\tilde{\boldsymbol{x}}),V_t^-(\tilde{\boldsymbol{x}}))dt\\ &=\alpha f(V_0^+,V_0^-)+\int_0^{\|x\|_\infty}f(V_t^+(\boldsymbol{x}),V_t^-(\boldsymbol{x}))dt\\ &=\alpha f(V_0^+,V_0^-)+f^L(\boldsymbol{x}).\end{aligned}$$

(2) For any $\lambda>0$, we have

$$\begin{aligned} f^L(\lambda\boldsymbol{x})&=\int_0^{\lambda\|x\|_\infty}f(V_t^+(\lambda\boldsymbol{x}),V_t^-(\lambda\boldsymbol{x}))dt\\ &=\int_0^{\|x\|_\infty}\lambda f(V_s^+(\boldsymbol{x}),V_s^-(\boldsymbol{x}))ds=\lambda f^L(\boldsymbol{x}).\end{aligned}$$

(3) It can be obtained directly by the linearity of integral operators.

(4) On one hand, if $f^L(\boldsymbol{x})$ is even, then for any $A,B\in\mathcal{P}_2(V)$, there holds

$$f(A,B)=f^L(\mathbf{1}_{A,B})=f^L(-\mathbf{1}_{A,B})=f^L(\mathbf{1}_{B,A})=f(B,A),$$

due to (2.7). Hence $f(A,B)$ is symmetric.
On the other hand, if $f(A,B)$ is symmetric, then using Proposition 2.3 leads to

$$\begin{aligned} f^L(\boldsymbol{x}) &= \int_0^{\|x\|_\infty} f(V_t^+(\boldsymbol{x}),V_t^-(\boldsymbol{x}))dt \\ &= \int_0^{\|x\|_\infty} f(V_t^-(\boldsymbol{x}),V_t^+(\boldsymbol{x}))dt \\ &= \int_0^{\|x\|_\infty} f(V_t^+(-\boldsymbol{x}),V_t^-(-\boldsymbol{x}))dt = f^L(-\boldsymbol{x}), \end{aligned}$$

i.e., $f^L(\boldsymbol{x})$ is even. ∎

Now we are in the position to give the proof of Theorem 1.3.

Proof. **(Proof of Theorem 1.3.)** On one hand, for any $(A,B)\in\mathcal{P}_2(V)\setminus\{(\varnothing,\varnothing)\}$, we have $f(A,B)=f^L(\mathbf{1}_{A,B})$ and $g(A,B)=g^L(\mathbf{1}_{A,B})$ due to (2.7), and then

$$\min_{(A,B)\in\mathcal{P}_2(V)\setminus\{(\varnothing,\varnothing)\}}\frac{f(A,B)}{g(A,B)} = \min_{(A,B)\in\mathcal{P}_2(V)\setminus\{(\varnothing,\varnothing)\}}\frac{f^L(\mathbf{1}_{A,B})}{g^L(\mathbf{1}_{A,B})} \ge \inf_{\boldsymbol{x}\neq\mathbf{0}}\frac{f^L(\boldsymbol{x})}{g^L(\boldsymbol{x})}. \tag{2.13}$$

On the other hand, for any $\boldsymbol{x}\neq\mathbf{0}$, we have

$$\frac{f^L(\boldsymbol{x})}{g^L(\boldsymbol{x})} = \frac{\sum_{i=0}^{n-1}(|x_{\sigma(i+1)}|-|x_{\sigma(i)}|)f(V^+_{\sigma(i)},V^-_{\sigma(i)})}{\sum_{i=0}^{n-1}(|x_{\sigma(i+1)}|-|x_{\sigma(i)}|)g(V^+_{\sigma(i)},V^-_{\sigma(i)})}.$$

Let $(C,D)\in\{(V^+_{\sigma(i)},V^-_{\sigma(i)})|0\le i\le n-1\}$ such that

$$\frac{f(C,D)}{g(C,D)} = \min_{0\le i\le n-1}\frac{f(V^+_{\sigma(i)},V^-_{\sigma(i)})}{g(V^+_{\sigma(i)},V^-_{\sigma(i)})},$$

and thus

$$\Pi_i := g(V^+_{\sigma(i)},V^-_{\sigma(i)})\left(\frac{f(V^+_{\sigma(i)},V^-_{\sigma(i)})}{g(V^+_{\sigma(i)},V^-_{\sigma(i)})} - \frac{f(C,D)}{g(C,D)}\right) \ge 0$$

holds for any $0\le i\le n-1$. Accordingly, we have

$$\frac{f^L(\boldsymbol{x})}{g^L(\boldsymbol{x})} - \frac{f(C,D)}{g(C,D)} = \frac{\sum_{i=0}^{n-1}(|x_{\sigma(i+1)}|-|x_{\sigma(i)}|)\Pi_i}{\sum_{i=0}^{n-1}(|x_{\sigma(i+1)}|-|x_{\sigma(i)}|)g(V^+_{\sigma(i)},V^-_{\sigma(i)})} \ge 0,$$

which directly implies

$$\min_{(A,B)\in\mathcal{P}_2(V)\setminus\{(\varnothing,\varnothing)\}}\frac{f(A,B)}{g(A,B)} = \frac{f^L(\mathbf{1}_{C,D})}{g^L(\mathbf{1}_{C,D})} \le \inf_{\boldsymbol{x}\neq\mathbf{0}}\frac{f^L(\boldsymbol{x})}{g^L(\boldsymbol{x})}. \tag{2.14}$$

Combining (2.13) and (2.14) finally yields

$$\min_{(A,B)\in\mathcal{P}_2(V)\setminus\{(\varnothing,\varnothing)\}}\frac{f(A,B)}{g(A,B)} = \inf_{\boldsymbol{x}\neq\mathbf{0}}\frac{f^L(\boldsymbol{x})}{g^L(\boldsymbol{x})}.$$

The proof for the maximum problem (1.6) is similar and thus skipped. ∎

We omit the proof of Proposition 1.1 because it is similar to that for Theorem 1.3. Next, we study the continuity of f^L.

THEOREM 2.2. *f^L is a Lipschitz continuous piecewise linear function.*

Proof. For a mapping $m:\{1,2,\dots,n\}\to\{-1,1\}$ and a permutation σ of $\{1,2,\dots,n\}$, one defines a closed convex cone as follows

$$\triangle_{m,\sigma} := \left\{\boldsymbol{x}\in\mathbb{R}^n : |x_{\sigma(1)}|\le\cdots\le|x_{\sigma(n)}| \text{ with } x_{\sigma(i)}m(i)\ge 0\right\},$$

and it can be readily seen that $\mathbb{R}^n=\bigcup_{m,\sigma}\triangle_{m,\sigma}$.

It suffices to prove that f^L is linear and Lipschitz continuous with a Lipschitz constant $2\max_{(A,B)\in\mathcal{P}_2(V)} f(A,B)$ on each $\triangle_{m,\sigma}$. In fact, for given m and σ and any $\boldsymbol{x}\in\triangle_{m,\sigma}$, we have

$$\begin{aligned} f^L(\boldsymbol{x}) &= \sum_{i=0}^{n-1}(m(i+1)x_{\sigma(i+1)}-m(i)x_{\sigma(i)})f(V^+_{x_{\sigma(i)}},V^-_{x_{\sigma(i)}}) \\ &= \sum_{i=1}^{n-1}x_{\sigma(i)}m(i)(f(V^+_{x_{\sigma(i-1)}},V^-_{x_{\sigma(i-1)}})-f(V^+_{x_{\sigma(i)}},V^-_{x_{\sigma(i)}})) \\ &\quad + x_{\sigma(n)}m(n)f(V^+_{x_{\sigma(n-1)}},V^-_{x_{\sigma(n-1)}}). \end{aligned}$$

Since $m(i)$ and $f(V^+_{x_{\sigma(i)}},V^-_{x_{\sigma(i)}})$ are constants for given m and σ, f^L is linear on $\triangle_{m,\sigma}$. Moreover, for any $\boldsymbol{x},\boldsymbol{y}\in\triangle_{m,\sigma}$,

$$\begin{aligned} |f^L(\boldsymbol{x})-f^L(\boldsymbol{y})| &\le \sum_{i=1}^{n-1}|x_{\sigma(i)}-y_{\sigma(i)}||f(V^+_{x_{\sigma(i-1)}},V^-_{x_{\sigma(i-1)}})-f(V^+_{x_{\sigma(i)}},V^-_{x_{\sigma(i)}})| \\ &\quad +|x_{\sigma(n)}-y_{\sigma(n)}|f(V^+_{x_{\sigma(n-1)}},V^-_{x_{\sigma(n-1)}}) \\ &\le 2\max_{(A,B)\in\mathcal{P}_2(V)} f(A,B)\|\boldsymbol{x}-\boldsymbol{y}\|_1. \end{aligned}$$

∎

The concept of submodular function was introduced by Lovász to characterize the convexity of its Lovász extension [16].

DEFINITION 2.3 (submodular function [16]). *A set-function $f:\mathcal{P}(V)\to\mathbb{R}$ is submodular if and only if, for all subsets $A,B\subset V$,*

$$f(A)+f(B)\ge f(A\cup B)+f(A\cap B).$$

THEOREM 2.3 ([16], Proposition 4.1). *f^L_o is convex if and only if f is submodular.*

Theorem 2.3 inspires us to consider the set-pair form of submodular function. A kind of set-pair submodular function was proposed in [3].

DEFINITION 2.4 (set-pair submodular function [3]). *Let*

$$\mathcal{P}_2'(V)=\{(X_I,X_O):X_I\subset X_O\subset V\}.$$

A function $p:\mathcal{P}_2'(V)\to\mathbb{R}$ is submodular if

$$p(X_I,X_O)+p(Y_I,Y_O)\ge p(X_I\cap Y_I,X_O\cap Y_O)+p(X_I\cup Y_I,X_O\cup Y_O) \tag{2.15}$$

for any $(X_I,X_O),(Y_I,Y_O)\in\mathcal{P}_2'(V)$.

By taking $p(X_I,X_O)=f(X_I,X_O\setminus X_I)$ *and* $f(A,B)=p(A,A\cup B)$, *we can transform from* $f:\mathcal{P}_2(V)\to\mathbb{R}$ *to* $p:\mathcal{P}_2'(V)\to\mathbb{R}$ *and vice versa. Such* f *and* p *are said to be equivalent.*

Now we show necessary and sufficient conditions for the convexity of f^L.

THEOREM 2.4. *Let* $f:\mathcal{P}_2(V)\to[0,+\infty)$ *is a set-pair functions satisfying* $f(\varnothing,\varnothing)=0$. *Then* f^L *is convex if and only if* $\forall (A,B),(C,D)\in\mathcal{P}_2(V)$

$$f(A,B)+f(C,D)\ge f((A\cup C)\setminus(B\cup D),(B\cup D)\setminus(A\cup C))+f(A\cap C,B\cap D); \tag{2.16}$$

if and only if $\forall (X_I,X_O),(Y_I,Y_O)\in\mathcal{P}_2'(V)$ *the equivalent function* p *satisifies*

$$p(X_I,X_O)+p(Y_I,Y_O)\ge p(X_I\cap Y_I,X_O\cap Y_O\setminus Z)+p((X_I\cup Y_I)\setminus Z,(X_O\cup Y_O)\setminus Z), \tag{2.17}$$

where $Z=(X_O\cap Y_I\setminus X_I)\cup(Y_O\cap X_I\setminus Y_I)$.

Three lemmas below are needed in proving Theorem 2.4.

LEMMA 2.1. *For* $\boldsymbol{x}\in\mathbb{R}^n$, $N\in\mathbb{N}^+$, *and* $N>2\|\boldsymbol{x}\|_\infty$, *let*

$$\hat{f}_N(\boldsymbol{x})=\min\left\{\sum_{(A,B)\in\mathcal{P}_2(V)}\lambda_{A,B}f(A,B)\,\middle|\,\begin{array}{l}\sum\lambda_{A,B}\mathbf{1}_{A,B}=\boldsymbol{x},\\ \sum\lambda_{A,B}\le N,\\ \lambda_{A,B}\ge 0.\end{array}\right\}. \tag{2.18}$$

Then $\hat{f}_N(\boldsymbol{x})$ *is convex.*

Proof. The fact that $\hat{f}_N(\boldsymbol{x})$ is well defined emerges from Proposition 2.4. Given $\boldsymbol{x},\boldsymbol{y}\in\mathbb{R}^n$, from (2.18), we deduce

$$\hat{f}_N(\boldsymbol{x})=\sum_{(A,B)\in\mathcal{P}_2(V)}\alpha_{A,B}f(A,B)$$

holds for some $\alpha_{A,B}\ge 0$ with $\sum\alpha_{A,B}\mathbf{1}_{A,B}=\boldsymbol{x}$. Similarly,

$$\hat{f}_N(\boldsymbol{y})=\sum_{(A,B)\in\mathcal{P}_2(V)}\beta_{A,B}f(A,B)$$

holds for some $\beta_{A,B}\ge 0$ with $\sum\beta_{A,B}\mathbf{1}_{A,B}=\boldsymbol{y}$.

Let $\lambda_{A,B}=t\alpha_{A,B}+(1-t)\beta_{A,B}$ with $t\in[0,1]$. Immediately, we have

$$\boldsymbol{z}:=t\boldsymbol{x}+(1-t)\boldsymbol{y}=\sum_{(A,B)\in\mathcal{P}_2(V)}\lambda_{A,B}\mathbf{1}_{A,B}$$

with $\sum\lambda_{A,B}\le N$ and $\lambda_{A,B}\ge 0$, and then

$$\hat{f}_N(\boldsymbol{z})\le\sum_{(A,B)\in\mathcal{P}_2(V)}\lambda_{A,B}f(A,B)=t\hat{f}_N(\boldsymbol{x})+(1-t)\hat{f}_N(\boldsymbol{y}).$$

∎

DEFINITION 2.5. *A set-pair function* $f:\mathcal{P}_2(V)\to[0,+\infty)$ *is said to be strictly submodular if, the inequality* (2.16) *holds. Moreover, the equality holds if and only if* $(A,B)\subset(C,D)$ *or* $(A,B)\supset(C,D)$.

LEMMA 2.2. *If f is strictly submodular, then $\hat{f}_N(\boldsymbol{x})=f^L(\boldsymbol{x})$ for $N>c\|\boldsymbol{x}\|_\infty$ with $c>1$.*

Proof. Given $\boldsymbol{x}\in\mathbb{R}^n$, according to Lemma 2.1, there exist $\lambda_{A,B}\geq 0,\forall(A,B)\in P_2(V)$ with $\sum\lambda_{A,B}\mathbf{1}_{A,B}=\boldsymbol{x}$ and $\sum\lambda_{A,B}\leq N$ such that

$$\hat{f}_N(\boldsymbol{x})=\sum_{(A,B)\in\mathcal{P}_2(V)}\lambda_{A,B}f(A,B).$$

Without loss of generality, we can assume $\lambda_{\varnothing,\varnothing}=0$.

We claim: if $\lambda_{A,B}\geq\lambda_{C,D}>0$, then either $(A,B)\subset(C,D)$ or $(C,D)\subset(A,B)$. Suppose the contrary and let

$$\begin{cases}\lambda'_{A,B}=\lambda_{A,B}-\lambda_{C,D},\\ \lambda'_{C,D}=0,\\ \lambda'_{A',B'}=\lambda_{A',B'}+\lambda_{C,D},\\ \lambda'_{C',D'}=\lambda_{C',D'}+\lambda_{C,D},\\ \lambda'_{E,F}=\lambda_{E,F},\forall(E,F)\in\mathcal{P}_2(V)\setminus\{(A,B),(C,D),(A',B'),(C',D')\},\end{cases}$$

where

$$A'=(A\cup C)\setminus(B\cup D),B'=(B\cup D)\setminus(A\cup C),C'=A\cap C,D'=B\cap D.$$

Then it can be easily verified that

$$\sum_{(P,Q)\in\mathcal{P}_2(V)}\lambda'_{P,Q}=\sum_{(P,Q)\in\mathcal{P}_2(V)}\lambda_{P,Q}\quad\text{and}\quad\sum_{(P,Q)\in\mathcal{P}_2(V)}\lambda'_{P,Q}\mathbf{1}_{P,Q}=\boldsymbol{x}.$$

Direct calculation shows

$$\begin{aligned}&\sum_{(P,Q)\in\mathcal{P}_2(V)}\lambda'_{P,Q}f(P,Q)-\sum_{(P,Q)\in\mathcal{P}_2(V)}\lambda_{P,Q}f(P,Q)\\ &=\sum_{(P,Q)\in\mathcal{P}_2(V)}(\lambda'_{P,Q}-\lambda_{P,Q})f(P,Q)\\ &=\lambda_{C,D}(-f(A,B)-f(C,D)+f(A',B')+f(C',D'))<0,\end{aligned}$$

provided the strict submodularity of f. This contradicts the minimality of $\hat{f}(\boldsymbol{x})$. According to the mathematical induction we obtain $(\varnothing,\varnothing)\neq(V_p^+,V_p^-)\subset\cdots\subset(V_0^+,V_0^-)$ with

$$\hat{f}_N(\boldsymbol{x})=\sum_{i=0}^{p}\lambda_{V_i^+,V_i^-}f(V_i^+,V_i^-).$$

Moreover, we have $\sum_{i=0}^{p}\lambda_{V_i^+,V_i^-}=\|\boldsymbol{x}\|_\infty$ via (2.12). After Proposition 2.4, $\hat{f}_N(\boldsymbol{x})=f^L(\boldsymbol{x})$. ∎

LEMMA 2.3. *The function*

$$g(A,B):=\sqrt{|A|+|B|}$$

is strictly submodular.

Proof. Given $(A,B),(C,D)\in\mathcal{P}_2(V)$, let $A'=A\setminus D$, $B'=B\setminus C$, $C'=C\setminus B$ and $D'=D\setminus A$. Then

$$\begin{aligned}&g((A\cup C)\setminus(B\cup D),(B\cup D)\setminus(A\cup C))+g(A\cap C,B\cap D)\\&=\sqrt{|A'\cup C'|+|B'\cup D'|}+\sqrt{|A'\cap C'|+|B'\cap D'|}\\&\le\sqrt{|A'|+|B'|}+\sqrt{|C'|+|D'|}\\&\le\sqrt{|A|+|B|}+\sqrt{|C|+|D|}=g(A,B)+g(C,D),\end{aligned}$$

where the first inequality holds since the function $\sqrt{t}$ is strictly convex. Meanwhile, we can easily see that the equality holds if and only if $(A,B)\subset(C,D)$ or $(C,D)\subset(A,B)$. The proof is thus completed. ∎

Proof. **(Proof of Theorem 2.4.)** Suppose that f satisfies (2.16). For any $\alpha>0$, $f+\alpha g$ is strictly submodular according to Lemma 2.3. Thus, by Lemma 2.2, we have

$$f^L+\alpha g^L=(f+\alpha g)^L=\widehat{f+\alpha g}\ge\hat{f}. \tag{2.19}$$

Given $\boldsymbol{x}\in\mathbb{R}^n$, set $\hat{f}=\hat{f}_N$ for fixed $N>2\|\boldsymbol{x}\|_\infty$. Hence, (2.18) leads to

$$\hat{f}(\boldsymbol{x})=\sum_{(A,B)\in\mathcal{P}_2(V)}\lambda_{A,B}f(A,B)$$

for some $\lambda_{A,B}\ge0$ with $\sum\lambda_{A,B}\mathbf{1}_{A,B}=\boldsymbol{x}$, $\sum\lambda_{A,B}<N$. Then

$$\begin{aligned}f^L(\boldsymbol{x})+\alpha g^L(\boldsymbol{x})=\widehat{f+\alpha g}(\boldsymbol{x})&\le\sum_{(A,B)\in\mathcal{P}_2(V)}\lambda_{A,B}(f(A,B)+\alpha g(A,B))\\&=\hat{f}(\boldsymbol{x})+\alpha\sum_{(A,B)\in\mathcal{P}_2(V)}\lambda_{A,B}g(A,B)\\&\le\hat{f}(\boldsymbol{x})+\alpha\sum_{(A,B)\in\mathcal{P}_2(V)}\lambda_{A,B}\sqrt{n}\\&\le\hat{f}(\boldsymbol{x})+\alpha N\sqrt{n}.\end{aligned}\tag{2.20}$$

Letting $\alpha\to0$ in (2.19) and (2.20) yields

$$f^L(\boldsymbol{x})=\hat{f}(\boldsymbol{x}),$$

and by Lemma 2.1, $\hat{f}(\boldsymbol{x})$ is convex, so is $f^L(\boldsymbol{x})$.

On the other hand, if $f^L(\boldsymbol{x})$ is convex, then

$$\begin{aligned}&f((A\cup C)\setminus(B\cup D),(B\cup D)\setminus(A\cup C))+f(A\cap C,B\cap D)\\&=f^L(\mathbf{1}_{(A\cup C)\setminus(B\cup D),(B\cup D)\setminus(A\cup C)}+\mathbf{1}_{A\cap C,B\cap D})\quad\text{[Proposition 2.4]}\\&=f^L(\mathbf{1}_{A,B}+\mathbf{1}_{C,D})\\&=2f^L((\mathbf{1}_{A,B}+\mathbf{1}_{C,D})/2)\quad\text{[Proposition 2.5: (2)]}\\&\le f^L(\mathbf{1}_{A,B})+f^L(\mathbf{1}_{C,D})\quad\text{[convexity]}\\&=f(A,B)+f(C,D),\quad\text{[Equation (2.7)]}\end{aligned}$$

where we have used $\mathbf{1}_{A,B}+\mathbf{1}_{C,D}=\mathbf{1}_{(A\cup C)\setminus(B\cup D),(B\cup D)\setminus(A\cup C)}+\mathbf{1}_{A\cap C,B\cap D}$ in the third line. Thus, (2.16) is true.

Finally, let $X_I=A$, $X_O=A\cup B$, $Y_I=C$, $Y_O=C\cup D$, then

$$Z=(X_O\cap Y_I\setminus X_I)\cup(Y_O\cap X_I\setminus Y_I)=(B\cap C)\cup(D\cap A).$$

By taking $f(A,B)=p(A,A\cup B)$, we can translate inequality (2.17) into

$$\begin{aligned}
&f(A,B)+f(C,D)=p(X_I,X_O)+p(Y_I,Y_O)\\
\geq&p(X_I\cap Y_I,X_O\cap Y_O\setminus Z)+p((X_I\cup Y_I)\setminus Z,(X_O\cup Y_O)\setminus Z)\\
=&p(A\cap C,(A\cap C)\cup(B\cap D))\\
&+p((A\cup C)\setminus(B\cup D),((A\cup C)\setminus(B\cup D))\cup((B\cup D)\setminus(A\cup C)))\\
=&f(A\cap C,B\cap D)+f((A\cup C)\setminus(B\cup D),(B\cup D)\setminus(A\cup C)),
\end{aligned}$$

which means (2.16) and (2.17) are equivalent.

Hence f^L is convex if and only if either (2.16) or (2.17) holds. ∎

Comparing (2.17) of Theorem 2.4 to (2.15) of Definition 2.4, we are able to deduce that the submodularity introduced in Definition 2.4 for a set-pair function fails to ensure an extension of the equivalence stated in Theorem 2.3 between convexity and submodularity for f_o^L into f^L (i.e., Definition 2.4 is neither necessary nor sufficient for f^L to be convex), whereas (2.16) succeeds. In such sense, we might call the set-pair function satisfying (2.16) or (2.17) to be submodular. However, both (2.16) and (2.17) are not so easy-looking that we give a concise necessary condition for f^L to be convex.

Definition 2.6. *A set-pair-function $f:\mathcal{P}_2(V)\to\mathbb{R}$ is partially submodular if and only if it is submodular for each component, i.e.,*

$$f(A,B)+f(A,D)\geq f(A,B\cup D)+f(A,B\cap D),$$

$$f(A,B)+f(C,B)\geq f(A\cup C,B)+f(A\cap C,B),$$

for all subsets $A,B,C,D\subset V$ with $A\cap B=A\cap D=C\cap B=\varnothing$.

Corollary 2.1. *If f^L is convex, then f must be partially submodular.*

Proof. If f^L is convex, then f must satisfy (2.16). Setting $C=A$ and $D=B$ respectively in (2.16), we can find that f is partially submodular. ∎

Similar to Corollary 2.1, if p is submodular, then its equivalent function f must be partially submodular.

Finally, we compare the set-pair Lovász extension to the original one:

(1) In contrast to the succinct integral form (2.8) of f^L, the integral form (2.5) of f_o^L has an extra remainder term.

(2) The original Lovász extension is unable directly to deal with graph 3-cut problems such as the dual Cheeger-cut problem, whereas the set-pair Lovász extension works.

(3) The characterization of the convexity of f_o^L is easier than f^L.

3. Applications to graph cut

A straightforward application of the original Lovász extension (2.2) into a graph 3-cut problem such as the dual Cheeger problem is not feasible. Instead, we will show in this section that the set-pair Lovász extension (1.4) can succeed to find an explicit and equivalent continuous optimization problem for graph 3-cut. To be more specific, the set-pair Lovász extension of the five set-pair functions is summarized in Table 3.1.

Object function	Set-pair Lovász extension
$F_1(A,B)=\lvert\partial A\rvert+\lvert\partial B\rvert$	$F_1^L(\boldsymbol{x})=I(\boldsymbol{x})$
$F_2(A,B)=\lvert E(A,B)\rvert$	$F_2^L(\boldsymbol{x})=\frac{1}{2}\lVert\boldsymbol{x}\rVert-\frac{1}{2}I^+(\boldsymbol{x})$
$G_1(A,B)=\mathrm{vol}(V)$	$G_1^L(\boldsymbol{x})=\mathrm{vol}(V)\lVert\boldsymbol{x}\rVert_\infty$
$G_2(A,B)=\mathrm{vol}(A)+\mathrm{vol}(B)$	$G_2^L(\boldsymbol{x})=\lVert\boldsymbol{x}\rVert$
$G_3(A,B)=\sum_{X\in\{A,B\}}\min_{Y\in\{X,X^c\}}\mathrm{vol}(Y)$	$G_3^L(\boldsymbol{x})=\min_{\alpha\in\mathbb{R}}\lVert\lvert\boldsymbol{x}\rvert-\alpha\mathbf{1}\rVert$

TABLE 3.1. *Set-pair Lovász extension of five object functions.*

The first four functions in Table 3.1 can be calculated directly according to Proposition 2.3. In fact,

$$F_1^L(\boldsymbol{x})=\int_0^{\lVert\boldsymbol{x}\rVert_\infty}\lvert\partial V_t^+(\boldsymbol{x})\rvert+\lvert\partial V_t^-(\boldsymbol{x})\rvert dt. \tag{3.1}$$

Then substituting

$$\begin{aligned}\lvert\partial V_t^+(\boldsymbol{x})\rvert&=\sum_{i<j}w_{ij}(\chi_{x_i\le t<x_j}+\chi_{x_j\le t<x_i}),\\ \lvert\partial V_t^-(\boldsymbol{x})\rvert&=\sum_{i<j}w_{ij}(\chi_{x_i<-t\le x_j}+\chi_{x_j<-t\le x_i})\end{aligned}$$

into Equation (3.1) yields

$$\begin{aligned}F_1^L(\boldsymbol{x})&=\sum_{i<j}w_{ij}\int_0^{\lVert\boldsymbol{x}\rVert_\infty}\chi_{x_i\le t<x_j}+\chi_{x_j\le t<x_i}+\chi_{x_i<-t\le x_j}+\chi_{x_j<-t\le x_i}dt\\ &=\sum_{i<j}w_{ij}\int_{-\lVert\boldsymbol{x}\rVert_\infty}^{\lVert\boldsymbol{x}\rVert_\infty}\chi_{x_i<t<x_j}+\chi_{x_j<t<x_i}dt\\ &=\sum_{i<j}w_{ij}\lvert x_i-x_j\rvert=I(\boldsymbol{x}),\end{aligned}$$

where the integral equalities hold when '$<$' is replaced by '$\le$'.

Hereafter the endpoints of intervals in the integral form (2.8) are dropped for convenience. Thus, it gives the form of set-pair Lovász extension of F_1 in Table 3.1.

Applying Proposition 2.3 to F_2, we get

$$\begin{aligned}F_2^L(\boldsymbol{x})&=\int_0^{\lVert\boldsymbol{x}\rVert_\infty}F_2(V_t^+(\boldsymbol{x}),V_t^-(\boldsymbol{x}))dt\\ &=\int_0^{\lVert\boldsymbol{x}\rVert_\infty}\lvert E(V_t^+(\boldsymbol{x}),V_t^-(\boldsymbol{x}))\rvert dt\\ &=\int_0^{\lVert\boldsymbol{x}\rVert_\infty}\sum_{i=1}^n\sum_{j=1}^n w_{ij}\chi_{x_i>t}\chi_{x_j<-t}dt\\ &=\sum_{i=1}^n\sum_{j=1}^n w_{ij}\int_0^{\lVert\boldsymbol{x}\rVert_\infty}\chi_{x_i>t}\chi_{x_j<-t}dt\end{aligned}$$

$$=\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}w_{ij}\min\{x_i+|x_i|,|x_j|-x_j\}$$
$$=\frac{1}{4}\sum_{i=1}^{n}\sum_{j=1}^{n}w_{ij}(x_i+|x_i|+|x_j|-x_j-|x_i+x_j+|x_i|-|x_j||),$$

where we used the fact that $\min\{a,b\}=\frac{1}{2}(a+b-|a-b|)$ in the last equality. Further, we can obtain

$$\begin{aligned}F_2^L(\boldsymbol{x})&=\frac{1}{4}\sum_{i<j}w_{ij}(2(|x_i|+|x_j|)-\big|x_i+x_j+|x_i|-|x_j|\big|-\big|x_i+x_j-|x_i|+|x_j|\big|)\\&=\frac{1}{4}\sum_{i<j}w_{ij}(2|x_i|+2|x_j|-2\max\{|x_i+x_j|,\big||x_i|-|x_j|\big|\})\\&=\frac{1}{2}\|\boldsymbol{x}\|-\frac{1}{2}\sum_{i<j}|x_i+x_j|=\frac{1}{2}\|\boldsymbol{x}\|-\frac{1}{2}I^+(\boldsymbol{x}),\end{aligned}\tag{3.2}$$

where Equation (3.2) utilizes the fact that $2\max\{|a|,|b|\}=|a+b|+|a-b|$.

Correspondingly, the set-pair extensions of G_1 and G_2 in Table 3.1 can be verified in the following way:

$$\begin{aligned}G_1^L(\boldsymbol{x})&=\int_0^{\|\boldsymbol{x}\|_\infty}G_1(V_t^+(\boldsymbol{x}),V_t^-(\boldsymbol{x}))dt=\mathrm{vol}(V)\|\boldsymbol{x}\|_\infty,\\G_2^L(\boldsymbol{x})&=\int_0^{\|\boldsymbol{x}\|_\infty}G_2(V_t^+(\boldsymbol{x}),V_t^-(\boldsymbol{x}))dt\\&=\int_0^{\|\boldsymbol{x}\|_\infty}\mathrm{vol}(V_t^+(\boldsymbol{x}))+\mathrm{vol}(V_t^-(\boldsymbol{x}))dt\\&=\int_0^{\|\boldsymbol{x}\|_\infty}\sum_{i=1}^{n}d_i\chi_{|x_i|>t}dt=\|\boldsymbol{x}\|.\end{aligned}$$

Now, let us focus on G_3. Direct calculation shows

$$\begin{aligned}G_3^L(\boldsymbol{x})&=\int_0^{\|\boldsymbol{x}\|_\infty}G_3(V_t^+(\boldsymbol{x}),V_t^-(\boldsymbol{x}))dt\\&=\int_0^{\|\boldsymbol{x}\|_\infty}\min\{\mathrm{vol}(V_t^+(\boldsymbol{x})),\mathrm{vol}(V_t^+(\boldsymbol{x})^c)\}+\min\{\mathrm{vol}(V_t^-(\boldsymbol{x})),\mathrm{vol}(V_t^-(\boldsymbol{x})^c)\}dt\\&=\int_0^{\|\boldsymbol{x}\|_\infty}\min\{\mathrm{vol}(V_t^+(\boldsymbol{x})),\mathrm{vol}(V_t^+(\boldsymbol{x})^c)\}+\min\{\mathrm{vol}(V_{-t}^+(\boldsymbol{x})^c),\mathrm{vol}(V_{-t}^+(\boldsymbol{x}))\}dt\\&=\int_{-\|\boldsymbol{x}\|_\infty}^{\|\boldsymbol{x}\|_\infty}\min\{\mathrm{vol}(V_t^+(\boldsymbol{x})),\mathrm{vol}(V_t^+(\boldsymbol{x})^c)\}dt.\end{aligned}$$

Let σ be a permutation of $\{1,2,\dots,n\}$ such that $x_{\sigma(1)}\le x_{\sigma(2)}\le\cdots\le x_{\sigma(n)}$. Then there exists $k_0\in\{1,2,\dots,n\}$ satisfying

$$\sum_{i=1}^{k_0-1}d_{\sigma(i)}<\frac{1}{2}\mathrm{vol}(V)\le\sum_{i=1}^{k_0}d_{\sigma(i)}.\tag{3.3}$$

Consequently, it reveals that

$$\min\{\mathrm{vol}(V_t^+(\boldsymbol{x})),\mathrm{vol}(V_t^+(\boldsymbol{x})^c)\}=\begin{cases}\mathrm{vol}(V_t^+(\boldsymbol{x})), & \text{if } t<x_{\sigma(k_0)},\\ \mathrm{vol}(V_t^+(\boldsymbol{x})^c), & \text{if } t\geq x_{\sigma(k_0)},\end{cases}\tag{3.4}$$

and

$$\begin{aligned}G_3^L(\boldsymbol{x})&=\int_{x_{\sigma(1)}}^{x_{\sigma(k_0)}}\mathrm{vol}(V_t^+(x)^c)dt+\int_{x_{\sigma(k_0)}}^{x_{\sigma(n)}}\mathrm{vol}(V_t^+(x))dt\\ &=\sum_{i=1}^{k_0-1}(x_{\sigma(i+1)}-x_{\sigma(i)})\sum_{j=1}^{i}d_{\sigma(j)}+\sum_{i=k_0}^{n-1}(x_{\sigma(i+1)}-x_{\sigma(i)})\sum_{j=i+1}^{n}d_{\sigma(j)}\\ &=\sum_{j=1}^{k_0-1}d_{\sigma(j)}\sum_{i=j}^{k_0-1}(x_{\sigma(i+1)}-x_{\sigma(i)})+\sum_{j=k_0+1}^{n}d_{\sigma(j)}\sum_{i=k_0}^{j-1}(x_{\sigma(i+1)}-x_{\sigma(i)})\\ &=\sum_{i=1}^{n}d_{\sigma(i)}|x_{\sigma(i)}-x_{\sigma(k_0)}|=\|\boldsymbol{x}-x_{\sigma(k_0)}\mathbf{1}\|.\end{aligned}$$

On the other hand, $\|\boldsymbol{x}-\alpha\mathbf{1}\|$ is convex in α and satisfies

$$p_\alpha:=-\sum_{i=1}^{n}d_{\sigma(i)}\mathrm{sign}(x_{\sigma(i)}-\alpha)\in\partial_\alpha\|\boldsymbol{x}-\alpha\mathbf{1}\|$$

and then

$$\begin{cases}p_\alpha\leq\sum_{j=1}^{k_0-1}d_{\sigma(j)}-\sum_{j=k_0}^{n}d_{\sigma(j)}\leq 0, & \text{if } \alpha<x_{\sigma(k_0)},\\ p_\alpha\geq\sum_{j=1}^{k_0}d_{\sigma(j)}-\sum_{j=k_0+1}^{n}d_{\sigma(j)}\geq 0, & \text{if } \alpha>x_{\sigma(k_0)}.\end{cases}$$

This implies that $\|\boldsymbol{x}-\alpha\mathbf{1}\|$ is decreasing with respect to α in $(-\infty,x_{\sigma(k_0)})$ and increasing in $(x_{\sigma(k_0)},+\infty)$. Thus, we obtain that

$$x_{\sigma(k_0)}\in\arg\min_{\alpha\in\mathbb{R}}\|\boldsymbol{x}-\alpha\mathbf{1}\|.$$

Therefore,

$$G_3^L(\boldsymbol{x})=\|\boldsymbol{x}-x_{\sigma(k_0)}\mathbf{1}\|=\min_{\alpha\in\mathbb{R}}\|\boldsymbol{x}-\alpha\mathbf{1}\|.\tag{3.5}$$

3.1. Graph 3-cut problems. It is straightforward to utilize the proposed set-pair Lovász extension to deal with the combination optimizations in a set-pair form, for example, the dual Cheeger and max 3-cut problems. Now, we are in a position to answer Question 1.2.

Proof. **(Proof of Theorem 1.4.)** Applying Theorem 1.3 in the dual Cheeger cut problem (1.1) yields

$$h^+(G)=\sup_{\boldsymbol{x}\neq\mathbf{0}}\frac{2F_2^L(\boldsymbol{x})}{G_2^L(\boldsymbol{x})}=1-\inf_{\boldsymbol{x}\neq\mathbf{0}}\frac{I^+(\boldsymbol{x})}{\|\boldsymbol{x}\|},\tag{3.6}$$

where we have used F_2^L and G_2^L in Table 3.1. ∎

Proof. **(Proof of Theorem 1.5.)** Since $A\cap B=B\cap C=C\cap A=\varnothing$ and $A\cup B\cup C=V$, we can suppose that $A\cup B\neq\varnothing$. Then, by Theorem 1.3 and Proposition 2.5, we have

$$\begin{aligned}\frac{1}{2}h_{\max,3}(G) &= \max_{A,B,C}\frac{|E(A,B)|+|E(B,C)|+|E(C,A)|}{\mathrm{vol}(V)}\\ &= \max_{A,B,C}\frac{|E(A,B)|+|E(A\cup B,C)|}{\mathrm{vol}(V)}\\ &= \max_{(A,B)\in\mathcal{P}_2(V)\setminus\{(\varnothing,\varnothing)\}}\frac{|E(A,B)|+|\partial(A\cup B)|}{\mathrm{vol}(V)}\\ &= \max_{(A,B)\in\mathcal{P}_2(V)\setminus\{(\varnothing,\varnothing)\}}\frac{|\partial A|+|\partial B|-|E(A,B)|}{\mathrm{vol}(V)}\\ &= \max_{(A,B)\in\mathcal{P}_2(V)\setminus\{(\varnothing,\varnothing)\}}\frac{F_1(A,B)-F_2(A,B)}{G_1(A,B)}\\ &= \sup_{\boldsymbol{x}\neq\boldsymbol{0}}\frac{F_1^L(\boldsymbol{x})-F_2^L(\boldsymbol{x})}{G_1^L(\boldsymbol{x})},\end{aligned}$$

where F_1^L and G_1^L are given in Table 3.1. Thus

$$h_{\max,3}(G)=\sup_{\boldsymbol{x}\neq\boldsymbol{0}}\frac{2I(\boldsymbol{x})-\|\boldsymbol{x}\|+I^+(\boldsymbol{x})}{\mathrm{vol}(V)\|\boldsymbol{x}\|_\infty}. \tag{3.7}$$

It follows from $|a-b|+|a+b|=2\max\{|a|,|b|\}=||a|-|b||+|a|+|b|$ for any $a,b\in\mathbb{R}$ that

$$\begin{aligned}I(\boldsymbol{x}) &= \sum_{i<j}w_{ij}|x_i-x_j| = \sum_{i<j}w_{ij}(|x_i|+|x_j|+||x_i|-|x_j||-|x_i+x_j|)\\ &= \sum_{i=1}^n d_i|x_i|+\sum_{i<j}w_{ij}||x_i|-|x_j||-\sum_{i<j}w_{ij}|x_i+x_j|\\ &= \|\boldsymbol{x}\|+\hat{I}(\boldsymbol{x})-I^+(\boldsymbol{x}),\end{aligned}$$

and thus

$$2I(\boldsymbol{x})-\|\boldsymbol{x}\|+I^+(\boldsymbol{x})=2\hat{I}(\boldsymbol{x})+\|\boldsymbol{x}\|-I^+(\boldsymbol{x})=I(\boldsymbol{x})+\hat{I}(\boldsymbol{x}).$$

Finally, Equation (3.7) turns out to be

$$h_{\max,3}(G)=\sup_{\boldsymbol{x}\neq\boldsymbol{0}}\frac{I(\boldsymbol{x})+\hat{I}(\boldsymbol{x})}{\mathrm{vol}(V)\|\boldsymbol{x}\|_\infty}.$$

∎

Proof. **(Proof of Theorem 1.6.)** According to Theorem 1.3, we have

$$\begin{aligned}h_{\max,3,I}(G) &= \max_{(A,B)\in\mathcal{P}_2(V)}\frac{F_1(A,B)-F_2(A,B)}{G_2(A,B)}\\ &= \sup_{\boldsymbol{x}\neq\boldsymbol{0}}\frac{\|\boldsymbol{x}\|-I^+(\boldsymbol{x})+2\hat{I}(\boldsymbol{x})}{\|\boldsymbol{x}\|}.\end{aligned}$$

∎

Proof. **(Proof of Theorem 1.7.)** Let

$$G(A,B)=\min\{\mathrm{vol}(A\cup B),\mathrm{vol}((A\cup B)^c)\}, \tag{3.8}$$

and σ be a permutation of $\{1,2,\dots,n\}$ such that $|x_{\sigma(1)}|\le|x_{\sigma(2)}|\le\cdots\le|x_{\sigma(n)}|$. By Definition 1.2, the set-pair Lovász extension of $G(A,B)$ is

$$\begin{aligned}G^L(\boldsymbol{x}) &= \sum_{i=0}^{n-1}(|x_{\sigma(i+1)}|-|x_{\sigma(i)}|)G(V^+_{\sigma(i)},V^-_{\sigma(i)})\\ &= \sum_{i=0}^{n-1}(|x_{\sigma(i+1)}|-|x_{\sigma(i)}|)\min\{\mathrm{vol}(V^+_{\sigma(i)}\cup V^-_{\sigma(i)}),\mathrm{vol}((V^+_{\sigma(i)}\cup V^-_{\sigma(i)})^c)\}\\ &= \sum_{i=0}^{n-1}(|x_{\sigma(i+1)}|-|x_{\sigma(i)}|)G_3(V^+_{\sigma(i)}\cup V^-_{\sigma(i)},\varnothing)\\ &= G_3^L(|\boldsymbol{x}|)=\min_{\alpha\in\mathbb{R}}\||\boldsymbol{x}|-\alpha\mathbf{1}\|,\end{aligned}$$

where $|\boldsymbol{x}|=(|x_1|,\dots,|x_n|)$ and Equation (3.5) is applied in the last line. Finally, applying Theorem 1.3 into Equation (1.18) leads to

$$\begin{aligned}h_{\max,3,II}(G) &= \sup_{\boldsymbol{x}\neq\mathbf{0}}\frac{2F_1^L(\boldsymbol{x})-2F_2^L(\boldsymbol{x})}{G_1^L(\boldsymbol{x})-G^L(\boldsymbol{x})}\\ &= \sup_{\boldsymbol{x}\neq\mathbf{0}}\frac{2I(\boldsymbol{x})-\|\boldsymbol{x}\|+I^+(\boldsymbol{x})}{\mathrm{vol}(V)\|\boldsymbol{x}\|_\infty-\min_{\alpha\in\mathbb{R}}\sum_{i=1}^n d_i\,||x_i|-\alpha|}.\end{aligned}$$

□

3.2. Graph k-cut ($k>3$) problems. In this section, we present a preliminary attempt to a graph k-cut problem. The main idea is to transfer a graph k-cut problem to a 3-cut one on a larger graph. To this end, let us start from

$$\mathcal{P}_k([n])=\{(A_1,\dots,A_k)\big|A_i\cap A_j=\varnothing,A_i\subset[n]\},\tag{3.9}$$

$$\mathcal{H}_{k+1}([n])=\{(A_1,\dots,A_{k+1})\big|A_i\cap A_j=\varnothing,\bigcup_{i=1}^{k+1}A_i=[n]\},\tag{3.10}$$

where $[n]=\{1,2,\dots,n\}$. Obviously $\mathcal{P}_k([n])$ is equivalent to $\mathcal{H}_{k+1}([n])$, $\mathcal{P}_k([n])\simeq\mathcal{H}_{k+1}([n])$. Then a bijection between $\mathcal{P}_2([ln])$ and $\mathcal{H}_{3^l}([n])$ can be obtained via

$$\mathcal{P}_2([ln])\simeq\mathcal{H}_3([ln])\simeq\prod_{i=1}^{l}\mathcal{H}_3([n])\simeq\mathcal{H}_{3^l}([n]).\tag{3.11}$$

For a family $\mathcal{P}$ consisting of set-tuples, we use $C(\mathcal{P}):=\{f:\mathcal{P}\to\mathbb{R}\}$ to denote the collection of real valued functions on $\mathcal{P}$, and then have the following commutative diagrams for any $k<3^l$:

In Figure 3.1, h_1 is the natural injective mapping from $\mathcal{H}_{3^l}([n])$ to $\mathcal{P}_k([n])$ by choosing only the last k parts from each element in $H_{3^l}([n])$. Therefore, given $f\in C(\mathcal{P}_k([n]))$, there exists $F\in C(\mathcal{P}_2([ln]))$ and an injective mapping h from $\mathcal{P}_2([ln])$ to $\mathcal{P}_k([n])$ such that $F=f\circ h$. For convenience, the set-pair Lovász extension of F is again called the Lovász extension of f. That is, there exist $F_1\in C(H_{3^l}([n]))$ and $F_2\in C(\prod_{i=1}^l H_3([n]))$ such that

$$F_1=f\circ h_1,\ F_2(\prod_{i=1}^{l}(T_0^i,T_1^i,T_2^i))=F_1((A_0,A_1,\dots,A_{3^l-1})),\tag{3.12}$$

$$
\begin{tikzcd}
\mathcal{P}_k([n]) & \mathcal{P}_2([ln]) \arrow[l, "h"'] \arrow[d, "h_3"] \\
\mathcal{H}_{3^l}([n]) \arrow[u, "h_1"] & \prod_{i=1}^{l} \mathcal{H}_3([n]) \arrow[l, "h_2"']
\end{tikzcd}
\qquad
\begin{tikzcd}
C(\mathcal{P}_k([n])) \arrow[r, "h^*"] \arrow[d, "h_1^*"'] & C(\mathcal{P}_2([ln])) \\
C(\mathcal{H}_{3^l}([n])) \arrow[r, "h_2^*"] & C(\prod_{i=1}^{l} \mathcal{H}_3([n])) \arrow[u, "h_3^*"']
\end{tikzcd}
$$

FIG. 3.1. *The commutative diagram on the right is indeed the dual diagram of the left one in some sense.*

where $(T_0^i,T_1^i,T_2^i)\in H_3([n])$, $i=1,\dots,l$, and $A_j=\bigcap_{i=1}^l T_{a_i}^i$ with $(a_l\dots a_1)_3$ being the ternary representation of j for $j=0,1,\dots,3^l-1$. In other words, there exists an injection h_2 from $\prod_{i=1}^l H_3([n])$ to $H_{3^l}([n])$ such that

$$F_2=F_1\circ h_2. \tag{3.13}$$

Finally, we can take an injection h_3 from $\mathcal{P}_2([ln])$ to $H_{3^l}([n])$ by

$$h(T_1,T_2)=\prod_{i=1}^{l}(T_0^i,T_1^i,T_2^i),$$

where

$$T_a^i=\{t\in[n]|t+n(i-1)\in T_a\},\quad T_0=(T_1\cap T_2)^c,\quad a=0,1,2. \tag{3.14}$$

Thus, the correspondence is well established between the function f defined on $P_k([n])$ and the function F on $P_2([ln])$ by letting $h=h_1\circ h_2\circ h_3$ and

$$F=F_2\circ h_3=F_1\circ h_2\circ h_3=f\circ h.$$

Applying Theorem 1.3, we are able to give an equivalent continuous optimization for the max k-cut problem.

DEFINITION 3.1 (max k-cut [11]). *Given a connected graph $G=(V,E)$ with $V=[n]$, the max k-cut problem is to determine a graph k-cut by solving*

$$h_k(G)=\min_{(A_1,A_2,\dots,A_k)\in\mathcal{H}_k([n])}\frac{\sum_{i=1}^k|\partial A_i|}{\sum_{i=1}^k \mathrm{vol}(A_i)}. \tag{3.15}$$

We can find $F,G\in C(\mathcal{P}_2([ln]))$ such that

$$F(T_1,T_2)=\sum_{j=3^l-k}^{3^l-1}|\partial A_j|,\ \text{and}\ G(T_1,T_2)=\sum_{j=3^l-k}^{3^l-1}\mathrm{vol}(A_i), \tag{3.16}$$

where $A_j=\bigcap_{i=1}^l T_{a_i}^i$, $T_{a_i}^i$ is given in Equation (3.14), and $(a_l\dots a_1)_3$ denotes the ternary representation of j for $j\in\{0,1,\dots,3^l-1\}$.

Let us write down the functions F^L and G^L explicitly. In fact,

$$F^L(\prod_{i=1}^{l}\boldsymbol{x}^{(i)})=\int_0^{\|\boldsymbol{x}\|_\infty}F(V_t^+,V_t^-)dt \tag{3.17}$$

is a continuous function defined on $\mathbb{R}^{nl}$. Here $V_t^{\pm}=\{(i-1)n+j|\pm x_j^{(i)}>t\}$ for any $\boldsymbol{x}=\prod_{i=1}^{l}\boldsymbol{x}^{(i)}\in\mathbb{R}^{nl}$, and $\boldsymbol{x}^{(i)}=(x_1^{(i)},\ldots,x_n^{(i)})$.

PROPOSITION 3.1.

$$F^L(\prod_{i=1}^{l}\boldsymbol{x}^{(i)})=\sum_{j=1}^{n}d_j z_j-2\sum_{i<j}^{n}w_{ij}\sum_{(a_l\ldots a_1)_2=3^l-k}^{3^l-1}z_{ij}^{(a_l\ldots a_1)_3},$$

where

$$\begin{aligned}
z_j&=\min\Big\{t\geq 0\Big|(a_l\ldots a_1)_3<3^l-k,\,a_i=\mathbf{1}_{x_j^{(i)}>t}+2\mathbf{1}_{-x_j^{(i)}>t}\Big\},\\
z_{ij}^{(a_l\ldots a_1)_3}&=\min\Big\{-(-1)^{a_\alpha}x_{i'}^{(\alpha)}-|x_{j'}^{(\beta)}|\Big|a_\alpha>0,a_\beta=0,i',j'\in\{i,j\}\Big\}_+,\\
z_+&=\max\{z,0\}.
\end{aligned}$$

Proof. A direct calculation leads to

$$\begin{aligned}
F^L(\prod_{i=1}^{l}\boldsymbol{x}^{(i)})&=\int_0^{\|\boldsymbol{x}\|}F(V_t)dt=\int_0^{\|\boldsymbol{x}\|}\sum_{(a_l\ldots a_1)_3=3^l-k}^{3^l-1}|\partial A_{a_l\ldots a_1}(t)|dt\\
&=\int_0^{\|\boldsymbol{x}\|}\sum_{(a_l\ldots a_1)_3=3^l-k}^{2^l-1}|\mathrm{vol}(A_{a_l\ldots a_1}(t))|-2|E(A_{a_l\ldots a_1}(t))|dt\\
&=\mathrm{I}-\mathrm{II},
\end{aligned}$$

where $E(A)$ is the set of all edges with endpoints in A and

$$\begin{aligned}
\mathrm{I}&=\int_0^{\|\boldsymbol{x}\|}\sum_{(a_l\ldots a_1)_3=3^l-k}^{3^l-1}|\mathrm{vol}(A_{a_l\ldots a_1}(t))|dt,\\
\mathrm{II}&=2\int_0^{\|\boldsymbol{x}\|}\sum_{(a_l\ldots a_1)_3=3^l-k}^{3^l-1}|E(A_{a_l\ldots a_1}(t))|dt.
\end{aligned}$$

It is easy to check the first part

$$\begin{aligned}
\mathrm{I}&=\int_0^{\|\boldsymbol{x}\|}\sum_{(a_l\ldots a_1)_3=3^l-k}^{3^l-1}\sum_{i<j}w_{ij}[\mathbf{1}_{i\in A_{a_l\ldots a_1}(t)}+\mathbf{1}_{j\in A_{a_l\ldots a_1}(t)}]dt\\
&=\sum_{i<j}w_{ij}\int_0^{\|\boldsymbol{x}\|}\sum_{(a_l\ldots a_1)_3=3^l-k}^{3^l-1}\mathbf{1}_{i\in A_{a_l\ldots a_1}(t)}+\mathbf{1}_{j\in A_{a_l\ldots a_1}(t)}dt\\
&=\sum_{i<j}w_{ij}\int_0^{\|\boldsymbol{x}\|}\mathbf{1}_{i\in\bigcup_{(a_l\ldots a_1)_3=3^l-k}^{3^l-1}A_{a_l\ldots a_1}(t)}+\mathbf{1}_{j\in\bigcup_{(a_l\ldots a_1)_3=3^l-k}^{3^l-1}A_{a_l\ldots a_1}(t)}dt\\
&=\sum_{i<j}w_{ij}(z_i+z_j)=\sum_{j=1}^{n}d_j z_j,
\end{aligned}$$

and the second part

$$
\begin{aligned}
\mathrm{II} &= 2\int_0^{\|\boldsymbol{x}\|} \sum_{(a_l\ldots a_1)_3=3^l-k}^{3^l-1} \sum_{i<j} w_{ij}\mathbf{1}_{i,j\in A_{a_l\ldots a_1}(t)}dt \\
&= 2\sum_{i<j} w_{ij} \sum_{(a_l\ldots a_1)_3=3^l-k}^{3^l-1} \int_0^{\|\boldsymbol{x}\|} \mathbf{1}_{i,j\in A_{a_l\ldots a_1}(t)}dt \\
&= 2\sum_{i<j} w_{ij} \sum_{(a_l\ldots a_1)_3=3^l-k}^{3^l-1} \int_0^{\|\boldsymbol{x}\|} \prod_{\substack{\alpha:a_\alpha>0,\\ i'\in\{i,j\}}} \mathbf{1}_{t\le -(-1)^{a_\alpha} x_{i'}^{(\alpha)}} \prod_{\substack{\beta:,a_\beta=0,\\ j'\in\{i,j\}}} \mathbf{1}_{|x_{j'}^{(\beta)}|<t}dt \\
&= 2\sum_{i<j} w_{ij} \sum_{(a_l\ldots a_1)_3=3^l-k}^{3^l-1} z_{ij}^{(a_l\ldots a_1)_3}.
\end{aligned}
$$

Thus, we complete the proof. □

REMARK 3.1. If $k=3^l-1$, then

$$F^L\Big(\prod_{i=1}^{l}\boldsymbol{x}^{(i)}\Big)=\sum_{j=1}^{n} d_i \max_s |x_j^{(s)}| - 2\sum_{i<j}^{n} w_{ij} \sum_{(a_l\ldots a_1)_3=1}^{3^l-1} z_{ij}^{(a_l\ldots a_1)_3}.$$

It can be readily verified that:

PROPOSITION 3.2.

$$G^L\Big(\prod_{i=1}^{l}\boldsymbol{x}^{(i)}\Big)=\sum_{j=1}^{n} d_j z_j = \mathrm{I}.$$

Accordingly, we get

PROPOSITION 3.3.

$$h_k = \sup_{\boldsymbol{x}\in\mathbb{R}^{nl}\setminus K} \frac{F^L(\boldsymbol{x})}{G^L(\boldsymbol{x})}, \tag{3.18}$$

where F^L and G^L are defined in Propositions 3.1 and 3.2, respectively, and $K=\{\boldsymbol{x}|z_j=0,\forall j=1,2,\ldots,n\}$.

REMARK 3.2. It should be noted that we can follow the same procedure to derive a continuous representation of the dual Cheeger cut problem staring from the original Lovász extension. However, it requires $2n$ variables while the equivalent continuous formulation in Equation (1.4) obtained from the set-pair Lovász extension only needs n variables, where n gives the order of graph.

3.3. Graph 2-cut problems. With the help of the following lemma, the proposed set-pair Lovász extension also works for some graph 2-cut problems.

LEMMA 3.1. *Suppose $f,g:\mathcal{P}(V)\to[0,+\infty)$ are two symmetric functions with $g(A)>0$ for any $A\in\mathcal{P}(V)$. Let $F(A,B)=f(A)+f(B)$ and $G(A,B)=g(A)+g(B)$. Then*

$$\min_{A\in\mathcal{P}(V)} \frac{f(A)}{g(A)} = \min_{(A,B)\in\mathcal{P}_2(V)} \frac{F(A,B)}{G(A,B)}, \tag{3.19}$$

$$\max_{A\in\mathcal{P}(V)}\frac{f(A)}{g(A)}=\max_{(A,B)\in\mathcal{P}_2(V)}\frac{F(A,B)}{G(A,B)}. \tag{3.20}$$

Proof. First we prove (3.19). On one hand, let $(A_0,B_0)\in\mathcal{P}_2(V)$ be the minimizer of $\frac{f(A)+f(B)}{g(A)+g(B)}$. Without loss of generality, we may assume $\frac{f(A_0)}{g(A_0)}\le\frac{f(B_0)}{g(B_0)}$. Then

$$\frac{f(A_0)+g(B_0)}{g(A_0)+g(B_0)}-\frac{f(A_0)}{g(A_0)}=\frac{f(B_0)g(A_0)-f(A_0)g(B_0)}{g(A_0)(g(A_0)+g(B_0))}\ge 0,$$

which follows that

$$\min_{(A,B)\in\mathcal{P}_2(V)}\frac{f(A)+f(B)}{g(A)+g(B)}=\frac{f(A_0)+g(B_0)}{g(A_0)+g(B_0)}\ge\frac{f(A_0)}{g(A_0)}\ge\min_{A\in\mathcal{P}(V)}\frac{f(A)}{g(A)}.$$

On the other hand, let $A_1\subset V$ be the minimizer of $\frac{f(A)}{g(A)}$. Then we have

$$\min_{(A,B)\in\mathcal{P}_2(V)}\frac{f(A)+f(B)}{g(A)+g(B)}\le\frac{f(A_1)+f(A_1^c)}{g(A_1)+g(A_1^c)}=\frac{f(A_1)}{g(A_1)}=\min_{A\in\mathcal{P}(V)}\frac{f(A)}{g(A)},$$

and hence,

$$\min_{A\in\mathcal{P}(V)}\frac{f(A)}{g(A)}=\min_{(A,B)\in\mathcal{P}_2(V)}\frac{f(A)+f(B)}{g(A)+g(B)}.$$

It is also true if we replace 'min' by 'max', i.e., (3.20) holds. □

Proof. **(Proof of Theorem 1.8.)** The proof involves F_1 and G_1. Let

$$f(A)=|\partial A|\ \text{ and }\ g(A)=\frac{1}{2}\mathrm{vol}(V).$$

Since f and g are symmetric functions, by Lemma 3.1 and Theorem 1.3, we have

$$\begin{aligned}h_{\max}(G)&=\max_{S\in\mathcal{P}(V)}\frac{2|\partial S|}{\mathrm{vol}(V)}=\max_{S\in\mathcal{P}(V)}\frac{f(S)}{g(S)}=\max_{(A,B)\in\mathcal{P}_2(V)}\frac{F_1(A,B)}{G_1(A,B)}\\&=\max_{(A,B)\in\mathcal{P}_2(V)\setminus\{(\varnothing,\varnothing)\}}\frac{F_1(A,B)}{G_1(A,B)}=\sup_{\boldsymbol{x}\neq\boldsymbol{0}}\frac{F_1^L(\boldsymbol{x})}{G_1^L(\boldsymbol{x})}.\end{aligned}$$

Accordingly, we have

$$h_{\max}(G)=\sup_{\boldsymbol{x}\neq\boldsymbol{0}}\frac{F_1^L(\boldsymbol{x})}{G_1^L(\boldsymbol{x})}=\sup_{\boldsymbol{x}\neq\boldsymbol{0}}\frac{I(\boldsymbol{x})}{\mathrm{vol}(V)\|\boldsymbol{x}\|_\infty}.$$

□

Proof. **(Proof of Theorem 1.9.)** Let $f(A)=|\partial A|$ and $g(A)=\min\{\mathrm{vol}(A),\mathrm{vol}(A^c)\}$. Since f, g are symmetric functions, by Equation (1.7), Lemma 3.1 and Proposition 1.1, we have

$$\begin{aligned}h(G)&=\min_{S\subset V}\frac{f(S)}{g(S)}=\min_{(A,B)\in\mathcal{P}_2(V)\setminus\{(\varnothing,\varnothing),(\varnothing,V),(V,\varnothing)\}}\frac{F_1(A,B)}{G_2(A,B)}\\&=\inf_{\boldsymbol{x}\text{ nonconstant}}\frac{F_1^L(\boldsymbol{x})}{G_2^L(\boldsymbol{x})}=\inf_{\boldsymbol{x}\text{ nonconstant}}\sup_{c\in\mathbb{R}}\frac{I(\boldsymbol{x})}{\sum_{i=1}^n d_i|x_i-c|},\end{aligned}$$

where F_1^L and G_2^L have been presented in Table 3.1. □

Proof. **(Proof of Theorem 1.10.)** Let

$$f(A)=|\partial A| \quad \text{and} \quad g(A)=\max\{\mathrm{vol}(A),\mathrm{vol}(A^c)\}. \tag{3.21}$$

Since f, g are symmetric functions, by Lemma 3.1 and Theorem 1.3, we obtain

$$\begin{aligned} h_{\text{anti}}(G) &= \max_{S\in\mathcal{P}(V)} \frac{f(S)}{g(S)} = \max_{(A,B)\in\mathcal{P}_2(V)\}} \frac{F_1(A,B)}{G(A,B)} \\ &= \max_{(A,B)\in\mathcal{P}_2(V)\setminus\{(\varnothing,\varnothing)\}\}} \frac{F_1(A,B)}{G(A,B)} = \sup_{\boldsymbol{x}\neq\boldsymbol{0}} \frac{F_1^L(\boldsymbol{x})}{G^L(\boldsymbol{x})}, \end{aligned}$$

where $G(A,B)=2G_1(A,B)-G_3(A,B)$, and $G^L=2G_1^L-G_3^L$. Thus,

$$h_{\text{anti}}(G)=\sup_{\boldsymbol{x}\neq\boldsymbol{0}} \frac{F_1^L(\boldsymbol{x})}{2G_1^L(\boldsymbol{x})-G_3^L(\boldsymbol{x})} = \sup_{\boldsymbol{x}\neq\boldsymbol{0}} \frac{I(\boldsymbol{x})}{2\mathrm{vol}(V)\|\boldsymbol{x}\|_\infty - \min\limits_{\alpha\in\mathbb{R}}\|\boldsymbol{x}-\alpha\mathbf{1}\|}.$$

□

Acknowledgements. This research was supported by the National Key R&D Program of China (No. 2020AAA0105200), the National Natural Science Foundation of China (Nos. 11822102, 11421101). SS is partially supported by Beijing Academy of Artificial Intelligence (BAAI). DZ is partially supported by the Project funded by China Postdoctoral Science Foundation (No. BX201700009). The authors are sincerely grateful to the handling editor and referees for their patience and valuable suggestions.

REFERENCES

[1] F. Bach, *Learning with submodular functions: A convex optimization perspective*, Found. Trends Mach. Learn., 6:145–373, 2013. 2.1, 2.2

[2] A. Billionnet, S. Elloumi, and A. Lambert, *Extending the QCR method to general mixed-integer programs*, Math. Program., 131(1-2):381–401, 2012. 1

[3] K. Bérczi and A. Frank, *Variations for Lovász' submodular ideas*, in M. Grötschel and G.O.H. Katona (eds.), Building Bridges between Mathematics and Computer Science, Springer, 137–164, 2008. 2, 2.4

[4] F. Bauer and J. Jost, *Bipartite and neighborhood graphs and the spectrum of the normalized graph laplace operator*, Commun. Anal. Geom., 21:787–845, 2013. 1, 1.1

[5] T. Bühler, S.S. Rangapuram, S. Setzer, and M. Hein, *Constrained fractional set programs and their application in local clustering and community detection*, in Proceedings of the 30th International Conference on Machine Learning, 624–632, 2013. 1, 1.7, 2.1, 2.1

[6] S. Burer, *On the copositive representation of binary and continuous nonconvex quadratic programs*, Math. Program., 120(2):479–495, 2009. 1, 1

[7] K.C. Chang, *Spectrum of the* 1*-Laplacian and Cheeger's constant on graphs*, J. Graph Theor., 81:167–207, 2016. 1

[8] K.C. Chang, S. Shao, and D. Zhang, *The 1-Laplacian Cheeger cut: Theory and algorithms*, J. Comput. Math., 33:443–467, 2015. 1, 1

[9] K.C. Chang, S. Shao, and D. Zhang, *Spectrum of the signless* 1*-Laplacian and the dual Cheeger constant on graphs*, arXiv preprint, arXiv:1607.00489, 2016. 1, 1

[10] K.C. Chang, S. Shao, D. Zhang, and W. Zhang, *A simple iterative algorithm for maxcut*, arXiv preprint, arXiv:1803.06496v1, 2018. 1, 1

[11] A. Frieze and M. Jerrum, *Improved approximation algorithms for MAX k-CUT and MAX BISECTION*, Algorithmica, 18:67–81, 1997. 1.3, 3.1

[12] M.X. Goemans and D.P. Williamson, *Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming*, J. Assoc. Comput. Mach., 42:1115–1145, 1995. 1

[13] M. Hein and T. Bühler, *An inverse power method for nonlinear eigenproblems with applications in 1-spectral clustering and sparse PCA*, Adv. Neur. Inf. Process. Syst., 847–855, 2010. 1

[14] R. Horst and N.V. Thoai, *DC programming: Overview*, J. Optimiz. Theory App., 103:1–43, 1999. 1
[15] R.M. Karp, *Reducibility among combinatorial problems*, in R.E. Miller, J.W. Thatcher, and J.D. Bohlinger (eds.), Complexity of Computer Computations, Springer, 85–103, 1972. 1.6
[16] L. Lovász, *Submodular functions and convexity*, in A. Bachem, M. Grötschel, and B. Korte (eds.), Mathematical Programming: The State of the Art, Springer, 235–257, 1983. 1, 2.1, 2, 2.3, 2.3
[17] E. Lawler and D. Wood, *Branch-and-bound methods: A survey*, Oper. Res., 14(4):699–719, 1966. 1
[18] L. Qi, *Directed submodularity, ditroids and directed submodular flows*, Math. Program., 42:579–599, 1988. 2.3
[19] H.D. Sherali and W.P. Adams, *A hierarchy of relaxations between the continuous and convex hull representations for zero-one programming problems*, SIAM J. Discrete Math., 3(3):411–430, 1990. 1, 1
[20] S. Schaible and T. Ibaraki, *Fractional programming*, Eur. J. Oper. Res., 12(4):325–338, 1983. 1, 1
[21] L. Trevisan, *Max cut and the smallest eigenvalue*, SIAM J. Comput., 41:1769–1786, 2012. 1, 1.1
[22] B. Xu, *Graph partitions: Recent progresses and some open problems*, Adv. Math. (China), 45:1–20, 2016. 1.8