Stud. Univ. Babeş-Bolyai Math. 66(2021), No. 1, 211–222
DOI: 10.24193/subbmath.2021.1.17

An optimization problem for continuous submodular functions

Laszlo Csirmaz

Dedicated to Professor Gheorghe Moroşanu on the occasion of his 70th anniversary.

Abstract. Real continuous submodular functions, as a generalization of the corresponding discrete notion to the continuous domain, gained considerable attention recently. The analog notion for entropy functions requires additional properties: a real function defined on the non-negative orthant of $\mathbb{R}^n$ is entropy-like (EL) if it is submodular, takes zero at zero, non-decreasing, and has the Diminishing Returns property. Motivated by problems concerning the Shannon complexity of multipartite secret sharing, a special case of the following general optimization problem is considered: find the minimal cost of those EL functions which satisfy certain constraints. In our special case the cost of an EL function is the maximal value of the n partial derivatives at zero. Another possibility could be the supremum of the function range. The constraints are specified by a smooth bounded surface S cutting off a downward closed subset. An EL function is feasible if at the internal points of S the left and right partial derivatives of the function differ by at least one. A general lower bound for the minimal cost is given in terms of the normals of the surface S. The bound is tight when S is linear. In the two-dimensional case the same bound is tight for convex or concave S. It is shown that the optimal EL function is not necessarily unique. The paper concludes with several open problems.

Mathematics Subject Classification (2010): 90C26, 46N10, 49Q10.

Keywords: Continuous submodular optimization, entropy method, secret sharing.

1. Introduction

Continuous submodularity is a generalization of the discrete notion of submodularity to the continuous domain. It has gained considerable attention recently [2, 4] as efficient convex optimization methods can be extended to find the minimal and maximal value of special multivariable continuous submodular functions over a compact and convex domain. Such optimization algorithms have important applications in many areas of computer science and applied mathematics such as training deep neural networks [5], design of online experiments [6], or budget allocation [12]. For more information see [1].

Interestingly, the same class of continuous submodular functions arises when the continuous version of multipartite secret sharing schemes is considered. In classical secret sharing [3] each participant receives a piece of information – their *shares* – such that a qualified subset of participants can recover the secret from the shares they received, while unqualified subsets – based on their shares only – should have no information on the secret's value at all. In the multipartite case [8, 9] participants are in n disjoint groups, and members in the same group have equal roles. In particular, a qualified subset is described uniquely by the n numbers telling how many members this subset has from each group. The main question in secret sharing is the efficiency – also called complexity – of the scheme, which is typically defined as the worst-case ratio of the size of any of the shares (measured by their Shannon entropy) and the size of the secret. Keeping track of the total entropy of different subsets of shares, traditional entropy inequalities imply a lower bound on the complexity [9, 10] known as the *Shannon-bound.* No general method is known which would effectively determine, or even estimate, the Shannon bound for an arbitrary collection of qualified subsets, and numerical computation is intractable even for moderately sized problems. Investigating the same question in the continuous domain allows applying analytical tools, and results achieved this way might shed light on the discrete case. This paper, based partly on the last section of [7], is an attempt to initiate such a line of research.

No notion from secret sharing or from information theory will be used later as they only serve as motivation for the definitions. The family of real functions corresponding to the (normalized) multipartite entropy will be called *entropy-like* functions and abbreviated as EL. This function family is defined in Section 2; actually it is the family of pointed, increasing, submodular functions with the "Diminishing Returns" property, see [4].

The optimization problem corresponding to finding an optimal multipartite secret sharing scheme is discussed in Section 3. It differs from the well-studied optimization problem for submodular functions [2, 4], where some member of the continuous submodular function family is given, and the task is to find its maximal (minimal) value over a compact, convex set. In our case the optimization problem asks to find an EL function with the smallest cost satisfying certain constraints. Two cost functions are considered. The first one corresponds to the discrete worst case complexity discussed above, and it is the maximal partial derivative of the EL function at the origin. The second possibility is the supremum of the function range; it corresponds to another frequently investigated complexity measure in the discrete case: the total randomness used by the scheme. In Section 3 a general lower bound for the worst case complexity is given as Theorem 3.4. This bound is tight when the constraints are specified by some linear surface.

Section 4 presents results for the bipartite, two-dimensional case. General constructions show that the lower bound of Theorem 3.4 is also tight for strictly convex or strictly concave constraint curves. An alternate construction shows that the optimal

EL function is not necessarily unique. Finally, Section 5 concludes the paper with a list of open problems.

2. Submodular and entropy-like functions

A real function f defined on subsets of a set is *submodular* if

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

for arbitrary subsets A and B, see [2] and references therein. The same notion extended to an arbitrary lattice requires

$$f(A) + f(B) \geq f(A \wedge B) + f(A \vee B)$$

for any two lattice members A and B. In particular, the n-variable real function f is *submodular* if it is submodular in the lattice determined by the partial order on $\mathbb{R}^n$ defined by $x \leq y$ if and only if $x_i \leq y_i$ for all coordinates $1 \leq i \leq n$. In this case $x \wedge y = \min(x, y)$ and $x \vee y = \max(x, y)$ where minimization (maximization) is taken coordinatewise, and the submodularity condition rewrites to

$$f(x) + f(y) \geq f(\min(x, y)) + f(\max(x, y)).$$

Entropy-like real functions, also called EL functions, share additional properties with discrete Shannon entropy functions [13], and are defined as follows.

Defnition 2.1. The n-variable real function f is *entropy-like*, or EL function for short, if it satisfies properties (a) – (e) below.

(a) f is defined on the non-negative orthant $\mathbb{R}^n_{\geqslant 0} = \{x \in \mathbb{R}^n : x \geq 0\}$.
(b) f is submodular.
(c) $f(0) = 0$ (f is pointed).
(d) f is non-decreasing: if $0 \leq x \leq y$ then $f(x) \leq f(y)$.
(e) f has the "Diminishing Returns" property [4]. It means that for two points $0 \leq x \leq y$ differing only in their i-th coordinate, increasing that coordinate at x and also at y by the same amount ε, the gain at y is never bigger than the gain at x. Formally, if e_i is the i-th unit vector and $y = x + \lambda e_i$ for some $\lambda > 0$, then for every $\varepsilon > 0$,

$$f(x + \varepsilon e_i) - f(x) \geq f(y + \varepsilon e_i) - f(y). \tag{2.1}$$

The "Diminishing Returns" property models the natural expectation that adding one more unit of some resource contributes more in the case when one has less available amount of that resource.

The left and right partial derivatives of the n-variable function f at $x \in \mathbb{R}^n$ are denoted by $f_i^-(x)$ and $f_i^+(x)$, respectively, and their definition goes as

$$f_i^-(x) = \lim_{\varepsilon \to +0} \frac{f(x) - f(x - \varepsilon e_i)}{\varepsilon}$$

and

$$f_i^+(x) = \lim_{\varepsilon \to +0} \frac{f(x + \varepsilon e_i) - f(x)}{\varepsilon}$$

assuming that the corresponding limits exist. Here e_i is the i-th unit vector.

The following claim summarizes some basic properties of EL functions.

Claim 2.2. *Let f be an n-variable EL function.*

(a) *f is continuous.*

(b) *f is concave along any positive direction: if $0 \le x \le y$ and $0 \le \lambda \le 1$ then*

$$\lambda f(x) + (1-\lambda) f(y) \le f(\lambda x) + (1-\lambda)y).$$

(c) *The Diminishing Returns property (2.1) holds for arbitrary pair of points $0 \le x \le y$.*

(d) *f has both left and right partial derivatives at every point of its domain.*

(e) *The partial derivatives are non-negative and non-increasing along any positive direction.*

Proof. (a) It is enough to show that f is continuous along every coordinate. By property (d) it is monotone increasing. The left limit $\lim_{\varepsilon\to+0} f(x-\varepsilon e_i)$ cannot be strictly smaller than the right limit $\lim_{\varepsilon\to+0} f(x+\varepsilon e_i)$ as this would contradict the Diminishing Returns property.

(b) Continuity and the Diminishing Returns property ensures that f is concave along each coordinate. It means that statement (b) is true when points x and y share $n-1$ coordinates. Suppose we have two points sharing i coordinates, and the claim has been established for point pairs sharing $i+1$ or more coordinates. Denote these points by (c,x,a) and (d,y,a) where a stands for the joint i coordinates, x and y are real numbers, and c and d are the remaining tuples. The linear combination $\lambda(c,x,a) + (1-\lambda)(d,y,a)$ is shortened to $(c \mathbin{\phi} d, x \mathbin{\phi} y, a)$. Using $(c,x,a) \le (d,y,a)$ and the induction hypothesis for $n-1$ (first line) and for $i+1$ (next two lines) we have

$$\begin{aligned}
\lambda f(c \mathbin{\phi} d, x, a) + (1-\lambda) f(c \mathbin{\phi} d, y, a) &\le f(c \mathbin{\phi} d, x \mathbin{\phi} y, a), \\
\lambda f(c, x, a) + (1-\lambda) f(d, x, a) &\le f(c \mathbin{\phi} d, x, a), \\
\lambda f(c, y, a) + (1-\lambda) f(d, y, a) &\le f(c \mathbin{\phi} d, y, a).
\end{aligned}$$

From here the required inequality

$$\lambda f(c,x,a) + (1-\lambda) f(d,y,a) \le f(c \mathbin{\phi} d, x \mathbin{\phi} y, a)$$

follows as the submodularity for the points (c,y,a) and (d,x,a) gives

$$f(c,y,a) + f(d,x,a) \ge f(c,x,a) + f(d,y,a).$$

(c) Similarly to (b) by induction on how many coordinates x and y have in common. Observe that if x and y do not differ at their i-th coordinate then (2.1) is equivalent to submodularity.

(d) This is immediate as f is continuous and non-decreasing.

(e) Non-negativity is clear. Monotonicity: if $x \le y$ then, for example,

$$\begin{aligned}
f_i^+(x) &= \lim_{\varepsilon\to+0} \frac{f(x+\varepsilon e_i) - f(x)}{\varepsilon} \\
&\ge \lim_{\varepsilon\to+0} \frac{f(y+\varepsilon e_i) - f(y)}{\varepsilon} = f_i^+(y),
\end{aligned}$$

where the inequality follows from (c). Other cases are similar. □

The next lemma follows easily from the fact that along each coordinate f is increasing and concave, and is given without proof.

Lemma 2.3. *If $\varepsilon \to +0$, then $f_i^+(x+\varepsilon e_i) \to f_i^+(x)$, and $f_i^+(x-\varepsilon e_i) \to f_i^-(x)$.* □

Remark 2.4. The family of EL functions is closed for non-negative linear combination and truncation: if f_1, f_2 are EL, then so is $\lambda_1 f_1 + \lambda_2 f_2$ for λ_1, $\lambda_2 \geq 0$; if f is EL and $M \geq 0$ then $\min(f, M)$ is EL. Consequently

$$f(x) = \min\big(\textstyle\sum c_i x_i, M\big)$$

is EL for positive c_i and M. Similarly, if f is EL and $a \geq 0$, then $g(x) = f(\min(x,a))$ is EL again. Further examples of EL functions will be given in Section 4.

Remark 2.5. If the sequence f_k of EL functions converge pointwise, then the limit f is also an EL function, moreover

$$f_i^+(x) \leq \liminf_k \,(f_k)_i^+(x) \leq \limsup_k \,(f_k)_i^-(x) \leq f_i^-(x).$$

3. The optimization problem

According to the intuition discussed in Section 1 the value of n-variable EL function f at $x \in \mathbb{R}^n_{\geqslant 0}$ can be considered as the value of the (scaled) entropy of the set of shares assigned to a subset of participants which has members from the i-th group proportional to the i-th coordinate of x. The right derivative $f_i^+(x)$ can be interpreted as the (scaled) entropy increase if one more member from the i-th group joins this subset, and $f_i^-(x)$ as the entropy decrease when one member from the i-th group leaves the subset (defined only if $x_i > 0$). Consequently the share size of a single participant from group i can be identified to $f_i^+(0)$, the i-th right partial derivative of f at zero. Accordingly, the cost function corresponding to the maximal share size is

$$\mathrm{Cost}(f) = \max\{f_1^+(0), f_2^+(0), \ldots, f_n^+(0)\,\}.$$

While this cost function will be considered in this paper, there are other possibilities. In the discrete cases the total entropy (the amount of randomness needed to generate the whole scheme) is used frequently, this would correspond to the cost function $\sup\{f(x) : x \in \mathrm{Dom}(f)\}$.

In secret sharing the shares of a qualified subset determine the secret, while the same secret is (statistically) independent of the shares of an unqualified subset. We call the point $x \in \mathbb{R}^n_{\geqslant 0}$ *qualified* if the corresponding subset is qualified. When decreasing an unqualified subset it remains unqualified, thus the set of unqualified points are downward closed: if x is unqualified and $0 \leq y \leq x$ then y is unqualified as well. Suppose the unqualified and qualified points are separated by the smooth $(n-1)$-dimensional surface S. Downward closedness means that the normal vectors of S pointing outwards (towards qualified points) have non-negative coordinates. This surface S specifies the secret sharing problem, namely which subsets of the participants are qualified and which are not, and thus the optimization problem as well. The definition below requires slightly stronger properties from such a separating surface excluding certain problematic cases.

Defnition 3.1. An *s-surface* (secret sharing surface) is a smooth $(n-1)$-dimensional surface S in the non-negative orthant $\mathbb{R}^n_{\geqslant 0}$ satisfying the following properties:

(a) S avoids 0,
(b) S is compact, and
(c) for every $x \in S$ the normal vector $\nabla S(x)$ pointing outwards has strictly positive coordinates.

Consider the subset of participants which corresponds to the point $x \in S$ of the s-surface S. If any member from the i-th group leaves this subset, then the subset becomes unqualified – and then the secret must be independent of the joint collection of the associated shares. If any new member from the i-th group joins that subset, it becomes qualified – meaning that the new share collection determines the secret. Thus the difference between the before and after entropy changes, namely $f_i^-(x) - f_i^+(x)$, must cover the entropy of the secret. The entropy of the secret can be taken to be 1 as this changes all values up to a scaling factor only. The following definition summarizes this discussion.

Defnition 3.2. The EL function f is *feasible* for S, or S-feasible, if for every positive $x \in S$ (that is, $x_i > 0$ for all $1 \le i \le n$),

$$f_i^-(x) - f_i^+(x) \ge 1 \ (1 \le i \le n). \tag{3.1}$$

(Positivity of x is ensures the existence of $f_i^-(x)$.)

Optimization problems considered in this paper are of this form: given the s-surface S, find the minimal cost of the S-feasible functions.

Defnition 3.3. For a given s-surface $S \subseteq \mathbb{R}^n_{\geqslant 0}$ OPT(S) is the optimization problem

$$\begin{cases} \text{minimize:} & \mathrm{Cost}(f) \\ \text{subject to:} & f \text{ is an } S\text{-feasible EL function.} \end{cases}$$

By an abuse of notation, both the problem and its solution – the infimum of the costs of S-feasible functions – will be denoted by OPT(S).

As an example let us consider the case when S is the intersection of the hyperplane

$$c_1x_1 + c_2x_2 + \cdots + c_nx_n = M$$

and the non-negative orthant, here c_i and M are positive constants. Observe that the normal at every $x \in S$ is $\nabla S(x) = (c_1, \ldots, c_n)$. Feasible EL functions will be searched among the one-parameter family

$$f(y) = k \cdot \min\left\{\textstyle\sum c_iy_i, M\right\}$$

with positive k. All of them are EL functions by Remark 2.4. Pick the positive point $x \in S$ and consider $f(x + \varepsilon e_i)$ as a function of ε. It has the constant value $k \cdot M$ for $\varepsilon \ge 0$, and it is linear with slope $k \cdot c_i$ for $\varepsilon \le 0$. Consequently

$$f_i^-(x) - f_i^+ = k \cdot c_i,$$

which is ≥ 1 if $k \geq 1/\min\{c_i\}$. At zero the partial derivatives of f are $k \cdot c_i$, therefore $\mathrm{Cost}(f) = k \cdot \max\{c_i\}$. The $k = 1/\min\{c_i\}$ choice gives an S-feasible EL function with cost $\max\{c_i\}/\min\{c_i\}$, thus

$$\mathrm{OPT}(S) \leq \frac{\max\{c_i\}}{\min\{c_i\}}.$$

According to Theorem 3.4 below the optimal value is actually equal to this amount, as in this case $\nabla S_i(x) = c_i$ for every $x \in S$.

Theorem 3.4. *For every s-surface S, inner point $x \in S$ and $1 \leq i, j \leq n$ the following inequality holds:*

$$\mathrm{OPT}(S) \geq \frac{\nabla S_j(x)}{\nabla S_i(x)}.$$

Proof. By assumption S behaves linearly on a small neighborhood of x, thus for every small enough positive w there is a unique positive h such that $y = x - we_i + he_j \in S$, and

$$\lim_{w \to +0} \frac{h}{w} = \frac{\nabla S_j(x)}{\nabla S_i(x)}.$$

Let f be any S-feasible EL function, $u = \min(x, y) = x - we_i$ and $v = \max(x, y) = x + he_j$. The following inequalities follow from the facts that f is monotone and concave along each coordinate by Claim 2.2:

$$\begin{aligned} w \cdot f_i^+(u) &\geq f(x) - f(u), \\ h \cdot f_j^+(x) &\geq f(v) - f(x), \\ f(y) - f(u) &\geq h \cdot f_j^-(y), \\ f(v) - f(y) &\geq 0. \end{aligned}$$

Their sum proves the first inequality in the sequence

$$\begin{aligned} w \cdot f_i^+(u) &\geq h\big(f_j^-(y) - f_j^+(x)\big) \\ &\geq h\big(1 + f_j^+(y) - f_j^+(x)\big) \\ &\geq h\big(1 + f_j^+(v) - f_j^+(x)\big). \end{aligned}$$

The second inequality follows from $y \in S$ and that f is an S-feasible function. The third one uses the monotonicity of the derivatives from Claim 2.2 (e). Letting $w \to +0$, $f_i^+(u) \to f_i^-(x)$ and $f_j^+(v) \to f_j^+(x)$ by Lemma 2.3, thus

$$f_i^-(x) \geq \frac{\nabla S_j(x)}{\nabla S_i(x)}.$$

From here the theorem follows as $\mathrm{Cost}(f) \geq f_i^+(0) \geq f_i^-(x)$ by the monotonicity of the derivatives. □

Theorem 3.5. *Suppose $\mathrm{OPT}(S) < +\infty$ for an s-surface S. The optimal value is taken by some S-feasible function f, that is, $\mathrm{Cost}(f) = \mathrm{OPT}(S)$.*

Proof. Let $\mathrm{OPT}(S) < M$, and choose the sequence of S-feasible functions f_k such that $\mathrm{Cost}(f_k) < M$ and $\lim_k \mathrm{Cost}(f_k) = \mathrm{OPT}(S)$. Also pick a point $a \in \mathbb{R}^n_{\geqslant 0}$ such that S is contained completely in the box $B = \{x \in \mathbb{R}^n_{\geqslant 0} : x \le a\}$. The functions $g_k(x) = f(\min(x, a))$ are EL by Remark 2.4, and $\mathrm{Cost}(g_k) = \mathrm{Cost}(f_k)$. Each g_k is clearly S-feasible and is bounded by $M\cdot(a_1+\cdots+a_n)$. The sequence $\{g_k\}$ is uniformly equicontinuous as all partial derivatives are bounded by M, thus the Arzelà–Ascoli theorem [11] guarantees a subsequence which converges uniformly on B – and then converges everywhere. Denote this subsequence also by $\{g_k\}$, and let the pointwise limit be g. By Remark 2.5 g is an EL function and $\mathrm{Cost}(g) \le \liminf_k \mathrm{Cost}(g_k) = \mathrm{OPT}(S)$. Also, each g_k is S-feasible, that is, at the points of S the difference between the left and right derivatives is at least 1:

$$(g_k)_i^-(x) - (g_k)_i^+(x) \ge 1,\ x \in S.$$

By Remark 2.5 the same is true for the limit function g. Thus there is an S-feasible function g with $\mathrm{Cost}(g) \le \mathrm{OPT}(S)$, which proves the theorem. □

4. Two-dimensional cases

We have seen that the bound provided by Theorem 3.4 is sharp when S is linear. We show that, at least in the two-dimensional case, it is also sharp when S is strictly convex or strictly concave by constructing matching S-feasible EL functions.

In two dimensions S is a strictly decreasing continuous curve. Write S as $\{(x, \alpha(x)) : 0 \le x \le a\}$, and also as $\{(\beta(y), y) : 0 \le y \le b\}$, see Figure 1.

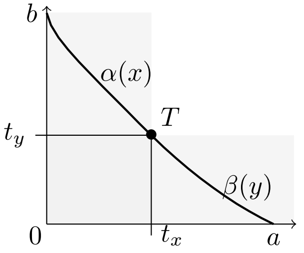

Figure 1. The curve S

If S is either convex or concave, then $\nabla S_i(x)/\nabla S_j(x)$ is increasing or decreasing along the curve, thus attains its maximal value at one of the endpoints.

First assume that S is strictly convex. In this case both α and β are convex functions. Let $T = (t_x, t_y)$ be the point on S where the normal is $(1, 1)$. On the $[0, t_x]$ interval the derivative $\alpha'(x)$ is ≤ -1, and, similarly, $\beta'(y) \le -1$ on $[0, t_y]$. The function f depicted on Figure 2 is defined as follows.

If both $x \ge t_x$ and $y \ge t_y$ then $f(x, y) = C$, otherwise

$$f(x,y) = \begin{cases} C + \min\{x - \beta(y), 0\} & \text{if } x \ge t_x, \\ C + \min\{y - \alpha(x), 0\} & \text{if } y \ge t_y, \\ a - \alpha(x) + b - \beta(y) & \text{otherwise,} \end{cases}$$

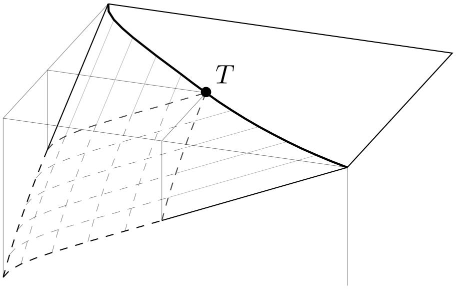

Figure 2. Convex case

where $C = a - t_x + b - t_y$. Clearly f has a flat plateau of height C beyond the curve S. It is a routine to check that f is an EL function; one has to use that $-\alpha(x)$ and $-\beta(y)$ are concave functions and have derivative 1 at $x = t_x$ and $y = t_y$, respectively. The left and right partial derivatives of f at $(x, y) \in S$ are $(1, 0)$ and $(-\beta'(y), 0)$ when $x \geq t_x$, and $(-\alpha'(x), 0)$ and $(1, 0)$ when $y \geq t_y$. In all cases the values in the pair differ by at least one, thus f is a feasible S-function. The partial derivatives of f at zero are $-\alpha'(0)$ and $-\beta'(0)$, thus

$$\mathrm{Cost}(f) = \max\{-\alpha'(0), -\beta'(0)\}$$

matching the lower bound of Theorem 3.4.

In the case when no point on S has normal $(1, 1)$ the simpler construction using only the first (or second) line in the definition of the function f works.

A different construction is illustrated on Figure 3 which also meets the lower bound of Theorem 3.4. It also shows that the optimal EL function, if exists, is not necessarily unique. Using the same notation as above,

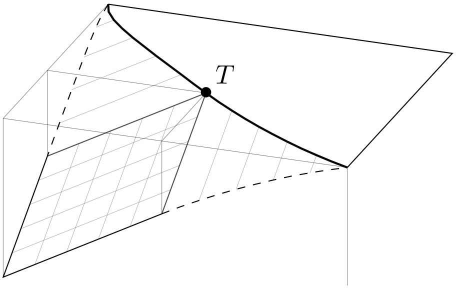

Figure 3. Alternate construction for the convex case

the function f is defined analogously: $f(x, y) = C$ if $x \geq t_x$ and $y \geq t_y$, otherwise

$$f(x,y) = \begin{cases} C + \min\{y - \alpha(x), 0\} & \text{if } x \geq t_x, \\ C + \min\{x - \beta(y), 0\} & \text{if } y \geq t_y, \\ x + y & \text{otherwise,} \end{cases}$$

where $C = t_x + t_y$. This is again an EL function, its cost is clearly 1. The difference between the left and right partial derivatives at points of S are $-\alpha'(x)$ and 1 when $x \geq t_x$, and 1 and $-\beta'(y)$ when $y \ge t_y$, thus the difference is at least $k = \min\{-\alpha'(a), -\beta'(b)\}$. Consequently the EL function $k^{-1}f(x,y)$ is feasible for S, and its cost, $1/k$, matches the lower bound in Theorem 3.4.

The third construction, depicted on Figure 4, works for any strictly concave curve S. In this case the plateau is not flat any more. Using the same notations as before, the decreasing functions $\alpha(x)$ and $\beta(y)$ are strictly concave, and $T = (t_x, t_y)$ is the curve point with normal $(1,1)$. The function $f(x,y)$ is defined as $f(x,y) = t_x + t_y$ if both $x \ge t_x$ and $y \ge t_y$, otherwise

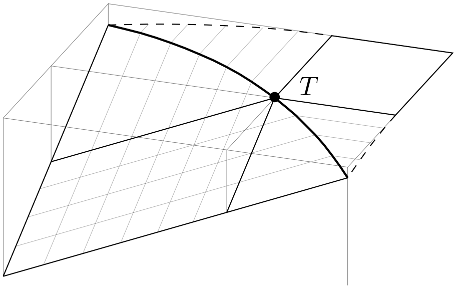

Figure 4. Concave case

$$f(x,y) = \begin{cases} y + \min\{x, \beta(y)\} & \text{if } x \ge t_x, \\ x + \min\{y, \alpha(x)\} & \text{if } y \ge t_y, \\ x + y & \text{otherwise.} \end{cases}$$

This is an EL function. For example, for a fixed $x \ge t_x$ it is increasing and concave as $y + \beta(y)$ is increasing on the $[0, t_y]$ interval ($\beta'(y) \le -1$ here), and is concave since β is concave. The left and right partial derivatives of f at a point (x,y) of S with $x \ge t_x$ are 1 and 0, and 1 and $1 + \beta'(y)$, respectively. The difference between the corresponding pairs is at least $-\beta'(y) \ge -\beta'(0)$. Choosing the multiplier k such that $k \cdot (-\alpha'(0)) \ge 1$ and $k \cdot (-\beta'(0)) \ge 1$, the EL function $k \cdot f$ will be S-feasible. The minimal such k gives a cost k EL function which again matches the lower bound of Theorem 3.4.

5. Conclusion

A continuous version of the discrete Shannon entropy functions, called *entropy-like*, or EL functions, has been defined in Definition 2.1. They form a natural subclass of multivariate continuous submodular functions which gained considerable attention recently [2]. Interestingly, the same subclass emerged as a crucial one when investigating possible parallelization of traditional submodular optimization algorithms [4].

Motivated by difficult problems in multipartite secret sharing [8], points in the non-negative orthant are flagged as either qualified or unqualified, separated by a *secret sharing surface* S, see Definition 3.1. An EL function is *feasible* for such a surface S if at internal points of S all partial derivatives drop by at least one when passing from left to right. The following optimization problem was considered: for a given s-surface S find the minimal cost of an S-feasible EL function. The first open problem is to prove that this function set is never empty.

Problem 5.1. Prove that for every s-surface S there exists at least one S-feasible function.

The cost of an EL function f is the maximum of its partial derivatives at zero, thus it can be $+\infty$. Definition 3.1 stipulates that for every S-surface there is a positive constant c such that $1/c < \nabla S_i(x) < c$ at each point $x \in S$. The value in Theorem 3.4 bounding the cost of any S-feasible function from below is smaller than c^2, thus it does not exclude the following strengthening of Problem 5.1:

Problem 5.2. Prove that for every s-surface S there is at least one S-feasible function with *finite cost.*

The lower bound on the cost of S-feasible EL functions proved in Theorem 3.4 was shown to be tight for linear s-surfaces, and also for two-dimensional convex and concave s-surfaces.

Problem 5.3. Find an s-surface S for which the bound in Theorem 3.4 is not tight.

As a strenghtening of Problem 5.3 we offer a bold conjecture which might easily turn out to be false.

Problem 5.4. If S is neither convex nor concave, then the bound of Theorem 3.4 is not tight.

Constructions in Section 4 settled the problem of finding the optimal values for two-dimensional convex and concave s-surfaces. It would be interesting to see optimal solutions for convex and concave surfaces in higher dimensions.

Problem 5.5. Determine the optimal costs of convex and concave s-surfaces in dimension > 2.

As mentioned in Section 3, the cost function considered in this paper stems from the *worst case complexity* of general secret sharing schemes. An alternate cost function corresponding to the total entropy would be $\mathrm{Cost}^t(f) = \sup\{\, f(x) : x \in \mathrm{Dom}(f)\,\}$. As an EL function can be truncated, the sup here can be limited to the points of S. The two costs functions are obviously related, but it is not clear how this relationship can be used to connect the corresponding optimization problems.

Problem 5.6. Prove lower bounds, similar to Theorem 3.4, for the optimization problem $\mathrm{OPT}^t(S)$ using the Cost^t function.

By Theorem 3.5, if there is any S-feasible function at all then there is one with minimal cost. The proof relied on the fact that finite cost EL functions have bounded derivatives. For Cost^t this property does not hold anymore.

Problem 5.7. If there is an S-feasible function, then there is one with

$$\mathrm{Cost}^t(f) = \mathrm{OPT}^t(S).$$

Finally, extend the quite meager collection of s-surfaces from Section 4 for which the exact bound is known.

Problem 5.8. Find optimal solutions for additional "interesting" s-surfaces for both cost functions.

Acknowledgment. The work of the author was supported by the GACR grant number 19-045798, and by the Lendület program of the HAS.

References

[1] Bach, F., *Learning with Submodular Functions: A Convex Optimization Perspective*, Foundations and Trends in Machine Learning, 2013.

[2] Bach, F., *Submodular functions: from discrete to continuous domains*, Math. Program., **175**(2019), no. 1-2, 419-459.

[3] Beimel, A., *Secret-sharing schemes: A survey*, In: Proceedings of the Third International Conference on Coding and Cryptology, IWCC'11, Springer-Verlag, Berlin, Heidelberg, 2011, 11-46.

[4] Bian, Y., Buhmann, J.M., Krause, A., *Continuous submodular function maximization*, arXiv preprint arXiv:2006.13474, 2020.

[5] Botton, L., Curtis, F.E., Nocedal, J., *Optimization methods for large-scale machine learning*, Siam Review, **60**(2018), no. 2, 223-311.

[6] Chen, L., Hassani, H., Karbasi, A., *Online continuous submodular maximization*, In: AISTATS, Playa Blanca, Lanzarote, Canary Islands, 2018.

[7] Csirmaz, L., Matús, F., Padró, C., *Bipartite secret sharing and staircases*, Unpublished manuscript, 2020.

[8] Farràs, O., Martí-Farré, J., Padró, C., *Ideal multipartite secret sharing schemes*, J. Cryptology, **25**(2012), no. 3, 434-463.

[9] Farràs, O., Metcalf-Burton, J.R., Padró, C., Vázquez, L., *On the optimization of bipartite secret sharing schemes*, Des. Codes Cryptog., **63**(2012), no. 2, 255-271.

[10] Metcalf-Burton, J.R., *Information rates of minimal non-matroid-related access structures*, CoRR, 2008.

[11] Rudin, W., *Principles of Mathematical Analysis*, McGraw-Hill Book Co., New York, 3rd edition, 1976.

[12] Staib, M., Jegelka, S., *Robust budget allocation via continuous submodular functions*, In: Doina Precup and Yee Whye Teh, editors, Proceedings of the 34th International Conference on Machine Learning, ICML 2017, Sydney, NSW, Australia, 6-11 August 2017, vol. 70 of Proceedings of Machine Learning Research, 3230-3240.

[13] Yeung, R.W., *A First Course in Information Theory*, Information Technology: Transmission, Processing and Storage, Springer US, 2012.

Laszlo Csirmaz
Alfréd Rényi Mathematical Institute, Budapest, Hungary
and
Institute of Information Theory and Automation,
The Czech Academy of Sciences, Pod Vodáreskou věží 4, 182 08 Praha 8, Czech Republic
e-mail: `csirmaz@renyi.hu`