Efficient algorithm for computing certain graph-based monotone integrals: the l_p-indices

Marek Gagolewski *†‡ Michał Dębski § Michał Nowakiewicz †

Abstract: *The Choquet, Sugeno and Shilkret integrals with respect to monotone measures are useful tools in decision support systems. In this paper we propose a new class of graph-based integrals that generalize these three operations. Then, an efficient linear-time algorithm for computing their special case, that is l_p-indices, $1 \le p < \infty$, is presented. The algorithm is based on R.L. Graham's routine for determining the convex hull of a finite planar set.*

Keywords: *monotone measures, Choquet, Sugeno and Shilkret integral, l_p-index, convex hull, Graham's scan, scientific impact indices.*

1 Introduction

Many practical situations, especially in decision making, face us with the problem of aggregating numeric sequences not necessarily having equal lengths. For example, we observe a gradually increasing interest in developing objective and fair research performance evaluation methods of individual scientists by means of citations number of authored papers. Scientometricians try to find quantitative indicators that might complement, or even replace, expert judgment. Such tools could be used in deciding upon employment, grant allocation, etc.

Let us assume that we are given a set of vectors with elements in $\mathbb{I} = [0, \infty]$ and we would like to construct an aggregation operator F defined on the space $\mathbb{I}^{1,2,\cdots} := \bigcup_{i=1}^{\infty} \mathbb{I}^n$, and such that it is nondecreasing in each variable, arity-monotonic, and symmetric, cf. [6]. The most common approach is to assume that F is zero-insensitive, i.e. that it holds $\mathsf{F}(\mathbf{x}) = \mathsf{F}(\mathbf{x}, 0)$. It may be shown, cf. [7], that in such setting, a vector $\mathbf{x} \in \mathbb{I}^{1,2,\cdots}$ may be projected to the space $\mathcal{S}$ of infinite-length, nonincreasing vectors, $\widetilde{\mathbf{x}} = (x_{\{1\}}, x_{\{2\}}, \dots, x_{\{n\}}, 0, 0, \dots)$, where $x_{\{i\}}$ denotes the ith greatest value in $\mathbf{x}$, and then the construction of F is equivalent to considering $\mathsf{E} : \mathcal{S} \to \mathbb{I}$, $\mathsf{E}(0, 0, \dots) = 0$, such that for all $\mathbf{x} \in \mathbb{I}^{1,2,\cdots}$ we have $\mathsf{F}(\mathbf{x}) = \mathsf{E}(\widetilde{\mathbf{x}})$.

In a very recent paper [7] we considered a uniform framework for the scientific impact assessment problem (and similar issues), where we have shown that most currently used bibliometric impact indices may be expressed by some universal integrals [10], see also [1, 2, 14] for other applications of monotone measures and integrals in scientometrics.

Recently, a very interesting class of so-called decomposition integrals [11] has been proposed. Some of these objects have a very nice graphical interpretation, which may be very important for the practitioners.

In this paper we propose another class of integrals that generalize the Sugeno, Choquet and Shilkret integrals, as well as some decomposition integrals. For example, they include the so-called "geometric" scientific impact indices proposed in [5]. Moreover, we introduce a linear-time algorithm for computing the l_p-indices and thus solve the open problem stated in [5]. The algorithm is an appealing modification of Graham's routine for the convex hull of a finite planar set [8].

*Corresponding author, *gagolews@ibspan.waw.pl*
†Systems Research Institute, Polish Academy of Sciences, Warsaw, Poland
‡Faculty of Mathematics and Information Science, Warsaw University of Technology, Warsaw, Poland
§Faculty of Mathematics, Informatics and Mechanics, University of Warsaw, Poland

2 Monotone measures and integrals

2.1 Monotone measures

Let $(Z, \mathcal{A})$ be a measurable space, i.e. a nonempty set Z equipped with a σ-algebra. We call $\mu : \mathcal{A} \to \mathbb{I}$ a *monotone measure* (a capacity), if (a) $\mu(\emptyset) = 0$, (b) $\mu(Z) > 0$, and (c) $\mu(U) \leq \mu(V)$ for $U \subseteq V$. Note that a monotone measure is not necessarily (σ-)additive. Moreover, let $\mathcal{M}^{(Z,\mathcal{A})}$ denote the set of all monotone measures.

A function $\mathsf{f} : Z \to \mathbb{I}$ is called $\mathcal{A}$-*measurable* if for each T in the σ-algebra of Borel subsets of $\mathbb{I}$, the inverse image $\mathsf{f}^{-1}(T) \in \mathcal{A}$. Let $\mathcal{F}^{(Z,\mathcal{A})}$ denote the set of all $\mathcal{A}$-measurable functions $\mathsf{f} : Z \to \mathbb{I}$.

2.2 Integrals

Let $\{u : \mathsf{f}(u) \geq t\} \in \mathcal{A}$ denote the so-called t-level set of $\mathsf{f} \in \mathcal{F}^{(Z,\mathcal{A})}$, $t \in \mathbb{I}$. It is easily seen that $\{u : \mathsf{f}(u) \geq t\}_{t\in\mathbb{I}}$ forms a left-continuous, nonincreasing chain (w.r.t. t). Thus, $\mathsf{h}^{(\mu,\mathsf{f})}(t) := \mu(\{u : \mathsf{f}(u) \geq t\})$ is a nonincreasing function of t, $\mathsf{h}^{(\mu,\mathsf{f})} : \mathbb{I} \to \mathbb{I}$.

Which function shall be called an integral of $\mathsf{f} \in \mathcal{F}^{(Z,\mathcal{A})}$ is still a disputable issue. Generally, it is agreed that an integral should map the space $\mathcal{M}^{(Z,\mathcal{A})} \times \mathcal{F}^{(Z,\mathcal{A})}$ into $\mathbb{I}$, should be at least nondecreasing with respect to each coordinate, and for $\mathsf{f} \equiv 0$ it should "output" the value 0.

Often integrals are defined as a function $\mathcal{I} : \mathcal{M}^{(Z,\mathcal{A})} \times \mathcal{F}^{(Z,\mathcal{A})} \to \mathbb{I}$ given by:

$$\mathcal{I}(\mu, \mathsf{f}) = \mathcal{J}(\mathsf{h}^{(\mu,\mathsf{f})}),$$

where $\mathcal{J} : \mathcal{F}^{(\mathbb{I},\mathcal{B}(\mathbb{I}))} \to \mathbb{I}$ is nondecreasing, $\mathcal{J}(0) = 0$.

For example, universal integrals, thoroughly discussed in [10], fulfill additional condition that for each $c, d \in \mathbb{I}$ we have $\mathcal{J}(d \cdot \mathbf{I}_{(0,c]}) = c \otimes d$, where $\otimes$ is some pseudo-multiplication.

2.3 Graph-based integrals

Let $\mathrm{Gr}(\mathsf{h}^{(\mu,\mathsf{f})}) = \{(x, y) \in \mathbb{I}^2 : y < \mathsf{h}^{(\mu,\mathsf{f})}(x)\}$. Recall that the Choquet integral is given by

$$\mathrm{Ch}(\mu, \mathsf{f}) = \int_{\mathbb{I}} \mu(\{u : \mathsf{f}(u) \geq t\})\, dt.$$

It may easily be shown that if $\mathrm{Gr}(\mathsf{h}^{(\mu,\mathsf{f})})$ is bounded and measurable, then $\mathrm{Ch}(\mu, \mathsf{f}) = \iint_{\mathbb{I}^2} d\mathrm{Gr}(\mathsf{h}^{(\mu,\mathsf{f})})$. Inspired with this fact and the notion of an decomposition integral [11], we introduce the so-called graph-based integrals.

Let $\mathcal{H} \subseteq 2^{2^{\mathbb{I}^2}}$, $\mathcal{H} \neq \emptyset$ be such that for all $\mathcal{P} \in \mathcal{H}$ it holds $p \cap p' = \emptyset$ for all $p, p' \in \mathcal{P}$, $p \neq p'$, i.e. it is a system of sets of disjoint Lebesgue-measurable subsets of $\mathbb{I}^2$. We define the **graph-based integral** corresponding to $\mathcal{H}$ as:

$$\mathrm{Gbi}_{\mathcal{H}}(\mu, \mathsf{f}) = \sup \left\{ \sum_{p\in\mathcal{P}} \lambda(p) : \mathcal{P} \in \mathcal{H}, \bigcup_{p\in\mathcal{P}} p \subseteq \mathrm{Gr}(\mathsf{h}^{(\mu,\mathsf{f})}) \right\}, \tag{1}$$

where λ is the Lebesgue measure, here in the space $(\mathbb{I}^2, \mathcal{B}(\mathbb{I}^2))$.

Intuitively, the calculation of a graph-based integral is done by finding the total area of the maximal "subcover" of $\mathrm{Gr}(\mathsf{h}^{(\mu,\mathsf{f})})$ by shapes from $\mathcal{H}$.

Here are some worth-noting instances of graph-based integrals.

Example 1. Let $\mathcal{H} = \{\{[x_1, x_2] \times [y_1, y_2]\}_{0\leq x_1\leq x_2, 0\leq y_1\leq y_2}\}$, i.e. each element of $\mathcal{H}$ is a set consisting of exactly one rectangle in $\mathbb{I}^2$. Then $\mathrm{Gbi}_{\mathcal{H}}(\mu, \mathsf{f})$ is equivalent to the Shilkret integral [12], $\mathrm{Sh}(\mu, \mathsf{f}) = \sup_{t\in\mathbb{I}}\{t \cdot \mu\{u : \mathsf{f}(u) \geq t\}\}$.

Example 2. Let $\mathcal{H} = \{\{[0,r] \times [0,r]\}_{0 \le r}\}$. Then $\sqrt{\mathrm{Gbi}_{\mathcal{H}}(\mu, \mathsf{f})}$ is equivalent to the Sugeno integral [13], $\mathrm{Su}(\mu, \mathsf{f}) = \sup_{t \in \mathbb{I}}\{t \wedge \mu\{u : \mathsf{f}(u) \ge t\}\}$. The same holds e.g. for $\mathcal{H} = \{\{[x, x+r] \times [y, y+r]\}_{0 \le r, 0 \le x, 0 \le y}\}$.

Example 3. Let $\mathcal{H}_k = \{\{[x_{1,i}, x_{2,i}] \times [y_{1,i}, y_{2,i}]\}_{i=1,\dots,k, 0 \le x_{1,i} \le x_{2,i}, 0 \le y_{1,i} \le y_{2,i}}\}$, such that for all $\mathcal{P} \in \mathcal{H}_k$, $p \cap p' = \emptyset$ for $p, p' \in \mathcal{P}$, $p \ne p'$, i.e. $\mathcal{P}$ is a set of k disjoint rectangles. Then each $\mathrm{Gbi}_{\mathcal{H}_k}(\mu, \mathsf{f})$ is an universal decomposition integral as defined in [11, Def. 4.4]. Moreover, $\lim_{k \to \infty} \mathrm{Gbi}_{\mathcal{H}_k}(\mu, \mathsf{f}) = \mathrm{Ch}(\mu, \mathsf{f})$, i.e. the Choquet integral [3].

2.4 The uniform model for bibliometric impact assessment

In [7] Gagolewski and Mesiar presented the following uniform model for bibliometric impact assessment problem. First of all, we need a transformation from the vector space $\mathcal{S}$ into the space $\mathcal{F}^{(\mathbb{I}, \mathcal{B}(\mathbb{I}))}$. Given $\mathbf{x} \in \mathcal{S}$, let $\langle \mathbf{x} \rangle \in \mathcal{F}^{(\mathbb{I}, \mathcal{B}(\mathbb{I}))}$ such that

$$\langle \mathbf{x} \rangle(t) = x_{\lfloor t+1 \rfloor}, \quad t \in \mathbb{I}.$$

Let us consider the family Φ of aggregation operators $\mathsf{F} : \mathcal{S} \to \mathbb{I}$ given by the equation:

$$\mathsf{F}(\mathbf{x}) = \eta\Big(\mathcal{I}\big(\mu, \langle \varphi(\mathbf{x}) \rangle\big)\Big) \tag{2}$$

where:

- $\varphi : \mathcal{S} \to \mathcal{S}$ – a function nondecreasing in each variable, $\varphi(0, 0, \dots) = (0, 0, \dots)$,
- $\mu : \mathcal{B}(\mathbb{I}) \to [0, \infty]$ – a monotone measure,
- $\mathcal{I}$ – an integral on $\mathcal{M}^{(\mathbb{I}, \mathcal{B}(\mathbb{I}))} \times \mathcal{F}^{(\mathbb{I}, \mathcal{B}(\mathbb{I}))}$,
- $\eta : \mathbb{I} \to \mathbb{I}$ – an increasing function, $\eta(0) = 0$.

It may be shown that an aggregation operator F may be expressed as (2) if and only if it is a zero-insensitive impact function, see [7]. Moreover, $\mathsf{h}^{(\mu, \langle \varphi(\mathbf{x}) \rangle)}$ is a nonincreasing step function.

Example 4. For $p \ge 1$ let $\mathcal{H}_p = \{\{\mathrm{B}_p(r)|_{\mathbb{I}^2}\}_{r \ge 0}\}$, where $\mathrm{B}_p(r) = \{(x, y) : \|(x, y)\|_p \le r\}$, i.e. it is a ball of radius r w.r.t. L_p distance, centered at $(0, 0)$. Then for μ being a Lebesgue measure and $\varphi = \mathrm{id}$, $\lfloor \sqrt{\mathrm{Gbi}_{\mathcal{H}_\infty}(\mu, \langle \varphi(\mathbf{x}) \rangle)} \rfloor$ is equivalent to the h-index [9], $\lfloor \sqrt{2\, \mathrm{Gbi}_{\mathcal{H}_1}(\mu, \langle \varphi(\mathbf{x}) \rangle)} \rfloor$ is the w-index [15] and, more generally, $\lfloor \sqrt{p\, \mathrm{Gbi}_{\mathcal{H}_p}(\mu, \langle \varphi(\mathbf{x}) \rangle)/B(1/p, 1 + 1/p)} \rfloor$ gives the r_p-index [5], where B is the Euler beta function.

Example 5. For $p \ge 1$ let $\mathcal{H}_p = \{\{\mathrm{Bs}_p(a, b)|_{\mathbb{I}^2}\}_{a > 0, b > 0}\}$, where $\mathrm{Bs}_p(a, b)$ is a scaled L_p ball (L_p ellipse): $\mathrm{Bs}_p(a, b) = \{(x, y) : \|(x/a, y/b)\|_p \le 1\}$, cf. Fig. 1. Then $p\, \mathrm{Gbi}_{\mathcal{H}_p}(\mu, \langle \varphi(\mathbf{x}) \rangle)/B(1/p, 1 + 1/p)$ is the (projected) l_p-index [5] if μ is the Lebesgue measure. Moreover, $\mathrm{Gbi}_{\mathcal{H}_\infty}(\mu, \mathsf{f})$ is equivalent to the Shilkret integral $\mathrm{Sh}(\mu, \mathsf{f})$ [12].

3 Determining the value of an l_p-index

As the definition of graph-based integrals for particular $\mathcal{H}$ sets may seem quite complicated, one should ask him- or herself a question whether there exist an algorithm that calculates the value of the integral efficiently. Of course, the Choquet, Sugeno, and Shilkret (and thus l_∞-index) integral for given $\mathsf{h}^{(\mu, \langle \varphi(\mathbf{x}) \rangle)}$ may be calculated in linear time, i.e. $O(n)$, where $n = |\{x_i : x_i > 0\}|$.

Calculation of some graph-based integrals (like decomposition integrals from Example 3 for some k) seem to be an NP-Complete problem. Here we will derive an algorithm for calculating $\mathrm{Gbi}_{\mathcal{H}_p}(\mu, \langle \varphi(\mathbf{x}) \rangle)$ for $1 \le p < \infty$, where $\mathcal{H}_p$ is given in Example 5, i.e. we will get an l_p-index in particular. Such method is of practical interest, as a naïve implementation has computational complexity of $O(n^3)$, which even for moderate values of n (> 100) may require too much of computer processor time.

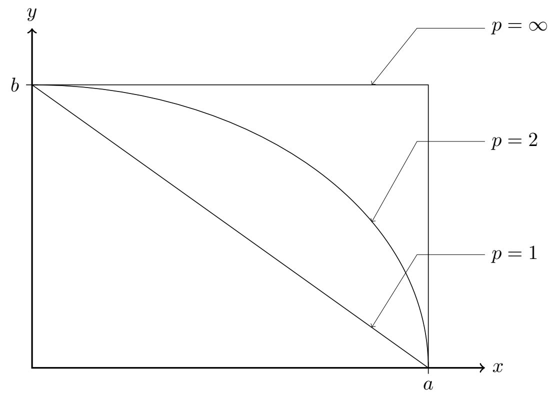

Figure 1: Boundaries of $\mathrm{Bs}_p(a,b)|_{\mathbb{I}^2}$ for different p.

Fix $\varphi, \mu, \mathbf{x}$ and $1 \le p < \infty$, with $\mathsf{h}^{(\mu,\langle\varphi(\mathbf{x})\rangle)} \not\equiv 0$. As $\mathsf{h}^{(\mu,\langle\varphi(\mathbf{x})\rangle)}$ is a lower semicontinuous step function, let $\mathbf{Q} = (\mathbf{q}_0, \ldots, \mathbf{q}_n)$, $\mathbf{q}_i = (q_{ix}, q_{iy}) \in \mathbb{I}^2$, $q_{0x} = 0$, q_{ix} be (all) such that $\mathsf{h}^{(\mu,\langle\varphi(\mathbf{x})\rangle)}(q_{ix}^-) \ne \mathsf{h}^{(\mu,\langle\varphi(\mathbf{x})\rangle)}(q_{ix}^+)$ for $i > 1$, and $q_{iy} = \mathsf{h}^{(\mu,\langle\varphi(\mathbf{x})\rangle)}(q_{ix})$ for $i = 0, 1, \ldots, n$. Moreover, we assume that $q_{ix} < q_{jx}$ for $i < j$.

Let $\mathbf{u} = (x_u, y_u)$ and $\mathbf{v} = (x_v, y_v)$ be arbitrary points in $\mathbb{I}^2$, for which $0 \le x_u < x_v$ and $y_u > y_v \ge 0$. Let $\mathrm{Bs}_p(\mathbf{u}, \mathbf{v})$ denote the L_p ellipse interpolating these two points. It may be easily shown that $\mathrm{Bs}_p(\mathbf{u}, \mathbf{v}) = \mathrm{Bs}_p(a, b)$, where

$$a = \left(\frac{c}{y_v^p - y_u^p}\right)^{\frac{1}{p}}, \quad b = \left(\frac{-c}{x_v^p - x_u^p}\right)^{\frac{1}{p}},$$

and $c = x_u^p y_v^p - x_v^p y_u^p$.

The following lemma states that the graph-based integral of our interest may be determined by calculating the measure of an l_p-ellipse interpolating some two points from $\mathbf{Q}$.

Lemma 1. *There exist i, k, $i < k$, such that*

$$\lambda(\mathrm{Bs}_p(\mathbf{q}_i, \mathbf{q}_k)) = \mathrm{Gbi}_{\mathcal{H}_p}(\mu, \langle\varphi(\mathbf{x})\rangle).$$

The proof is straightforward and therefore omitted.

The next lemma states that the boundaries of any two p-ellipses intersect in $\mathbb{I}^2$ at most in one point.

Lemma 2. *For any a, a', b, b' $|\partial\mathrm{Bs}_p(a,b)|_{\mathbb{I}^2} \cap \partial\mathrm{Bs}_p(a',b')|_{\mathbb{I}^2}| \le 1$.*

The proof is left to the reader.

Lemma 3. *Let $0 \le i < j < k \le n$, such that $\mathbf{q}_i \notin \mathrm{Bs}_p(\mathbf{q}_j, \mathbf{q}_k)$. Then*

(i) $\mathbf{q}_k \notin \mathrm{Bs}_p(\mathbf{q}_i, \mathbf{q}_j)$;

(ii) $\mathrm{Bs}_p(\mathbf{q}_i, \mathbf{q}_j)|_{[0,q_{jx}]\times\mathbb{I}} \supseteq \mathrm{Bs}_p(\mathbf{q}_j, \mathbf{q}_k)|_{[0,q_{jx}]\times\mathbb{I}}$;

(iii) $\mathrm{Bs}_p(\mathbf{q}_i, \mathbf{q}_j)|_{[q_{jx},\infty)\times\mathbb{I}} \subseteq \mathrm{Bs}_p(\mathbf{q}_j, \mathbf{q}_k)|_{[q_{jx},\infty)\times\mathbb{I}}$.

See Fig. 2 for an illustration of the lemma. The proof is omitted.

The proposed algorithm is given in Fig. 3. It is a modification of Graham's [8] routine for determining the convex hull of a finite planar set of points, also known as the Graham Scan.

The algorithm uses a stack, $\mathbf{S}$, i.e. a data structure on which the following operations may be performed: Push (adds an element to the top), Pop (removes the current top element) and $\#\mathbf{S}$ (returns the number of stored elements). Its elements may be accessed by an indexing operator $[\cdot]$, eg. $\mathbf{S}[\#\mathbf{S}]$ gets the element from the top of the stack.

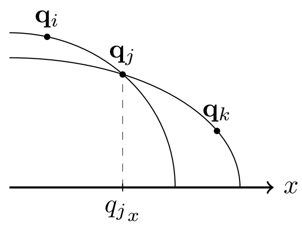

Figure 2: Illustration of Lemma 3.

Input: $p \in [1,\infty)$ and $\mathbf{Q} := (\mathbf{q}_0,\dots,\mathbf{q}_n)$ determined by $\mathbf{x},\mu,\varphi$ (see p. 3),
Result: $\mathrm{Gbi}_{\mathcal{H}_p}(\mu, \langle\varphi(\mathbf{x})\rangle)$.

1 Create an empty stack $\mathbf{S} \subseteq \mathbf{Q}$;
2 Push $\mathbf{q}_0$ into $\mathbf{S}$;
3 Let $i := 1$;
4 **while** $(i < n)$ **and** $(q_{i_y} = q_{0_y})$ **do**
5     $i := i + 1$;
6 Push $\mathbf{q}_i$ into $\mathbf{S}$;
7 **for** $j = i+1, i+2, \dots, n$ **do**
8   **if** $(\mathbf{S}[\#\mathbf{S}]_y \neq q_{j_y})$ **then** {
9   **while** $(\#\mathbf{S} \geq 2)$ **and** $(\mathbf{S}[\#\mathbf{S}-1] \in \operatorname{int}\mathrm{Bs}_p(\mathbf{S}[\#\mathbf{S}], \mathbf{q}_j))$ **do**
10   Pop from $\mathbf{S}$;
11   Push q_j into $\mathbf{S}$;
12   }
13 **return** $B(1/p, 1+1/p) \cdot \max\{\mathbf{S}[i]\cdot\mathbf{S}[i+1] : i = 1,2,\dots,\#\mathbf{S}-1\}/p$,
i.e. $\max\{\lambda(\mathrm{Bs}_p(\mathbf{S}[i],\mathbf{S}[i+1])) : i = 1,2,\dots,\#\mathbf{S}-1\}$;

Figure 3: Algorithm for computing $\mathrm{Gbi}_{\mathcal{H}_p}(\mu, \langle\varphi(\mathbf{x})\rangle)$, for $1 \leq p < \infty$.

The stack stores points in $\mathbb{I}^2$ which will be used to find the maximal p-ellipse. The algorithm scans through all the points from $\mathbf{Q}$, in the direction of increasing x (and nonincreasing y). At the j-th iteration, it repetitively pops elements from $\mathbf{S}$, until the p-ellipse interpolating $\mathbf{q}_j$ and the top-stack element does not contain the last-to-top element. After this process we consider only the p-ellipses obtained (by interpolation) from each two consecutive points on the stack. We state, that the p-ellipse of maximum area can be found among those.

Lemma 4. *Let $\mathbf{S}^*$ denote the contents of the stack after running the algorithm on arbitrary $\mathbf{Q} = Q(\mathsf{h}^{(\mu,\mathsf{f})})$. Then:*

(a) for all $i = 1,2,\dots,\#\mathbf{S}^*-1$ it holds that $\mathrm{Bs}_p(\mathbf{S}^*[i],\mathbf{S}^*[i+1]) \subseteq \mathrm{Gr}(\mathsf{h}^{(\mu,\mathsf{f})})$.

(b) for any $0 \leq i < k \leq n$ such that $\mathrm{Bs}_p(\mathbf{q}_i,\mathbf{q}_k) \subseteq \mathrm{Gr}(\mathsf{h}^{(\mu,\mathsf{f})})$ we have

$$\max\{\lambda(\mathrm{Bs}_p(\mathbf{S}[i],\mathbf{S}[i+1])) : i = 1,2,\dots,\#\mathbf{S}-1\} \geq \lambda(\mathrm{Bs}_p(\mathbf{q}_i,\mathbf{q}_k)).$$

Proof. (a) For brevity of notation, let $E_l := \mathrm{Bs}_p(\mathbf{S}^*[l],\mathbf{S}^*[l+1])$ for $l = 1,2,\dots,\#\mathbf{S}^*-1$. Let $\mathbf{q}_i,\mathbf{q}_j,\mathbf{q}_k \in \mathbf{S}^*$ be any three consecutive points from the stack. Also, let l be an integer such that $\mathbf{S}^*[l] = \mathbf{q}_j$. We have $E_l \equiv \mathrm{Bs}_p(\mathbf{q}_j,\mathbf{q}_k)$. Then for every m such that $j < m < k$ it holds $\mathbf{q}_m \notin \operatorname{int} E_l$, due to Lemma 2 and the fact that $\mathbf{q}_j \in \mathrm{Bs}_p(\mathbf{q}_m,\mathbf{q}_k)$ (otherwise $\mathbf{q}_m$ would not be removed from the stack at some step, cf. line 10 of the algorithm).

By Lemma 3, for any $i < m < j$ we have $\mathbf{q}_m \notin \operatorname{int} E_l$. As $\mathbf{S}^*[l-2] \notin \operatorname{int} E_{l-1}$ then also $\mathbf{S}^*[l-2] \notin E_l$. By induction, for every $m < j$, $\mathbf{q}_m \notin \operatorname{int} E_l$.

Similarly we may show that for every $m > k$, $\mathbf{q}_m \notin \operatorname{int} E_l$. As $\mathbf{S}^*[1] = \mathbf{q}_0$ and $\mathbf{S}^*[\#\mathbf{S}^*] = \mathbf{q}_n$,

(b) Let $i < k$ with $\mathbf{q}_i,\mathbf{q}_k$ not being two consecutive elements from $\mathbf{S}^*$. Consider two cases:

(i) Assume that $\mathbf{q}_i \in \mathbf{S}^*$ and $\mathbf{q}_k \in \mathbf{S}^*$. Let $\mathbf{q}_j$ be the element from the stack that directly precedes $\mathbf{q}_k$. By (a), $\mathbf{q}_j \neq \mathbf{q}_i$. We have $\mathbf{q}_i \notin \operatorname{int} \mathrm{Bs}_p(\mathbf{q}_j, \mathbf{q}_k)$. Lemma 2 states that p-ellipses $\mathrm{Bs}_p(\mathbf{q}_i, \mathbf{q}_j)$ and $\mathrm{Bs}_p(\mathbf{q}_j, \mathbf{q}_k)$ intersect only in $\mathbf{q}_j$. That implies $\mathbf{q}_j \in \operatorname{int} \mathrm{Bs}_p(\mathbf{q}_i, \mathbf{q}_k)$, thus $\mathrm{Bs}_p(\mathbf{q}_i, \mathbf{q}_k)$ cannot generate a proper solution to our task.

(ii) Assume that $\mathbf{q}_i \notin \mathbf{S}^*$. Let $j = \min\{m > i : \mathbf{q}_m \in \mathbf{S}^*\}$, i.e. $\mathbf{q}_j$ be the element from the stack the nearest to $\mathbf{q}_i$ on the right. Also, let l be an integer such that $\mathbf{S}^*[l] = \mathbf{q}_j$. By Lemma 2, we either have $\mathbf{q}_j \in \operatorname{int} \mathrm{Bs}_p(\mathbf{q}_i, \mathbf{q}_k)$ or $\mathbf{q}_j \notin \operatorname{int} \mathrm{Bs}_p(\mathbf{q}_i, \mathbf{q}_k)$ but $\mathbf{S}^*[l-1] \in \operatorname{int} \mathrm{Bs}_p(\mathbf{q}_i, \mathbf{q}_k)$, because $\mathbf{S}^*[l-1] \in \operatorname{int} \mathrm{Bs}_p(\mathbf{q}_i, \mathbf{q}_j)$ (line 10 of the algorithm). Thus,

$$\lambda(\mathrm{Bs}_p(\mathbf{q}_i, \mathbf{q}_k)) < \max\left\{\lambda(\mathrm{Bs}_p(\mathbf{S}[i], \mathbf{S}[i+1])) : i = 1, 2, \ldots, \#\mathbf{S} - 1\right\}.$$

(iii) Case $\mathbf{q}_k \notin \mathbf{S}^*$ is similar to the previous one, therefore the proof is complete.

□

Now we may approach to the final conclusion.

Theorem 5. *For any $1 \le p < \infty$, nonincreasingly sorted $\mathbf{x}$, μ, and φ, $\mathrm{Gr}(\mathsf{h}^{(\mu,\mathsf{f})})$ may be determined in linear time using the algorithm in Fig. 3.*

Proof. By Lemma 1, the maximal p-ellipse interpolates some two points from $\mathbf{Q}$. The algorithm finds in $O(n)$ time the only (Lemma 4) p-ellipses which may generate the desired solution. In the last step of the algorithm, the largest p-ellipse is determined in $O(n)$, thus the proof is complete. □

4 Conclusions

In this paper we introduced the notion of a graph-based integral, which generalizes the Choquet, Shilkret, and Sugeno integrals, as well as some decomposition integrals, and which have a very appealing graphical interpretation. Moreover, for a particular class of those integrals, the l_p-indices, we developed an efficient, linear-time algorithm. The routine being, on its own, an interesting modification of Graham's Scan, was shown to be potentially useful in practical applications. Its implementation for $\mu = \lambda$ and $\varphi = \mathrm{id}$ has been included in the agop package for R, see [4].

Future work should definitely explore formally the properties of graph-based integrals and their relation with a universal integral.

Acknowledgments

The contribution of Marek Gagolewski was partially supported by the FNP START 2013 Scholarship from the Foundation for Polish Science.

References

[1] G. Beliakov and S. James. Citation-based journal ranks: The use of fuzzy measures. *Fuzzy Sets and Systems*, 167:101–119, 2011.

[2] M. Bras-Amorós, J. Domingo-Ferrer, and V. Torra. A bibliometric index based on the collaboration distance between cited and citing authors. *Journal of Informetrics*, 5(2):248–264, 2011.

[3] G. Choquet. Theory of capacities. *Annales de l'institut Fourier*, 5:131–295, 1954.

[4] M. Gagolewski and A. Cena. *agop: Aggregation Operators and Preordered Sets in R*, 2013. http://agop.rexamine.com.

[5] M. Gagolewski and P. Grzegorzewski. A geometric approach to the construction of scientific impact indices. *Scientometrics*, 81(3):617–634, 2009.

[6] M. Gagolewski and P. Grzegorzewski. Possibilistic analysis of arity-monotonic aggregation operators and its relation to bibliometric impact assessment of individuals. *International Journal of Approximate Reasoning*, 52(9):1312–1324, 2011.

[7] M. Gagolewski and R. Mesiar. Monotone measures and universal integrals in a uniform framework for the scientific impact assessment problem, 2013. submitted.

[8] R.L. Graham. An efficient algorithm for determining the convex hull of a finite planar set. *Information Processing Letters*, 1:132–133, 1972.

[9] J.E. Hirsch. An index to quantify individual's scientific research output. *Proceedings of the National Academy of Sciences*, 102(46):16569–16572, 2005.

[10] E.P. Klement, R. Mesiar, and E. Pap. A universal integral as common frame for Choquet and Sugeno integral. *IEEE Transactions on Fuzzy Systems*, 18:178–187, 2010.

[11] R. Mesiar and A. Stupnňanová. Decomposition integrals. *International Journal of Approximate Reasoning*, 54(8):1252–1259, 2013.

[12] N. Shilkret. Maxitive measure and integration. *Indag. Math.*, 33:109–116, 1971.

[13] M. Sugeno. *Theory of fuzzy integrals and its applications*. PhD thesis, Tokyo Institute of Technology, 1974.

[14] V. Torra and Y. Narukawa. The h-index and the number of citations: Two fuzzy integrals. *IEEE Transactions on Fuzzy Systems*, 16(3):795–797, 2008.

[15] G.J. Woeginger. An axiomatic characterization of the Hirsch-index. *Mathematical Social Sciences*, 56(2):224–232, 2008.