Positivity (2023) 27:3
https://doi.org/10.1007/s11117-022-00957-3

Positivity

On the compounding of higher order monotonic pseudo-Boolean functions

Paul Ressel[1]

Received: 24 June 2021 / Accepted: 29 December 2021

Abstract
Compounding submodular monotone (i.e. 2-alternating) set functions on a finite set preserves this property, as shown in 2010. A natural generalization to k-alternating functions was presented in 2018, however hardly readable because of page long formulas. We give an easier proof of a more general result, exploiting known properties of higher order monotonic functions.

Keywords Submodular · Pseudo-Boolean function · Higher order monotonic · k-alternating · Multilinear polynomial · Set interval

Mathematics Subject Classification 06E30 · 26A48 · 26D07 · 26C99

1 Introduction

Let V be a finite non-empty set. A function $\varphi : \mathcal{P}(V) \longrightarrow \mathbb{R}$ on $\mathcal{P}(V)$, the set of all subsets of V, i.e. a socalled *pseudo-Boolean function*, is *submodular* if

$$\varphi(A \cup \{v\}) - \varphi(A) \geq \varphi(B \cup \{v\}) - \varphi(B) \tag{1}$$

for all $A \subseteq B$ and all $v \in V \backslash B$. And φ is *increasing* if $\varphi(A) \leq \varphi(B)$ whenever $A \subseteq B \subseteq V$. Condition (1) has in many applications the interpretation that the marginal effect expressed by φ decreases for larger subsets (the property of "diminishing returns"). It is not surprising that submodular increasing functions are modelling many situations, both technical and social, for example the influence in social networks. In this connection an interesting aggregation problem had been posed in [1]: does "local" submodularity imply the corresponding property "globally"? This was confirmed 10 years later in [2]. Now, in another "language", an increasing submod-

✉ Paul Ressel
paul.ressel@ku.de

1 Katholische Universität Eichstätt-Ingolstadt, Eichstätt, Germany

ular function φ on $\mathcal{P}(V)$ is “2-alternating” on $\{0,1\}^V \cong \mathcal{P}(V)$, and it seems natural to consider the more general case of “k-alternating” functions. For example, φ is 3-alternating if, in addition to being increasing and submodular, the difference between the left and right hand side in (1) is further diminished if one more element is added. This idea is suggested in the recent work [3], whose central mathematical result (Theorem 4) however is given a very complicated and hardly readable proof, with page-long formulas. We shall give a much more transparent proof, based on existing theorems about higher order monotonic functions. Our result is also considerably more general.

Notations

$\mathbb{N} = \{1,2,3,\ldots\}$, $\mathbb{N}_0 = \{0,1,2,\ldots,\}$, $\mathbb{R}_+ = [0,\infty[$, $\mathcal{P}(V) =$ set of all subsets of V,

$1_A(x) := \begin{cases} 1, x \in A \\ 0, x \notin A \end{cases}$, $[d] := \{1,2,\ldots,d\}$ for $d \in \mathbb{N}$, $\mathbf{1}_d := (1,1,\ldots,1) \in \mathbb{N}^d$,

$|\boldsymbol{n}| := \sum_{i=1}^d n_i$ for $\boldsymbol{n} \in \mathbb{N}_0^d$, $|\alpha| :=$ cardinality of a finite set α.

These two uses of the same symbol can hardly be mixed up; in fact, for $a = 1_\alpha \in \{0,1\}^d$ we have $|a| = |\alpha|$.

$(f \times g)(x,y) := (f(x), g(y))$ for mappings f, g,

$(f,g)(x) := (f(x), g(x))$ for mappings f, g with the same domain

$\langle \sigma, \tau \rangle := \{\gamma | \sigma \subseteq \gamma \subseteq \tau\}$ a *set-interval* where (usually) $\sigma \subseteq \tau$

$\langle \sigma, \sigma \rangle = \{\sigma\}$ is a special case

$A \dot\cup B$, $\dot\bigcup_j A_j$ for disjoint unions, d.f. = distribution function (of some measure),

$x \odot y := (x_1 y_1, x_2 y_2, \ldots)$ and $x \vee y := (x_1 \vee y_1, x_2 \vee y_2, \ldots)$ for two vectors x, y of equal dimension.

2 Multivariate higher order monotonicity

Let $I_1, \ldots, I_d \subseteq \mathbb{R}$ be non-degenerate intervals, $I := I_1 \times \cdots \times I_d$, and let $f : I \longrightarrow \mathbb{R}$ be any function. For $s \in I$, $h \in \mathbb{R}_+^d$ such that also $s + h \in I$ put

$$(E_h f)(s) := f(s+h)$$

and $\Delta_h := E_h - E_0$, i.e. $(\Delta_h f)(s) := f(s+h) - f(s)$, and $\nabla_h := -\Delta_h$.

Since $\{E_h\}$ is commutative (where defined), so are $\{\Delta_h\}$ and $\{\nabla_h\}$. In particular, with $e_1, \ldots, e_d$ denoting the standard unit vectors in $\mathbb{R}^d$, $\Delta_{h_1 e_1}, \ldots, \Delta_{h_d e_d}$ commute. As usual $\Delta_h^0 f := f =: \nabla_h^0 f$. For $\boldsymbol{n} = (n_1, \ldots, n_d) \in \mathbb{N}_0^d$ and $h \in \mathbb{R}_+^d$ we put

$$\Delta_h^{\boldsymbol{n}} := \Delta_{h_1 e_1}^{n_1} \Delta_{h_2 e_2}^{n_2} \ldots \Delta_{h_d e_d}^{n_d}$$

and similarly $\nabla_h^{\boldsymbol{n}}$. The multinomial theorem gives

$$\left(\Delta_h^{\boldsymbol{n}} f\right)(s) = \sum_{0 \le \boldsymbol{q} \le \boldsymbol{n}} (-1)^{|\boldsymbol{n}| - |\boldsymbol{q}|} \binom{\boldsymbol{n}}{\boldsymbol{q}} f(s + \boldsymbol{q} \odot h). \tag{2}$$

Note that $\Delta_h^{\mathbf{1}_d} \neq \Delta_h$ for $d > 1$ and $h \neq 0$. Also, $\Delta_h^{\boldsymbol{n}} = 0$ if $h_i = 0 < n_i$ for some $i \leq d$.

Definition $f : I \longrightarrow \mathbb{R}$ is called

(i) $\boldsymbol{n}$-*increasing* ("$\boldsymbol{n}$ - $\uparrow$") if

$$(\Delta_h^{\boldsymbol{p}} f)(s) \geq 0$$

for all $s \in I$, $h \in \mathbb{R}_+^d$, $\boldsymbol{p} \in \mathbb{N}_0^d$, $0 \neq \boldsymbol{p} \leq \boldsymbol{n}$ such that $s + \boldsymbol{p} \odot h \in I$

(ii) $\boldsymbol{n}$-*decreasing* ("$\boldsymbol{n}$ - $\downarrow$") if instead

$$\left(\nabla_h^{\boldsymbol{p}} f\right)(s) \geq 0$$

(iii) $\boldsymbol{n}$-*alternating* ("$\boldsymbol{n}$ - $\updownarrow$") if instead

$$\left(\nabla_h^{\boldsymbol{p}} f\right)(s) \leq 0.$$

It is easy to see that, using the notation $(f(-\cdot))(s) := f(-s)$,

$$\begin{aligned} f \text{ is } \boldsymbol{n}\text{-}\downarrow \text{ on } I &\iff f(-\cdot) \text{ is } \boldsymbol{n}\text{-}\uparrow \text{ on } -I \\ f \text{ is } \boldsymbol{n}\text{-}\updownarrow \text{ on } I &\iff -f(-\cdot) \text{ is } \boldsymbol{n}\text{-}\uparrow \text{ on } -I. \end{aligned}$$

For $\boldsymbol{n} \in \{0,1\}^d$ the I_j considered here need not be intervals, just non-empty subsets of $\mathbb{R}$, or even $\overline{\mathbb{R}}$. Right-continuous bounded non-negative $\mathbf{1}_d$ - $\uparrow$ functions on I are precisely the distribution functions ("d.f.s") of finite measures on $\overline{I}$ (closure in $\overline{\mathbb{R}}^d$), see [4] Theorem 7, a result which will be used later on. Functions which are $\mathbf{1}_d$ - $\uparrow$ ($\downarrow$, $\updownarrow$) are also called *fully d-increasing* (*-decreasing*, *-alternating*), and this notion will now be extended:

Definition Let $I_1, \ldots, I_d \subseteq \overline{\mathbb{R}}$ be any non-empty subsets, $I := I_1 \times \cdots \times I_d$, $1 \leq k \leq d$. Then $f : I \longrightarrow \mathbb{R}$ is *fully k-increasing* ("$\mathbf{1}_k$ - $\uparrow$") iff

$$(\Delta_h^{\boldsymbol{p}} f)(s) \geq 0 \quad \text{for each} \quad 0 \lneqq \mathbf{p} \leq \mathbf{1}_d \quad \text{with} \quad |\mathbf{p}| \leq k$$

and for each $s \in I$ and $h \in \mathbb{R}_+^d$ such that $s + \mathbf{p} \odot h \in I$.

If instead $(\nabla_h^{\boldsymbol{p}} f)(s) \geq 0$ we call f *fully k-decreasing* ("$\mathbf{1}_k$ - $\downarrow$") and if $(\nabla_h^{\boldsymbol{p}} f)(s) \leq 0$, f is by definition *fully k-alternating* ("$\mathbf{1}_k$ - $\updownarrow$").

For the important special case where $I_j = \{0,1\}$ $\forall j \leq d$, i.e. for ("pseudo-Boolean") functions on $\{0,1\}^d$, it is sometimes useful to identify $\{0,1\}^d$ with $\mathcal{P}([d]) := \{\alpha | \alpha \subseteq [d]\}$. Since $\Delta_0^1 = 0$ and $\Delta_1^0 = \text{id} = \Delta_0^0$, only $\Delta_h^{\boldsymbol{p}}$ with $h = \boldsymbol{p} \in \{0,1\}^d \backslash \{0\}$ have to be considered. It is then reasonable to use the simplified notation

$$\Delta_\alpha := \Delta_a^a \text{ for } a = 1_\alpha \in \{0,1\}^d \backslash \{0\}$$

(complemented by $\Delta_\emptyset = \mathrm{id}$).

We write likewise $E_\alpha := E_a$. Both $\Delta_\alpha f$ and $E_\alpha f$ have the domain $\{\gamma \subseteq [d] | \gamma \subseteq \alpha^c\}$, and for $\gamma \subseteq \alpha^c$

$$(\Delta_\alpha f)(\gamma) = ((E_\gamma \Delta_\alpha)(f))(\emptyset) = ((\Delta_\alpha E_\gamma)(f))(\emptyset).$$

Clearly $\Delta_\alpha \circ \Delta_\beta = \Delta_{\alpha\cup\beta}$ for disjoint α, β. Note that

$$(\Delta_\alpha f)(\emptyset) = f(\alpha) - \sum_{\substack{\gamma\subseteq\alpha \\ |\gamma|=|\alpha|-1}} f(\gamma) + \sum_{\substack{\gamma\subseteq\alpha \\ |\gamma|=|\alpha|-2}} f(\gamma) \mp \cdots + (-1)^{|\alpha|} f(\emptyset).$$

The following identity (for $x_1, \ldots, x_d \in \mathbb{R}$)

$$\prod_{i=1}^d x_i = \prod_{i=1}^d [(x_i - 1) + 1] = \sum_{\alpha\subseteq[d]} \prod_{i\in\alpha} (x_i - 1) \qquad \left(\prod_\emptyset := 1 \right)$$

holds of course also within the commutative algebra generated by $\{E_{\{i\}} | i \in [d]\}$, and leads to

$$\sum_{\alpha\subseteq[d]} \Delta_\alpha = \prod_{i=1}^d E_{\{i\}} = E_{[d]},$$

i.e. to $\sum_{\alpha\subseteq[d]} (\Delta_\alpha f)(\emptyset) = f([d])$.

Slightly more general, and of importance later on, for $\beta \subseteq \gamma \subseteq [d]$

$$\sum_{\alpha\in\langle\beta,\gamma\rangle} \Delta_\alpha = \Delta_\beta \sum_{\alpha\subseteq\gamma\backslash\beta} \Delta_\alpha = \Delta_\beta E_{\gamma\backslash\beta}. \tag{3}$$

We mention that fully k-alternating pseudo-Boolean functions are called "$AD-k$" in [3].

3 Multilinear polynomials

Any (pseudo-Boolean) function $f : \{0,1\}^d \longrightarrow \mathbb{R}$ has an extension $\tilde{f}$ to a socalled *multilinear polynomial*

$$\tilde{f}(x) := \sum_{\alpha\subseteq[d]} f(\alpha) x^\alpha (\mathbf{1} - x)^{\alpha^c}, \quad x \in \mathbb{R}^d \tag{4}$$

where we use the abbreviations $x^\alpha := \prod_{i\in\alpha} x_i$, $x^\emptyset := 1$ and $\mathbf{1} := \mathbf{1}_d$. "Multilinear" means here that no variable appears in a power > 1 in $\tilde{f}$; $\tilde{f}$ is therefore an affine function of each variable x_i. Note that $\tilde{f}(1_\alpha) = f(\alpha)\,\forall\, \alpha \subseteq [d]$, so $\tilde{f}$ is uniquely determined; or in other words, each multilinear polynomial is the extension of its restriction to $\{0,1\}^d$, where we freely identify $\alpha \subseteq [d]$ with $1_\alpha \in \{0,1\}^d$. It is immediate that $f \geq 0$ iff $\tilde{f} \mid [0,1]^d \geq 0$.

Let for $\emptyset \neq \beta \subseteq [d]$ the partial derivative of $\tilde{f}$ w.r. to x_i, $i \in \beta$ be $\partial^\beta \tilde{f}$ (every other partial derivative of $\tilde{f}$ is obviously 0). Then for any $p \in [d]$

$$(\partial^{\{p\}} \tilde{f})(x) = \sum_{\alpha \subseteq [d]\setminus\{p\}} (\Delta_{\{p\}} f)(\alpha) x^\alpha (\mathbf{1} - x)^{[d]\setminus(\alpha\cup\{p\})}$$

by an application of the product role, i.e. $\partial^{\{p\}} \tilde{f}$ is multilinear in x_i for $i \in [d]\setminus\{p\}$. By iteration we obtain for any $\emptyset \neq \beta \subseteq [d]$

$$(\partial^\beta \tilde{f})(x) = \sum_{\alpha \subseteq [d]\setminus\beta} (\Delta_\beta f)(\alpha) x^\alpha (\mathbf{1} - x)^{[d]\setminus(\alpha\cup\beta)} \tag{5}$$

including finally

$$(\partial^{[d]} \tilde{f})(x) = (\Delta_{[d]} f)(\emptyset), \quad \text{a constant.}$$

That is, $\partial^\beta \tilde{f}$ is the multilinear extension of $\Delta_\beta f$ on $\{0,1\}^{\beta^c}$.

Now (5) implies

$$(\partial^\beta \tilde{f})(0) = (\Delta_\beta f)(\emptyset), \quad \beta \subseteq [d], \tag{6}$$

and in the likewise "canonical" representation

$$\tilde{f}(x) = \sum_{\alpha \subseteq [d]} c_\alpha x^\alpha$$

we have obviously $c_\alpha = (\partial^\alpha \tilde{f})(0)$. Combining this with (6) we get

$$\tilde{f}(x) = \sum_\alpha (\Delta_\alpha f)(\emptyset) x^\alpha. \tag{7}$$

We'll need later on the following result:

Lemma 1 *For $f : \{0,1\}^d \longrightarrow \mathbb{R}$ and its multilinear extension $\tilde{f}$ we have*

$$f \ is\ \mathbf{1}_k\text{-}\uparrow \iff \tilde{f}\ is\ \mathbf{1}_k\text{-}\uparrow\ on\ [0,1]^d.$$

Proof $\tilde{f}$ is a polynomial, in particular C^∞. Therefore $\tilde{f}$ is $\mathbf{1}_k$ - $\uparrow$ (on $[0,1]^d$) if and only if

$$(\partial^\beta \tilde{f})(x) \geq 0 \quad \forall\, |\beta| \leq k,\ \forall\, x \in [0,1]^d$$

which, as we just saw, is equivalent with

$$(\Delta_\beta f)(\alpha) \geq 0 \quad \forall\, |\beta| \leq k,\ \forall\, \alpha \subseteq \beta^c,$$

the defining property of f being $\mathbf{1}_k$ - $\uparrow$. □

Example 1 For $d = 3, k = 2$ consider $f : \mathcal{P}([3]) \longrightarrow \mathbb{R}$ given by $f(\alpha) := |\alpha| \vee 1$. Then

$$\tilde{f}(x) = 1 + x_1x_2 + x_1x_3 + x_2x_3 - x_1x_2x_3$$

and

$$(\partial^{\{1\}} \tilde{f})(x) = x_2 + x_3 - x_2x_3 \ \text{ etc.}$$
$$(\partial^{\{1,2\}} \tilde{f})(x) = 1 - x_3 \ \text{ etc.}$$

are all non-negative on $[0, 1]^3$; however

$$\partial^{\{1,2,3\}} \tilde{f} = -1,$$

showing f to be $\mathbf{1}_2$ - $\uparrow$, but not $\mathbf{1}_3$ - $\uparrow$. Slightly more general, $f(\alpha) := |\alpha| \vee 1$ is for any $d \geq 3$ $\mathbf{1}_2$ - $\uparrow$ and not $\mathbf{1}_3$ - $\uparrow$: we have

$$\tilde{f}(x) = \sum_{i=1}^{d} x_i + \prod_{i=1}^{d} (1 - x_i),$$

whence

$$(\partial^{\{i\}} \tilde{f})(x) = 1 - \prod_{\ell \neq i} (1 - x_\ell),$$
$$(\partial^{\{i,j\}} \tilde{f})(x) = \prod_{\ell \neq i,j} (1 - x_\ell) \quad \text{for } i \neq j,$$

and

$$\partial^\alpha \tilde{f}(0) = -1 \quad \text{for } |\alpha| = 3.$$

4 A combinatorial intermezzo

The following Lemma (of combinatorial nature) will play a crucial role in the proof of the main result. We shall use *set-intervals* in $\mathcal{P}([d])$ of the form

$$\langle \sigma, \tau \rangle := \{\gamma \mid \sigma \subseteq \gamma \subseteq \tau\},$$

including as a special case singletons

$$\langle\sigma\rangle := \langle\sigma,\sigma\rangle = \{\sigma\}.$$

Note that $\langle\sigma,\tau\rangle \neq \emptyset$ iff $\sigma \subseteq \tau$, and that

$$\langle\sigma_1,\tau_1\rangle \cap \langle\sigma_2,\tau_2\rangle = \langle\sigma_1 \cup \sigma_2, \tau_1 \cap \tau_2\rangle.$$

Lemma 2 *Let $k, d \in \mathbb{N}, k \leq d$, and $x_1,\dots,x_d \in \mathbb{R}^k$. For non-empty $\alpha,\beta \subseteq [d]$ define*

$$\alpha \sim \beta :\Longleftrightarrow \max_{i\in\alpha} x_i = \max_{i\in\beta} x_i \quad (\in \mathbb{R}^k).$$

Then $\{\gamma \subseteq [d] \mid |\gamma| \geq k\}$ is the disjoint union of set-intervals $\langle\sigma_j,\tau_j\rangle$ with $|\sigma_j| = k, \sigma_j \subseteq \tau_j$ and $\sigma_j \sim \tau_j$ for each j.

Proof For $k = 1$ we may assume $x_1 \leq x_2 \leq \cdots \leq x_d$, and then

$$\{\gamma \subseteq [d] \mid |\gamma| \geq 1\} = \langle\{d\},[d]\rangle \,\dot\cup\, \langle\{d-1\},[d-1]\rangle \,\dot\cup\, \cdots \,\dot\cup\, \langle\{2\},[2]\rangle \,\dot\cup\, \langle\{1\}\rangle$$

has the required properties.

For $k \geq 2$ and $d = k+1$ choose $\alpha \subseteq [d]$ of size k such that $\alpha \sim [d]$ (which is evidently possible for any $d > k$). Then

$$\{\gamma \subseteq [d] \mid |\gamma| \geq k\} = \langle\alpha,[d]\rangle \,\dot\cup\, \dot\bigcup_{a\in\alpha} \langle[d]\backslash\{a\}\rangle.$$

We now proceed by induction and suppose the result to be true for some $d \geq 3$ and each $k \leq d$. Let $x_1,\dots,x_{d+1} \in \mathbb{R}^k$ be given, $k \geq 2$. It is no restriction to assume $k < d$ and

$$x_{d+1}(k) = \max_{i\leq d+1} x_i(k).$$

By assumption,

$$\{\gamma \subseteq [d] \mid |\gamma| \geq k\} = \dot\bigcup_j \langle\xi_j,\eta_j\rangle$$

is the disjoint union of set-intervals, with $\xi_j \subseteq \eta_j \subseteq [d], |\xi_j| = k$ and $\xi_j \sim \eta_j$ for each j. Let $y_i := (x_i(1),\dots,x_i(k-1)) \in \mathbb{R}^{k-1}$ be the projection of x_i, $i = 1,\dots,d$. Making use once more of the induction hypothesis we have

$$\{\gamma \subseteq [d] \mid |\gamma| \geq k-1\} = \dot\bigcup_p \langle\alpha_p,\beta_p\rangle$$

with $\alpha_p \subseteq \beta_p \subseteq [d], |\alpha_p| = k-1, \alpha_p \approx \beta_p$ for all p, where $\alpha \approx \beta$ means $\max_{i\in\alpha} y_i = \max_{i\in\beta} y_i (\in \mathbb{R}^{k-1})$.

We now put

$$\overline{\alpha}_p := \alpha_p \cup \{d+1\}, \quad \overline{\beta}_p := \beta_p \cup \{d+1\},$$

then $|\overline{\alpha}_p| = k$, $\overline{\alpha}_p \subseteq \overline{\beta}_p \subseteq [d+1]$, and $\overline{\alpha}_p \sim \overline{\beta}_p$, since for $\ell < k$

$$\begin{aligned} \max_{i \in \overline{\alpha}_p} x_i(\ell) &= \left(\max_{i \in \alpha_p} x_i(\ell) \right) \vee x_{d+1}(\ell) \\ &= \left(\max_{i \in \alpha_p} y_i(\ell) \right) \vee x_{d+1}(\ell) \\ &= \left(\max_{i \in \beta_p} y_i(\ell) \right) \vee x_{d+1}(\ell) \\ &= \max_{i \in \overline{\beta}_p} x_i(\ell) \end{aligned}$$

and

$$\max_{i \in \overline{\alpha}_p} x_i(k) = x_{d+1}(k) = \max_{i \in \overline{\beta}_p} x_i(k).$$

For any j and p we have

$$\langle \xi_j, \eta_j \rangle \cap \langle \overline{\alpha}_p, \overline{\beta}_p \rangle = \langle \xi_j \cup \overline{\alpha}_p, \eta_j \cap \overline{\beta}_p \rangle = \emptyset$$

since $d+1 \in \overline{\alpha}_p$, but $d+1 \notin \eta_j$.

For $p \neq q$ likewise

$$\langle \overline{\alpha}_p, \overline{\beta}_p \rangle \cap \langle \overline{\alpha}_q, \overline{\beta}_q \rangle = \emptyset$$

because otherwise $\alpha_p \cup \alpha_q \subseteq \beta_p \cap \beta_q$, contradicting the choice of α_p, β_p.

So, finally

$$\{\gamma \subseteq [d+1] \,|\, |\gamma| \geq k\} = \dot{\bigcup}_j \langle \xi_j, \eta_j \rangle \,\dot{\cup}\, \dot{\bigcup}_p \langle \overline{\alpha}_p, \overline{\beta}_p \rangle$$

is a partition into disjoint set-intervals as claimed. □

5 The main result

In [3], Theorem 4 the following is shown: let $f : \{0,1\}^d \longrightarrow [0,1]$ and $g_1, \ldots, g_d : \{0,1\}^k \longrightarrow [0,1]$ be all fully k-alternating ("$\mathbf{1}_k$ - $\updownarrow$"), then also their "compounding" $h : \{0,1\}^k \longrightarrow \mathbb{R}$, defined by

$$h(x) := \sum_{\alpha \subseteq [d]} f(\alpha) \prod_{i \in \alpha} g_i(x) \prod_{j \in \alpha^c} (1 - g_j(x))$$

has this property. The proof there is based on the multilinear extensions of f, $\{g_i\}$ and h, but it is hardly readable, with formulas longer than a page. Since the result is true (see below), I believe their proof is, too, although I didn't check it in detail—by lack of patience.

We will prove a more general result, allowing $g_1, \dots, g_d$ to be any $\mathbf{1}_k$ - $\updownarrow$ functions on an arbitrary product subset of $\overline{\mathbb{R}}^k$. Only f has to remain a pseudo-Boolean function. We shall first deal with $\mathbf{1}_k$ - $\uparrow$ functions (generalizing increasing supermodular functions), and then deduce from it the statement about $\mathbf{1}_k$ - $\updownarrow$ functions in a straightforward way.

We shall need the following approximation result.

Lemma 3 *Let $A = A_1 \times \cdots \times A_k$ be a product of non-empty subsets $A_j \subseteq \overline{\mathbb{R}}$, and let $g : A \longrightarrow [0, 1]$ be $\mathbf{1}_k$ - $\uparrow$ and such that $\sup g(A) = 1$. Then there is a net (g_α) of distribution functions of probability measures with finite support contained in A, which converges pointwise to g.*

Proof Let $\alpha_j \subseteq A_j$ be finite and non-empty, $1 \le j \le k$, and $\alpha := \alpha_1 \times \cdots \times \alpha_k$; we may assume the α_j so large that $g(\max \alpha) > 0$. The restriction $g \mid \alpha$ is $\mathbf{1}_k$ - $\uparrow$, (automatically right-continuous on α(!)), and so there exists by [4], Theorem 7 a finite measure ν_α on α with d.f. $g \mid \alpha$. We have $\nu_\alpha(\alpha) = g(\max \alpha) > 0$, hence $\mu_\alpha := \nu_\alpha / g(\max \alpha)$ is a probability measure on α, which is extended trivially to a probability measure on $\overline{A}$, with $\mu_\alpha(\overline{A} \backslash \alpha) := 0$. By $g_\alpha : A \longrightarrow [0, 1]$ we denote the d.f. of this extended μ_α.

In order to see that g_α converges pointwise to g, let $0 < \varepsilon < 1/2$ and some (finite, non-empty) product set $\alpha_0 \subseteq A$ be given. Choose $\alpha \supseteq \alpha_0$ (a product set, too) so large, such that $g(\max \alpha) \ge 1 - \varepsilon$. Then for any $a \in \alpha$

$$
\begin{aligned}
|g_\alpha(a) - g(a)| = \left| \frac{g(a)}{g(\max \alpha)} - g(a) \right| &= g(a) \cdot \frac{1 - g(\max \alpha)}{g(\max \alpha)} \\
&\le g(a) \frac{\varepsilon}{1 - \varepsilon} \le 2\varepsilon.
\end{aligned}
$$

Noting that the family of finite product sets in A is upwards filtering, the proof is complete. □

Theorem 1 *Let $k, d \in \mathbb{N}, k \le d, \emptyset \ne A_j \subseteq \overline{\mathbb{R}}$ for $j = 1, \dots, k$, $A := A_1 \times \cdots \times A_k$. Let $g_i : A \longrightarrow [0, 1]$ for $i = 1, \dots, d$ and $f : \{0, 1\}^d \longrightarrow \mathbb{R}$ be given. Define $h : A \longrightarrow \mathbb{R}$ by*

$$
h(x) := \sum_{\alpha \subseteq [d]} f(\alpha) \prod_{i \in \alpha} g_i(x) \prod_{j \in \alpha^c} [1 - g_j(x)].
$$

Then, if $g_1, \dots g_d$ and f are all $\mathbf{1}_k$ - $\uparrow$, so is h.

Proof With $\tilde{f}$ as the multilinear extension of f, and $g := (g_1, \dots, g_d) : A \longrightarrow [0, 1]^d$, we have $h = \tilde{f} \circ g$, and by Lemma 1 $\tilde{f}$ is also $\mathbf{1}_k$ - $\uparrow$ on $[0, 1]^d$.

We first consider the case that g_i is the d.f. of some one-point measure ε_{a_i}, where $a_i \in A$. Then

$$g_i = 1_{[a_i,\infty]\cap A}$$

and for $\emptyset \neq \alpha \subseteq [d]$

$$\prod_{i\in\alpha} g_i = 1_{[\max\limits_{i\in\alpha} a_i,\infty]\cap A},$$

and then by (7)

$$\begin{aligned}\tilde{f}\circ g &= \sum_{\alpha\subseteq[d]} (\Delta_\alpha f)(\emptyset)\cdot\prod_{i\in\alpha} g_i\\ &= \sum_{\alpha\subseteq[d]} (\Delta_\alpha f)(\emptyset)\cdot 1_{[\max_{i\in\alpha} a_i,\infty]\cap A}.\end{aligned}$$

For $k = d$ we have $(\Delta_\alpha f)(\emptyset) \geq 0$ for each $\alpha \subseteq [d]$, implying directly that $\tilde{f}\circ g$ is $\mathbf{1}_k$ - $\uparrow$, too. For $k < d$ we apply Lemma 2, i.e.

$$\{\gamma \subseteq [d] \mid |\gamma| \geq k\} = \dot{\bigcup}_j \langle\sigma_j, \tau_j\rangle$$

is a disjoint union of set intervals, where $\sigma_j \subseteq \tau_j$, $\sigma_j \sim \tau_j$ and $|\sigma_j| = k$ for each j. Remember that $\sigma \sim \tau$ means $\max\limits_{\sigma} x_i = \max\limits_{\tau} x_i$ (in $\mathbb{R}^k$). Since by (3)

$$\sum_{\alpha\in\langle\sigma_j,\tau_j\rangle} (\Delta_\alpha f)(\emptyset) = (\Delta_{\sigma_j}\circ E_{\tau_j\backslash\sigma_j})(f)(\emptyset) = (\Delta_{\sigma_j} f)(\tau_j\backslash\sigma_j) \geq 0$$

(because of $|\sigma_j| = k$), we get

$$\begin{aligned}\tilde{f}\circ g = &\sum_{\substack{\alpha\subseteq[d]\\|\alpha|<k}} (\Delta_\alpha f)(\emptyset)\cdot 1_{[\max\limits_{\alpha} x_i,\infty]\cap A}\\ &+\sum_j (\Delta_{\sigma_j} f)(\tau_j\backslash\sigma_j)\cdot 1_{[\max\limits_{\sigma_j} x_i,\infty]\cap A}\end{aligned}$$

which is $\mathbf{1}_k$ - $\uparrow$.

In the next step we let g_1 be the d.f. of some probability measure with finite support in A, say $\sum_{\ell=1}^n \lambda_\ell \varepsilon_{a_{1,\ell}}$ with $\lambda_\ell \geq 0$, $\sum_\ell \lambda_\ell = 1$ and $a_{1,\ell} \in A$. Since $\tilde{f}$ is affine as a function of x_1,

$$\tilde{f}\circ g = \sum_\ell \lambda_\ell\, \tilde{f}\circ(g_{1,\ell}, g_2, \ldots, g_d)$$

is again $\mathbf{1}_k$ - $\uparrow$. This procedure is then repeated for $g_2, g_3, \ldots, g_d$, showing our result to be true if each g_i is the d.f. of some probability measure with finite support in A.

Invoking Lemma 3 we may extend the validity to $\mathbf{1}_k$ - $\uparrow$ functions $g_1, \ldots, g_d$ for which $c_i := \sup g_i(A) = 1$ for each i, making use also of the continuity of $\tilde{f}$.

In general we have $c_i \in [0, 1]$, where we may assume $c_i > 0$ for each i. Then $\varphi(x) := \tilde{f}(c \odot x)$ is still multilinear and $\mathbf{1}_k$ - $\uparrow$, so that

$$\varphi \circ (g_1/c_1, \ldots g_d/c_d) = \tilde{f} \circ g$$

is $\mathbf{1}_k$ - $\uparrow$, thereby finishing our proof. □

Theorem 1 deals with fully k-increasing functions, generalizing the case $k = 2$ of increasing super-modular functions. In [3] fully k-alternating functions are dealt with, for which we offer the following general result:

Theorem 2 *If in the situation of Theorem the functions $g_1, \ldots, g_d$ and f are $\mathbf{1}_k$ - $\updownarrow$, then so is h.*

Proof We make use of the very close direct connection between $\mathbf{n}$ - $\uparrow$ and $\mathbf{n}$ - $\updownarrow$ functions in full generality — see [5], Remark (d):

$$\begin{aligned} \varphi : A \longrightarrow \mathbb{R} \text{ is } \mathbf{n}\text{ - }\updownarrow &\iff -\varphi(-\cdot) \text{ is } \mathbf{n}\text{ - }\uparrow \quad \text{on} \quad -A \\ &\iff c - \varphi(-\cdot) \text{ is } \mathbf{n}\text{ - }\uparrow \quad \text{on} \quad -A \quad \forall\, c \in \mathbb{R} \end{aligned}$$

where in our situation (i.e. $n_j \in \{0, 1\}\, \forall j$) $\quad A = \prod_{j=1}^{k} A_j$ with arbitrary non-empty subsets $A_j \subseteq \overline{\mathbb{R}}$. We apply this to $g_1, \ldots, g_d$ and to f :

$$\begin{aligned} g_1, \ldots, g_d : A \longrightarrow [0,1] \text{ are } \mathbf{1}_k\text{ - }\updownarrow \quad &\text{and} \quad f : \{0,1\}^d \longrightarrow [0,1] \text{ is } \mathbf{1}_k\text{ - }\updownarrow \\ \iff 1 - g_i(-\cdot) : -A \longrightarrow [0,1] \text{ is } \mathbf{1}_k\text{ - }\uparrow \ \forall i \\ \text{and } 1 - f(-\cdot) : \{-1, 0\}^d \longrightarrow [0,1] \text{ is } \mathbf{1}_k\text{ - }\uparrow, \end{aligned}$$

where the last statement is equivalent with $1 - f(\mathbf{1}_d - \cdot)$ being $\mathbf{1}_k$ - $\uparrow$ on $\{0, 1\}^d$. By Theorem 1

$$1 - f(\mathbf{1}_d - (\mathbf{1}_d - g(-\cdot))) = 1 - f \circ g(-\cdot) \text{ is } \mathbf{1}_k\text{ - }\uparrow,$$

or, equivalently, $f \circ g$ is $\mathbf{1}_k$ - $\updownarrow$. □

Remark 1 In [3] the functions g_i may be defined on $\{0, 1\}^\ell$ for $\ell \geq k$. This is of course only superficially more general, since by the very definition, being fully k-increasing or alternating, only k variables are considered simultaneously.

Remark 2 For $k = d$ the polynomial $\tilde{f}$ has but non-negative coefficients (cf. (7) above), hence not only $\tilde{f} \circ (g_1, \ldots, g_d)$ is $\mathbf{1}_d$ - $\uparrow$, but even $\tilde{f} \circ (g_1 \times \cdots \times g_d)$ is $\mathbf{1}_{d^2}$ - $\uparrow$ on A^d. For $k < d$ this cannot be expected: take $k = 1, d = 2, f : \{0, 1\}^2 \longrightarrow \mathbb{R}$ defined by $f(0, 0) = 0, f(1, 0) = f(0, 1) = 1 = f(1, 1)$. Then f (and $\tilde{f}$) is increasing, but not $\mathbf{1}_2$-$\uparrow$, we have $\tilde{f}(x_1, x_2) = x_1 + x_2 - x_1x_2$. For increasing functions $g_1, g_2 : [0, 1] \longrightarrow [0, 1]$ the composed, map $\tilde{f} \circ (g_1, g_2) = g_1 + g_2 - g_1g_2$ is still increasing, however the bivariate

$$(\tilde{f} \circ (g_1 \times g_2))(s, t) = g_1(s) + g_2(t) - g_1(s)g_2(t)$$

is not in general $\mathbf{1}_2$-$\uparrow$, for ex. in the case $g_1 = g_2 = id$, with $(\tilde{f} \circ (g_1 \times g_2))(s, t) = s + t - st$.

Example 2 In Example 1 we considered the case $k = 2, d = 3$ and $f(\alpha) := |\alpha| \vee 1$, with

$$\tilde{f}(x) = 1 + x_1x_2 + x_1x_3 + x_2x_3 - x_1x_2x_3$$

f and $\tilde{f}$ are $\mathbf{1}_2$-$\uparrow$.

If now g_1, g_2, g_3 are $\mathbf{1}_2$-$\uparrow$, with values in $[0, 1]$, then by Theorem 1

$$\tilde{f} \circ (g_1, g_2, g_3) = 1 + g_1g_2 + g_1g_3 + g_2g_3 - g_1g_2g_3$$

is also $\mathbf{1}_2$-$\uparrow$, hence a distribution function in case g_1, g_2, g_3 are of this type. Proving this directly, say for differentiable g_i, should be possible, but is certainly cumbersome.

Example 3 Again $k = 2, d = 3$. Let $f : \mathcal{P}([3]) \longrightarrow \mathbb{R}$ be given by $f(\emptyset) = 0, f(\{i\}) = 2$ $(i = 1, 2, 3)$, $f(\{i, j\}) = 4$ $(i \neq j)$ and $f([3]) = 5$. Then f is easily seen to be $\mathbf{1}_2$-$\updownarrow$, but not $\mathbf{1}_3$-$\updownarrow$. Here

$$\tilde{f}(x) = 2(x_1 + x_2 + x_3) - x_1x_2x_3,$$

and for increasing submodular g_1, g_2, g_3 with values in $[0, 1]$ also $\tilde{f} \circ (g_1, g_2, g_3)$ is again increasing and submodular by Theorem 2.

Remark 3 A natural question is to know which (univariate) functions φ "operate" on fully k-increasing (resp. alternating) functions, i.e. have the property that $\varphi \circ f$ is $\mathbf{1}_k$-$\uparrow$ ($\updownarrow$) whenever f is $\mathbf{1}_k$-$\uparrow$ ($\updownarrow$), supposing of course $\varphi \circ f$ to be defined. The answer is provided by Theorem 12 in [6], later (in [5], p. 250) called *Monotone Composition Theorem*: if φ is k-$\uparrow$ ($\updownarrow$) and f is $\mathbf{1}_k$-$\uparrow$ ($\updownarrow$) then also $\varphi \circ f$ is $\mathbf{1}_k$-$\uparrow$ ($\updownarrow$). For a pseudo-Boolean function f on $\{0, 1\}^d$ we saw in Lemma 1 that f is $\mathbf{1}_k$-$\uparrow$ iff $\tilde{f}$ is (on $[0, 1]^d$), and this is likewise true for fully k-alternating f. Hence for k-$\uparrow$ ($\updownarrow$) φ also $\varphi \circ f$, $\varphi \circ \tilde{f}$ and $(\varphi \circ f)^{\sim}$ are $\mathbf{1}_k$-$\uparrow$ ($\updownarrow$), the latter two being different in general. In Example 3 above, with $\varphi(t) := \sqrt{t}$ (k-$\updownarrow$ $\forall k \in \mathbb{N}$), we have

$$\begin{aligned}(\sqrt{f})^{\sim}(x) = \sqrt{2}(x_1 + x_2 + x_3) - 2(\sqrt{2} - 1)(x_1x_2 + x_1x_3 + x_2x_3) \\ +(3\sqrt{2} + \sqrt{5} - 6)\, x_1x_2x_3,\end{aligned}$$

which is $\mathbf{1}_2$-$\updownarrow$ on $[0, 1]^3$, as is also $\sqrt{\tilde{f}(x)}$.

Note that $\varphi : \mathbb{R}_+ \longrightarrow \mathbb{R}_+$ is 2 - $\updownarrow$ iff φ is increasing and concave. And φ is k - $\updownarrow$ $\forall k \in \mathbb{N}$ iff it is a socalled Bernstein function.

Remark 4 When looking at Theorem 1 one might believe that perhaps each $\mathbf{1}_k$ - $\uparrow$ function on $[0, 1]^d$ has the property shown there for multilinear polynomials $\tilde{f}$ arising from a $\mathbf{1}_k$ - $\uparrow$ pseudo-Boolean function f. This is not the case:

Let again $k = 2, d = 3$ and $a := (\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$; the d.f. φ of ε_a (on $[0, 1]^3$) is given by $\varphi = 1_{[a,1]}$, it is (even) $\mathbf{1}_3$ - $\uparrow$. Let further $g_1 = g_2 = g_3$ be the d.f. of the uniform distribution of $[0, 1]^2$, i.e. $g_i(s, t) = st$. Then $\varphi \circ (g_1, g_2, g_3) = 1_A$ with

$$A := \left\{ \{(s, t) \in [0, 1]^2 \mid st \geq \frac{1}{2} \right\}$$

and this function is not $\mathbf{1}_2$ - $\uparrow$ because

$$\left(\Delta^{(1,1)}_{(\frac{1}{2}, \frac{1}{2})} 1_A \right) \left(\frac{1}{2}, \frac{1}{2} \right) = -1.$$

Funding Open Access funding enabled and organized by Projekt DEAL.

References

1. Kempe, D., Kleinberg, J., Tardos, É.: Maximising the spread of influence through a social network. In: KDD, pp. 137–146. ACM (2003)
2. Mossel, E., Roch, S.: Submodularity of influence in social networks: from local to global. SIAM J. Comput. **39**(6), 2176–2188 (2010)
3. Chen, W., Li, Q., Shan, X., Sun, X., Zhang, J.: Higher order monotonicity and submodularity of influence in social networks: from local to global. arXiv: 1803.00666v1 [cs. SI] (2018)
4. Ressel, P.: Monotonicity properties of multivariate distribution and survival functions—with an application to Lévy-frailty copulas. J. Multivar. Anal. **102**, 393–404 (2011)
5. Ressel, P.: Copulas, stable tail dependence functions, and multivariate monotonicity. Depend. Model. **7**, 247–258 (2019)
6. Ressel, P.: Higher order monotonic functions of several variables. Positivity **18**(2), 257–285 (2014)