A survey on the arity gap

Miguel Couceiro, Erkko Lehtonen, Tamas Waldhauser

This paper appeared in J. Mult.-Valued Logic Soft Comput. **24** (2015), 223–249.

A SURVEY ON THE ARITY GAP

MIGUEL COUCEIRO, ERKKO LEHTONEN, AND TAMÁS WALDHAUSER

Abstract. The arity gap of a function of several variables is defined as the minimum decrease in the number of essential variables when essential variables of the function are identified. We present a brief survey on the research done on the arity gap, from the first studies of this notion up to recent developments, and discuss some natural extensions and related problems.

1. Introduction

Let A and B be arbitrary nonempty sets. A *function of several variables from A to B* is a map $f\colon A^n \to B$ for some integer $n \geq 1$ called the *arity* of f. If $A = B$, then we speak of *operations on A.* Operations on the two-element set $\{0, 1\}$ are called *Boolean functions.* We denote the set of all finitary functions from A to B by

$$\mathcal{F}_{AB} := \bigcup_{n \geq 1} B^{A^n}.$$

We say that the i-th variable of $f\colon A^n \to B$ ($1 \leq i \leq n$) is *essential,* if there exist n-tuples

$$(a_1, \ldots, a_{i-1}, a_i, a_{i+1}, \ldots, a_n), (a_1, \ldots, a_{i-1}, a_i', a_{i+1}, \ldots, a_n) \in A^n \tag{1}$$

that only differ in the i-th position, such that

$$f(a_1, \ldots, a_{i-1}, a_i, a_{i+1}, \ldots, a_n) \neq f(a_1, \ldots, a_{i-1}, a_i', a_{i+1}, \ldots, a_n).$$

We refer to the pair of n-tuples in (1) as a *witness of essentiality* of the i-th variable. If the i-th variable of f is not essential, then we say that it is *inessential.* The number of essential variables of f is called the *essential arity* of f and it is denoted by $\operatorname{ess} f$. If $\operatorname{ess} f = m$, then we say that f is *essentially m-ary.*

For $i, j \in \{1, \ldots, n\}$, $i \neq j$, the function $f_{i \leftarrow j}\colon A^n \to B$ given by the rule

$$f_{i \leftarrow j}(a_1, \ldots, a_n) = f(a_1, \ldots, a_{i-1}, a_j, a_{i+1}, \ldots, a_n),$$

for all $a_1, \ldots, a_n \in A$, is called a *variable identification minor* of f, obtained by identifying the i-th variable with the j-th variable. Note that the i-th variable of $f_{i \leftarrow j}$ is necessarily inessential.

Let $f\colon A^n \to B$ be a function with at least two essential variables. The *arity gap* of f is the quantity

$$\operatorname{gap} f := \min_{i \neq j} (\operatorname{ess} f - \operatorname{ess} f_{i \leftarrow j}),$$

where i and j range over the set of indices of essential variables of f. Note that, by definition, $1 \leq \operatorname{gap} f \leq \operatorname{ess} f$.

We say that f and g are *equivalent* if each one can be obtained from the other by permutation of variables and addition or deletion of inessential variables. Whenever we consider the arity gap of a function, we can assume that all of its variables are essential. This is not a significant restriction, because every nonconstant function is equivalent to a function with no inessential variables, and equivalent functions have the same arity gap.

Key words and phrases. functions of several variables; essential variable; arity gap; quasi-arity.

Example 1. Let F be an arbitrary field. Consider the polynomial function $f\colon F^3 \to F$ induced by $x_1x_3 - x_2x_3$. It is clear that all variables of f are essential, i.e., $\operatorname{ess} f = 3$. Let us consider the various variable identification minors of f:

$$\begin{aligned} f_{1\leftarrow 2} &= 0, & f_{2\leftarrow 1} &= 0, \\ f_{1\leftarrow 3} &= x_3^2 - x_2x_3, & f_{3\leftarrow 1} &= x_1^2 - x_1x_2, \\ f_{2\leftarrow 3} &= x_1x_3 - x_3^2, & f_{3\leftarrow 2} &= x_1x_2 - x_2^2. \end{aligned}$$

We have that

$$\begin{aligned} &\operatorname{ess} f_{1\leftarrow 2} = \operatorname{ess} f_{2\leftarrow 1} = 0, \\ &\operatorname{ess} f_{1\leftarrow 3} = \operatorname{ess} f_{3\leftarrow 1} = \operatorname{ess} f_{2\leftarrow 3} = \operatorname{ess} f_{3\leftarrow 2} = 2. \end{aligned}$$

Hence $\operatorname{gap} f = 1$.

Example 2. Let $f\colon \{0,1\}^n \to \{0,1\}$ be the Boolean function induced by the polynomial $x_1 + x_2 + \cdots + x_n$ over the two-element field. Then for each $i \neq j$ we have that $f_{i\leftarrow j}$ is induced by the polynomial

$$\sum_{\ell \in \{1,\ldots,n\} \setminus \{i,j\}} x_\ell.$$

Thus $\operatorname{ess} f = n$ and $\operatorname{ess} f_{i\leftarrow j} = n - 2$ for all $i \neq j$; hence $\operatorname{gap} f = 2$.

Example 3. Let A be a finite set with $k \geq 2$ elements, say, $A = \{0, 1, \ldots, k-1\}$. Let $f\colon A^n \to A$, $2 \leq n \leq k$, be given by the rule

$$f(a_1, \ldots, a_n) := \begin{cases} 1 & \text{if } (a_1, \ldots, a_n) = (0, 1, \ldots, n-1), \\ 0 & \text{otherwise.} \end{cases}$$

It is easy to see that all variables of f are essential, and for all $i \neq j$, the function $f_{i\leftarrow j}$ is identically 0. Hence $\operatorname{gap} f = n$.

As shown by the examples above, every positive integer is the arity gap of some function of several variables. Are all positive integers possible as the arity gaps of functions of several variables from A to B for a fixed domain A and codomain B? Does the size of the domain or the codomain have any influence on the set of possible arity gaps? Or even, could one hope to classify functions according to their arity gap? These questions have been raised and studied by several authors. In this paper, we survey the research work done on this topic over the years. In the bibliography, we indicate literature relevant to the topics of essential variables (see [3, 4, 15, 16, 22, 24, 27]), variable identification minors and variants (see [2, 5, 14, 17, 18, 19, 20, 23, 28, 31]) and arity gap and its variants (see [6, 7, 8, 9, 10, 11, 13, 24, 25, 26, 30]).

This survey is organized as follows. We start in Section 2 with basic notions and classifications of special types of functions (namely, Boolean and pseudo-Boolean functions) according to their arity gap; we also give a general description of arity gap, which serves as a basis for the later sections. Section 3 reports on natural decomposition schemes which arise from arity gap and which are then used in Section 4 to enumerate functions with prescribed arity and arity gap. In Section 5, we make the description of arity gap more explicit for some special classes of functions (namely, order-preserving functions and polynomial functions over fields). Then, in Section 6, we briefly discuss the order which naturally arises from variable identification (namely, simple minor relation) and some parametrized variants of the arity gap. We conclude this survey by mentioning some open problems and proposing new directions for future research in Section 7.

2. From the beginnings to a complete classification

To the best of our knowledge, the first study of arity gap appeared in print in the 1963 paper by Salomaa [24]. In that paper, he addressed the question how the number of essential variables of a function is affected by substitution of constants for variables or by identification of variables. Concerning identification of variables, his main result was the following.

Theorem 1 (Salomaa [24]). *Let $f\colon \{0,1\}^n \to \{0,1\}$ be a Boolean function with at least two essential variables. Then* $\operatorname{gap} f \leq 2$.

As illustrated by Example 2, this upper bound is attained. On the other hand, Example 1 provides a Boolean function of arity gap 1. Since the arity gap is always at least 1, there are no other possible values for the arity gap of Boolean functions. This calls for a complete classification of Boolean functions into those with arity gap 1 and those with arity gap 2. Such a strengthening of Salomaa's result was obtained in [6].

Theorem 2 ([6]). *Let $f\colon \{0,1\}^n \to \{0,1\}$ be a Boolean function with at least two essential variables. Then* $\operatorname{gap} f = 2$ *if and only if f is equivalent to one of the following functions:*

(1) $x_1 + x_2 + \cdots + x_m + c$ *for some* $m \geq 2$,
(2) $x_1x_2 + x_1 + c$,
(3) $x_1x_2 + x_1x_3 + x_2x_3 + c$,
(4) $x_1x_2 + x_1x_3 + x_2x_3 + x_1 + x_2 + c$,

where $c \in \{0,1\}$. Otherwise $\operatorname{gap} f = 1$.

The number of n-ary Boolean functions f with $\operatorname{ess} f = n$ and $\operatorname{gap} f = 2$ can be straightforwardly read off from the list of functions given in Theorem 2. Since the number of functions $f\colon A^n \to B$ ($|A| = k$, $|B| = \ell$) with $\operatorname{ess} f = m$ ($0 \leq m \leq n$) is

$$\binom{n}{m} \sum_{i=0}^{m} (-1)^i \binom{m}{i} \ell^{k^{m-i}} \tag{2}$$

(see Wernick [29]), the number of n-ary Boolean functions with $\operatorname{ess} f = n$ and $\operatorname{gap} f = 1$ can be easily obtained.

Theorem 2 was later reproved using different techniques by Shtrakov in [25], where he also addressed the problem of counting the number of Boolean functions with a given arity gap.

We call functions of several variables from $\{0,1\}$ to an arbitrary nonempty set B *pseudo-Boolean functions.* Theorem 2 was extended for pseudo-Boolean functions in [7].

Theorem 3 ([7]). *Let $f\colon \{0,1\}^n \to B$ ($n \geq 2$) be a pseudo-Boolean function all of whose variables are essential. Then* $\operatorname{gap} f = 2$ *if and only if f satisfies one of the following conditions:*

(1) $n = 2$ *and f is a nonconstant function satisfying* $f(0,0) = f(1,1)$,
(2) $f = g \circ h$, *where $g\colon \{0,1\} \to B$ is injective and $h\colon \{0,1\}^n \to \{0,1\}$ is a Boolean function with* $\operatorname{gap} h = 2$, *as given by Theorem 2.*

Otherwise $\operatorname{gap} f = 1$.

A *partial function of several variables from A to B* is a map $f\colon S \to B$ where $S \subseteq A^n$ for some integer $n \geq 1$, called the *arity* of f. If $S = A^n$, we speak of *total functions.* Essential variables are defined for partial functions in the same way as for total functions, but the n-tuples of (1) forming a witness of essentiality are required to be in S. In this way, the notions of essential arity and arity gap can be naturally extended to partial functions.

The following notions were introduced by Berman and Kisielewicz [1]. Denote by $\mathcal{P}(A)$ the power set of A, and define the function oddsupp: $\bigcup_{n \geq 1} A^n \to \mathcal{P}(A)$ by

$$\operatorname{oddsupp}(a_1, \ldots, a_n) := \Big\{ a \in A : \big| \{ j \in \{1, \ldots, n\} : a_j = a \} \big| \text{ is odd} \Big\}.$$

We say that a partial function $f\colon S \to B$ ($S \subseteq A^n$) is *determined by* oddsupp if there exists a function $f^*\colon \mathcal{P}(A) \to B$ such that

$$f = f^* \circ \operatorname{oddsupp}|_S.$$

Willard showed that if $f\colon A^n \to B$, where A is finite, $\operatorname{ess} f = n > \max(|A|, 3)$ and $\operatorname{gap} f \geq 2$, then f is determined by oddsupp. It is easy to verify that if f is determined by oddsupp and $\operatorname{ess} f \geq 2$, then $\operatorname{gap} f = 2$.

Theorem 4 (Willard [30]). *Let A and B be finite nonempty sets, and let $k := |A|$. Suppose that $f\colon A^n \to B$ depends on all of its variables. If $n > k$, then $\operatorname{gap} f \leq 2$. Moreover, if $n > \max(k, 3)$, then $\operatorname{gap} f = 2$ if and only if f is determined by* oddsupp.

This theorem deals only with functions over finite domains with sufficiently large essential arities; if A and B are finite, then $\operatorname{gap} f \leq 2$ for almost all functions. The condition on essential arity was removed in [7], where a full classification of finite functions according to their arity gap was given in terms of so-called quasi-arity. As pointed out in [9], this classification theorem holds more generally – with no change needed in its proof – for functions with arbitrary domains. Thus, the study of arity gap culminated in Theorem 5, which will be presented below, after introducing some terminology.

For $n \geq 2$, let

$$A^n_{=} := \{(a_1, \ldots, a_n) \in A^n : a_i = a_j \text{ for some } i \neq j\}.$$

Let $f\colon A^n \to B$. Any function $g\colon A^n \to B$ satisfying $f|_{A^n_=} = g|_{A^n_=}$ is called a *support* of f. The *quasi-arity* of f, denoted $\operatorname{qa} f$, is defined as the minimum of the essential arities of all supports of f, i.e., $\operatorname{qa} f := \min_g \operatorname{ess} g$ where g ranges over all supports of f. If $\operatorname{qa} f = m$, then we say that f is *quasi-m-ary*. Note that if A is finite, then $A^n_= = A^n$ whenever $n > |A|$. Hence $\operatorname{qa} f = \operatorname{ess} f$ whenever $n > |A|$. Moreover, $\operatorname{qa} f = \operatorname{ess} f|_{A^n_=}$ whenever $n \neq 2$.

Theorem 5 ([7, 9]). *Let A and B be arbitrary sets with at least two elements. Suppose that $f\colon A^n \to B$, $n \geq 2$, depends on all of its variables.*

(1) *For $3 \leq p \leq n$, $\operatorname{gap} f = p$ if and only if $\operatorname{qa} f = n - p$.*
(2) *For $n \neq 3$, $\operatorname{gap} f = 2$ if and only if $\operatorname{qa} f = n-2$ or $\operatorname{qa} f = n$ and $f|_{A^n_=}$ is determined by* oddsupp.
(3) *For $n = 3$, $\operatorname{gap} f = 2$ if and only if there is a nonconstant unary function $h\colon A \to B$ and $i_1, i_2, i_3 \in \{0, 1\}$ such that*

$$f(x_1, x_0, x_0) = h(x_{i_1}),$$
$$f(x_0, x_1, x_0) = h(x_{i_2}),$$
$$f(x_0, x_0, x_1) = h(x_{i_3}).$$

(4) *Otherwise* $\operatorname{gap} f = 1$.

Theorem 5 gives answers to the questions posed at the end of Section 1. The set of possible arity gaps of functions of several variables from A to B is $\{n \in \mathbb{N} \setminus \{0\} : n \leq |A|\}$. This does not depend on the cardinality of the codomain B, as long as $|B| \geq 2$ (otherwise, all functions are constant and hence have no essential variables).

3. Decompositions of functions based on the arity gap

Theorem 5 provides a complete classification of functions according to their arity gap. Unfortunately, it is not as explicit as Theorems 2 and 3 that deal with Boolean and pseudo-Boolean functions. The main reason for this is the fact that no special structure was assumed on the sets A and B. By assuming that the codomain B has a group structure, more explicit descriptions can be obtained. In this direction, two cases are distinguished.

First, functions $f\colon A^n \to B$ with arity gap $p \geq 3$ are shown to be decomposable into a sum of a quasi-nullary function and an essentially $(n-p)$-ary function.

Theorem 6 ([9]). *Assume that $(B; +)$ is a group with neutral element 0. Let $f\colon A^n \to B$, $n \geq 3$, and $3 \leq p \leq n$. Then the following two conditions are equivalent:*

(1) $\operatorname{ess} f = n$ *and* $\operatorname{gap} f = p$.
(2) *There exist functions $g, h\colon A^n \to B$ such that $f = h + g$, $h|_{A^n_=} \equiv 0$, $h \not\equiv 0$, and* $\operatorname{ess} g = n - p$.

The decomposition $f = h + g$ given above is unique.

Similar decompositions were presented by Shtrakov and Koppitz [26], without proving uniqueness.

Second, we deal with functions with arity gap 2. In this case we need to further assume that B is a Boolean group (i.e., a group satisfying $x + x = 0$ for all $x \in B$). We also need to introduce some notation. Let $\varphi\colon A^{n-2} \to B$ be a function that is determined by oddsupp, i.e., $\varphi = \varphi^* \circ \operatorname{oddsupp}|_{A^{n-2}}$, for some function $\varphi^*\colon \mathcal{P}(A) \to B$. Let $\widetilde{\varphi}$ be the n-ary function defined by

$$\widetilde{\varphi}(x_1, \dots, x_n) := \sum_{\substack{k<n \\ 2|n-k}} \sum_{1 \leq i_1 < \dots < i_k \leq n} \varphi^*(\operatorname{oddsupp}(x_{i_1}, \dots, x_{i_k})). \tag{3}$$

Observe that each summand is a variable identification minor of φ, namely

$$\varphi^*(\operatorname{oddsupp}(x_{i_1}, \dots, x_{i_k})) = \varphi(x_{i_1}, \dots, x_{i_k}, y, \dots, y),$$

where the number of y's is $n - 2 - k$, which is an even number; therefore y is indeed an inessential variable of the function on the right-hand side; moreover, the order of the variables is irrelevant. The function $\widetilde{\varphi}$ is obviously totally symmetric, and it can be shown that $\widetilde{\varphi}|_{A^n_=}$ is determined by oddsupp; from these facts it follows that $\widetilde{\varphi}$ is determined by oddsupp as well (see [9] for details).

Theorem 7 ([9])**.** *Assume that $(B;+)$ is a Boolean group with neutral element 0. Let $f\colon A^n \to B$, $n \geq 4$. Then the following two conditions are equivalent:*

(1) $\operatorname{ess} f = n$ *and* $\operatorname{gap} f = 2$.
(2) *There exist functions $g, h\colon A^n \to B$ such that $f = h + g$, $h|_{A^n_=} \equiv 0$, and either*
  (a) $\operatorname{ess} g = n - 2$ *and* $h \not\equiv 0$, *or*
  (b) *$g = \widetilde{\varphi}$ for some nonconstant $(n-2)$-ary function φ that is determined by* oddsupp.

The decomposition $f = h + g$ given above is unique.

Theorem 2 shows that for all but finitely many Boolean functions, $\operatorname{gap} f = 2$ implies that f is a sum of unary functions. The following theorem provides an analogous result for functions $f\colon A^n \to B$, where A is a finite set and B is a Boolean group.

Theorem 8 ([12])**.** *Let A be a finite set, B a Boolean group, and let $f\colon A^n \to B$ be a function such that* $\operatorname{ess} f > \max(|A|, 3)$. *Then* $\operatorname{gap} f \leq 2$, *and if* $\operatorname{gap} f = 2$, *then f is a sum of functions of essential arity at most $|A| - 1$.*

Let f be a polynomial function over $\mathrm{GF}(q)$, the q-element finite field, with $\operatorname{ess} f = n > \max(q, 3)$ and $\operatorname{gap} f = 2$. By the above theorem, if q is even, then f is a sum of functions of arity at most $q - 1$. Since finite fields are functionally complete, we can represent each summand by a polynomial. This implies that f is a sum of monomials, where each monomial involves at most $q - 1$ variables. If $q = 2$, then we can conclude that every Boolean function f with $\operatorname{ess} f \geq 4$ and $\operatorname{gap} f = 2$ is linear, in accordance with Theorem 2.

The next example shows that Theorem 7 and Theorem 8 do not hold for arbitrary groups B.

Example 4 ([11])**.** Let q be an odd prime power, and let f be the following polynomial function over $\mathrm{GF}(q)$:

$$f(x_1, \dots, x_n) = \prod_{i=1}^{n} \left(x_i^{q-1} - \frac{1}{2} \right),$$

Here $\frac{1}{2}$ stands for the multiplicative inverse of $2 = 1 + 1$ (it exists, since $\mathrm{GF}(q)$ is of odd characteristic). Let us identify the first two variables of f:

$$\begin{aligned} f(x_1, x_1, x_3, \dots, x_n) &= \left(x_1^{q-1} - \frac{1}{2} \right)^2 \cdot \prod_{i=3}^{n} \left(x_i^{q-1} - \frac{1}{2} \right) \\ &= \left(x_1^{2q-2} - x_1^{q-1} + \frac{1}{4} \right) \cdot \prod_{i=3}^{n} \left(x_i^{q-1} - \frac{1}{2} \right) \\ &= \frac{1}{4} \cdot \prod_{i=3}^{n} \left(x_i^{q-1} - \frac{1}{2} \right), \end{aligned}$$

since $x_1^q = x_1$ holds identically in $\mathrm{GF}(q)$. We see that x_1 becomes an inessential variable, and this together with the symmetry of f shows that f is determined by oddsupp, hence $\operatorname{gap} f = 2$.

If f were a sum of at most $(n-1)$-ary functions, then every monomial of f would involve at most $n-1$ variables. However, this is clearly not the case, as the expansion of f involves the monomial $x_1^{q-1} \cdots x_n^{q-1}$, which will not be cancelled by any other monomial. Thus f cannot be expressed as a sum of functions of arity smaller than the arity of f. If $n > q$, then the only support of f is itself, thus a support g satisfying condition 2a) or 2b) of Theorem 7 cannot exist.

Theorem 8 and Example 4 ask for a characterization of those groups B for which Theorem 8 still holds. This question was partially answered in [12], where abelian groups B were classified according to whether every function $f\colon A^n \to B$ with $\operatorname{gap} f = 2$ can be decomposed into a sum of functions with a smaller number of essential variables. This result followed from a study of a hierarchy of function classes based on decomposability, which we now describe.

Let $(B;+)$ be an abelian group with neutral element 0. Recall that the *order* of a group element $b \in B$ is the smallest positive integer n such that $\underbrace{b + \cdots + b}_{n \text{ times}} = 0$. If such a number does not exist, the order of b is ∞. If the orders of the elements of B have a finite common upper bound, then the *exponent* of B is defined as the least upper bound (or, equivalently, the least common multiple) of the orders of its members; otherwise the exponent of B is ∞. For example, the exponent of a Boolean group is 2.

A function $f\colon A^n \to B$ is said to be *k-decomposable,* if it has an additive decomposition $f = f_1 + \cdots + f_s$, where each $f_i\colon A^n \to B$ $(1 \leq i \leq s)$ has essential arity at most k. If f is $(n-1)$-decomposable, we simply say that it is *decomposable.*

The following result provides a characterization of k-decomposable functions. For $I \subseteq [n]$ and $\mathbf{a} = (a_1, \dots, a_n) \in A^n$, let $\mathbf{x}_I^{\mathbf{a}}$ denote the n-tuple which is obtained from $\mathbf{x} = (x_1, \dots, x_n) \in A^n$ by replacing its i-th component by a_i, for each $i \in I$.

Proposition 1 ([12])**.** *Let $(B;+)$ be an abelian group, and let $c \in A$. A function $f\colon A^n \to B$ is k-decomposable if and only if for all $\mathbf{a} \in A^n$ and $I \subseteq [n]$ with $|I| > k$, we have*

$$\sum_{J \subseteq I} (-1)^{|I \setminus J|} f(\mathbf{c}_J^{\mathbf{a}}) = 0,$$

where $\mathbf{c} := (c, \dots, c)$.

Proposition 1 gave rise to the following result which reveals a dichotomy of abelian groups with respect the decomposability of functions determined by oddsupp.

Theorem 9 ([12])**.** *Let A be an arbitrary set with at least two elements, and let B be an abelian group. Then every function $f\colon A^n \to B$ that is determined by* oddsupp *is decomposable if and only if A is finite and the exponent of B is a power of 2. Moreover, if A is finite and the exponent of B is 2^e, then every function determined by* oddsupp *is $(|A| + e - 2)$-decomposable.*

The following example shows that the bound $(|A| + e - 2)$ in Theorem 9 cannot be improved.

Example 5. Let $A = \{0, 1, \dots, k-1\}$, and let B be an arbitrary abelian group of exponent 2^e. Fix an element $b \in B$ of order 2^e. Let $\varphi\colon \mathcal{P}(A) \to B$ be defined by

$$\varphi(T) = \begin{cases} b, & \text{if } T \supseteq A \setminus \{0\}, \\ 0, & \text{otherwise,} \end{cases}$$

let $n \geq k + e - 2$, and let $f\colon A^n \to B$ be given by $f(\mathbf{x}) = \varphi(\operatorname{oddsupp}(\mathbf{x}))$. Theorem 9 asserts that f is $(k+e-2)$-decomposable; however, it can be shown that f is not $(k+e-3)$-decomposable (see [12]).

4. The number of functions with a given arity gap

The unique decompositions provided by Theorems 6 and 7 enable us to actually count the number of functions $f\colon A^n \to B$ with $\operatorname{ess} f = n \geq 2$ and $\operatorname{gap} f = p$ for every $1 \leq p \leq n$. The problem of determining these numbers for operations on finite sets was raised by Shtrakov and Koppitz in [26], where upper bounds were provided. An answer to the counting problem with exact numbers was given in [9].

Let A and B be finite sets with $|A| = k$, $|B| = \ell$. Let us denote by $G^{k\ell}_{np}$ the number of functions $f\colon A^n \to B$ with $\operatorname{ess} f = n$ and $\operatorname{gap} f = p$. It is well known (see Wernick [29]) that the number of functions $g\colon A^n \to B$ that depend on exactly r variables ($0 \leq r \leq n$) is

$$U^{k\ell}_{nr} := \binom{n}{r} \sum_{i=0}^{r} (-1)^i \binom{r}{i} \ell^{k^{r-i}}.$$

Let $(m)_i$ denote the *falling factorial* $(m)_i := m(m-1)\cdots(m-(i-1))$. The number of functions $h\colon A^n \to B$ such that $h|_{A^n_=} \equiv 0$, $h \not\equiv 0$ is then given by

$$V^{k\ell}_n := \ell^{(k)_n} - 1.$$

Let us denote by $O^{k\ell}_n$ the number of functions $f\colon A^n \to B$ such that $\operatorname{ess} f = n$, $\operatorname{qa} f = n$ and $f|_{A^n_=}$ is determined by oddsupp. It can be shown (see [9]) that for $k \geq 2$, $\ell \geq 2$, $n \geq 2$,

$$O^{k\ell}_n = \begin{cases} \ell^{2^{k-1}} - \ell, & \text{if } n > k, \\ \ell^{(k)_n}(\ell^{S^k_n} - \ell), & \text{if } n \leq k, \end{cases}$$

where

$$S^k_n = \begin{cases} \sum_{i=0}^{\frac{n}{2}-1} \binom{k}{2i}, & \text{if } n \text{ is even}, \\ \sum_{i=0}^{\frac{n-1}{2}-1} \binom{k}{2i+1}, & \text{if } n \text{ is odd}. \end{cases}$$

With the notation and facts given above, we can now provide the number $G^{k\ell}_{np}$ of functions $f\colon A^n \to B$ with $\operatorname{ess} f = n$ and $\operatorname{gap} f = p$.

Theorem 10 ([9]). *Let $k \geq 2$, $\ell \geq 2$, $n \geq 2$.*

(1) *If $n > k$ and $3 \leq p \leq n$, then $G^{k\ell}_{np} = 0$.*
(2) *If $n > k$ and $n \geq 4$, then*

$$G^{k\ell}_{n2} = O^{k\ell}_n = \ell^{2^{k-1}} - \ell, \qquad G^{k\ell}_{n1} = U^{k\ell}_{nn} - G^{k\ell}_{n2}.$$

(3) *If $3 \leq n \leq k$ and $3 \leq p \leq n$, then $G^{k\ell}_{np} = U^{k\ell}_{n(n-p)} V^{k\ell}_n$.*
(4) *If $4 \leq n \leq k$, then*

$$G^{k\ell}_{n2} = U^{k\ell}_{n(n-2)} V^{k\ell}_n + O^{k\ell}_n, \qquad G^{k\ell}_{n1} = U^{k\ell}_{n(n-1)} V^{k\ell}_n + U^{k\ell}_{nn} \ell^{(k)_n} - V^{k\ell}_n \ell^{k^n} - O^{k\ell}_n$$

(5) $G^{k\ell}_{32} = (8\ell^{(k)_3} - 3)(\ell^k - \ell)$, $G^{k\ell}_{31} = U^{k\ell}_{33} - G^{k\ell}_{33} - G^{k\ell}_{32}$.
(6) $G^{k\ell}_{22} = \ell^{(k)_2+1} - \ell$, $G^{k\ell}_{21} = U^{k\ell}_{22} - G^{k\ell}_{22}$.

5. Special classes of functions

Theorems 5, 6 and 7 lack the explicitness that is enjoyed by Theorems 2 and 3 in the sense that in the former ones the arity gap of each function is described in terms of the essential arity of some related function. However, for some special classes of functions, our results become explicit. Examples include pseudo-Boolean functions (see Theorem 3), Lovász extensions (in particular, the so-called Choquet integrals), multilinear polynomial functions and lattice polynomial functions (in particular, the so-called Sugeno integrals). In this section we extend these results to other classes of functions, namely of order-preserving functions (Subsection 5.1) and of polynomial functions over fields (Subsection 5.2), and provide rather explicit descriptions of arity gap within these classes.

5.1. **Order-preserving functions.** Let $(A; \leq_A)$ and $(B; \leq_B)$ be partially ordered sets. A function $f\colon A^n \to B$ is said to be *order-preserving* (with respect to the partial orders $\leq_A$ and $\leq_B$) if for all $\mathbf{a}, \mathbf{b} \in A^n$, $f(\mathbf{a}) \leq_B f(\mathbf{b})$ whenever $\mathbf{a} \leq_A \mathbf{b}$, where $\mathbf{a} \leq_A \mathbf{b}$ denotes the componentwise ordering of tuples, i.e., $\mathbf{a} \leq_A \mathbf{b}$ if and only if $a_i \leq_A b_i$ for all $i \in \{1, \dots, n\}$. We say that $(A; \leq_A)$ is *bidirected* if every pair of elements of A has both an upper bound and a lower bound.

Theorem 11 ([10])**.** *Let $(A; \leq_A)$ be a bidirected poset, let $(B; \leq_B)$ be any poset, and let $f\colon A^n \to B$ $(n \geq 2)$ be an order-preserving function such that* $\operatorname{ess} f = n$. *Then* $\operatorname{gap} f = 2$ *if and only if $n = 3$ and there is a nonconstant order-preserving unary function $h\colon A \to B$ such that*
$$f(x_1, x_0, x_0) = f(x_0, x_1, x_0) = f(x_0, x_0, x_1) = h(x_0).$$
Otherwise, $\operatorname{gap} f = 1$.

By imposing stronger assumptions on the underlying posets, we obtain more stringent descriptions of order-preserving functions with arity gap 2. For example, if $(A; \leq_A)$ and $(B; \leq_B)$ are lattices and the function h occurring in Theorem 11 is a lattice homomorphism whose image is a distributive sublattice of B, then it can be shown that
$$f = \operatorname{med}\bigl(h(x_1), h(x_2), h(x_3)\bigr),$$
where med denotes the ternary median function on $\operatorname{Im} h$, given by
$$\begin{aligned} \operatorname{med}(a, b, c) &:= (a \wedge b) \vee (a \wedge c) \vee (b \wedge c) \\ &= (a \vee b) \wedge (a \vee c) \wedge (b \vee c). \end{aligned}$$
From this observation and Theorem 11, we obtain the following explicit classification of order-preserving functions defined on chains.

Corollary 1 ([10])**.** *Let $(A; \leq_A)$ be a chain and let $(B; \leq_B)$ be any lattice. Let $f\colon A^n \to B$ be an order-preserving function. Then* $\operatorname{gap} f = 2$ *if and only if $n = 3$ and $f = \operatorname{med}\bigl(h(x_1), h(x_2), h(x_3)\bigr)$ for some nonconstant order-preserving unary function $h\colon A \to B$ (here* med *denotes the median function on* $\operatorname{Im} h$*). Otherwise* $\operatorname{gap} f = 1$.

5.2. **Polynomial functions over fields.** As in the case of order-preserving functions, Theorem 5 can be refined for polynomial functions over fields. We summarize in this section the results obtained in [11] in this direction.

Polynomials over infinite fields are in one-to-one correspondence with polynomial functions. It is well known that every function over a finite field is a polynomial function, but the correspondence between polynomials and functions is not injective. This correspondence can be made bijective by assuming that, for a finite field $\mathrm{GF}(q)$, we only consider polynomials in which the exponent of each variable in every monomial is at most $q-1$; we call such polynomials *canonical.* Every polynomial over an infinite field is *canonical.*

Given a polynomial function $f\colon F^n \to F$, we denote by P_f the unique canonical polynomial which induces f, and given a polynomial $p \in F[x_1, \dots, x_n]$, we denote by $\overline{p}$ the function $f\colon F^n \to F$ induced by p. Note that $\overline{p} + \overline{q} = \overline{p + q}$ for all $p, q \in F[x_1, \dots, x_n]$.

Lemma 1 ([11])**.** *If f is a polynomial function over a field F, then the functions g and h in the decomposition $f = h + g$ given in Theorem 6 are also polynomial functions.*

Lemma 2 ([11])**.** *If h is an n-ary polynomial function over a field F, then $h|_{F^n_{=}} \equiv 0$ if and only if h is induced by a multiple of the polynomial*
$$\Delta_n = \prod_{1 \leq i < j \leq n} (x_i - x_j) \in F[x_1, \dots, x_n].$$

With an application of the two lemmas above, Theorem 6 about functions with arity gap at least 3 particularizes for polynomial functions as follows.

Theorem 12 ([11])**.** *Let F be a field and let $f\colon F^n \to F$ be a polynomial function of arity at least 4. Then* $\operatorname{gap} f = p \geq 3$ *if and only if there exist polynomials $P, Q \in F[x_1, \dots, x_n]$ such that $f = \overline{P} + \overline{Q}$, P is canonical, exactly $n - p$ variables occur in P, and Q is a multiple* *of the polynomial* Δ_n *such that* $\overline{Q}$ *is not identically* 0. *Moreover, if* $f = \overline{P'} + \overline{Q'}$, *where* P' *is canonical,* $n-p$ *variables occur in* P' *and* Q' *is a multiple of* Δ_n *such that* $\overline{Q'}$ *is not identically* 0, *then* $P' = P$ *and* $\overline{Q'} = \overline{Q}$.

Theorem 12 only deals with polynomial functions with arity gap at least 3, which thus asks for the study of the polynomial functions with arity gap at most 2. In particular, we consider polynomial functions determined by oddsupp, and we will see that the characteristic of the underlying field plays a crucial role here. We first introduce some notation that will be needed to state the subsequent results. Let F be an arbitrary field.

- If F is infinite, then N_F denotes the set of nonnegative integers, and $\oplus_F$ denotes the usual addition of nonnegative integers.
- If F has finite order q, then N_F denotes the set $\{0, 1, \ldots, q-1\}$, and $\oplus_F$ is the operation on N_F given by the following rules:
  - $0 \oplus_F 0 = 0$.
  - If $a \neq 0$ or $b \neq 0$, then $a \oplus_F b = c$, where c is the unique number in $\{1, \ldots, q-1\}$ such that $c \equiv a + b \pmod{q-1}$.

Proposition 2 ([11])**.** *Let* F *be a field, and let* $f \colon F^n \to F$ *be a polynomial function with*

$$P_f = \sum_{\mathbf{k}=(k_1,\ldots,k_n)\in N_F^n} c_{\mathbf{k}} x_1^{k_1} x_2^{k_2} \cdots x_n^{k_n}.$$

Then f *is determined by* oddsupp *if and only if*

(A) f *is symmetric, i.e.,* $c_{(k_1,\ldots,k_n)} = c_{(l_1,\ldots,l_n)}$ *whenever there exists a permutation* $\pi \in S_n$ *such that* $k_i = l_{\pi(i)}$ *for all* $i \in [n]$, *and*

(B) *for all* $(k, k_3, \ldots, k_n) \in N_F^{n-1}$ *with* $k \neq 0$,

$$\sum_{\substack{(a_1,a_2)\in N_F^2 \\ a_1 \oplus_F a_2 = k}} c_{(a_1,a_2,k_3,\ldots,k_n)} = 0.$$

In particular, if the characteristic of F *is* 2, *then* f *is determined by* oddsupp *if and only if condition* (A) *above holds together with*

(B_2) $c_{(k,k,k_3,\ldots,k_n)} = 0$ *for all* $(k, k, k_3, \ldots, k_n) \in N_F^n$ *with* $k \neq 0$.

In order to obtain more explicit descriptions of polynomial functions determined by oddsupp, we need to take into account the characteristic of the underlying field. We start with finite fields and fields of characteristic 2.

Proposition 3 ([11])**.** *Let* F *be a field, and let* $f \colon F^n \to F$ *be a polynomial function. If* F *is finite or the characteristic of* F *is* 2, *then* $f|_{F^n_=}$ *is determined by* oddsupp *if and only if there exist polynomials* $P, Q \in F[x_1, \ldots, x_n]$ *such that* $f = \overline{P} + \overline{Q}$, $\overline{P}$ *is determined by* oddsupp, *and* Q *is a multiple of the polynomial* Δ_n.

Theorem 13 ([11])**.** *Let* F *be a field of characteristic* 2, *possibly infinite, and let* $f \colon F^n \to F$ *be a polynomial function of arity at least* 4 *which depends on all of its variables. Then* $\operatorname{gap} f = p \geq 2$ *if and only if there exist polynomials* $P, Q \in F[x_1, \ldots, x_n]$ *such that* $f = \overline{P} + \overline{Q}$, P *is canonical,* Q *is a multiple of the polynomial* Δ_n, *and either*

(1) *exactly* $n - p$ *variables occur in* P *and* $\overline{Q} \neq 0$, *or*

(2) P *is not a constant polynomial and* $\overline{P}$ *satisfies conditions* (A) *and* (B_2) *of Proposition 2.*

Otherwise $\operatorname{gap} f = 1$.

Corollary 2 ([11])**.** *Let* $F = \mathrm{GF}(q)$, *where* q *is a power of* 2, *and let* $f \colon F^n \to F$ *be a polynomial function of essential arity* $n > \max(q, 3)$. *If* $\operatorname{gap} f = 2$, *then* f *can be decomposed into a sum of functions of essential arity at most* $q - 1$.

We now consider the case of fields of characteristic 0.

Lemma 3 ([11])**.** *Let* F *be a field of characteristic* 0, *let* $n \geq 2$, *and let* $f \colon F^n \to F$ *be a polynomial function. If* $f|_{F^n_=}$ *is determined by* oddsupp, *then* $f|_{F^n_=}$ *is constant, i.e.,* $\operatorname{qa} f = 0$.

Theorem 14 ([11]). *Let F be a field of characteristic 0, let $n \geq 2$, and let $P \in F[x_1, \ldots, x_n]$ be a polynomial such that all n variables occur in P. Then $\operatorname{gap} \overline{P} = p \geq 2$ if and only if there exist polynomials $Q, R \in F[x_1, \ldots, x_n]$ such that $P = Q + R$, exactly $n - p$ variables occur in Q, and R is a nonzero multiple of the polynomial Δ_n. Otherwise $\operatorname{gap} \overline{P} = 1$. Moreover, the decomposition $P = Q + R$, when it exists, is unique.*

As the following example illustrates, Proposition 3 and Lemma 3 do not extend to (infinite) fields of odd characteristic.

Example 6. Let F be an arbitrary field of characteristic 3, and let $f\colon F^3 \to F$ be the polynomial function induced by

$$2x^3 + 2y^3 + 2z^3 + yz^2 - xy^2 - xz^2 + y^2z + 2xyz. \tag{4}$$

It is straightforward to verify that

$$f(x, x, y) = f(x, y, x) = f(y, x, x) = 2y^3.$$

Hence $f|_{F^3_=}$ is determined by oddsupp but $f|_{F^3_=}$ is not constant. This shows that Lemma 3 does not hold if F has characteristic 3.

Next we show that Proposition 3 does not hold for infinite fields of characteristic 3. Assume now that F is infinite, and let f be induced by (4). Suppose that $g\colon F^3 \to F$ is a polynomial function determined by oddsupp induced by the canonical polynomial

$$\sum_{(k_1,k_2,k_3) \in \mathbb{N}^3} c_{(k_1,k_2,k_3)} x_1^{k_1} x_2^{k_2} x_3^{k_3}.$$

Condition (B) of Proposition 2 yields the following equalities:

$$\begin{aligned}
&c_{(3,0,0)} + c_{(2,1,0)} + c_{(1,2,0)} + c_{(0,3,0)} = 0,\\
&c_{(2,0,1)} + c_{(1,1,1)} + c_{(0,2,1)} = 0,\\
&c_{(1,0,2)} + c_{(0,1,2)} = 0.
\end{aligned}$$

Taking into account the total symmetry of g (condition (A)) and the fact that the characteristic of F is not 2, the only solution to this system of equations is $c_{(k_1,k_2,k_3)} = 0$ for all $(k_1, k_2, k_3) \in \mathbb{N}^3$ such that $k_1 + k_2 + k_3 = 3$. Thus, the canonical polynomial of $g(x, x, x)$ does not contain any cubic term; therefore it cannot coincide with $f(x, x, x) = 2x^3$, and we conclude that $f|_{F^3_=} \neq g|_{F^3_=}$.

6. Further directions

In this section, we briefly discuss two further research directions: the study of an order which naturally arises from variable identification (Subsection 6.1) and the consideration of extensions to the notion of arity gap (Subsection 6.2).

6.1. The simple minor order and its respective covering relation.

Let $f\colon A^m \to B$ and $g\colon A^n \to B$. We say that f is a *simple minor* of g and we write $f \leq g$, if there exists a function $\sigma\colon \{1, \ldots, n\} \to \{1, \ldots, m\}$ such that

$$f(a_1, \ldots, a_m) = g(a_{\sigma(1)}, \ldots, a_{\sigma(n)}),$$

for all $a_1, \ldots, a_m \in A$. Loosely speaking, f is a simple minor of g, if f can be obtained from g by permutation of variables, identification of variables, and addition and deletion of inessential variables.

The simple minor relation $\leq$ on $\mathcal{F}_{AB}$ is both reflexive and transitive, and thus it constitutes a quasiordering of $\mathcal{F}_{AB}$. As for quasiorders, $\leq$ induces an equivalence relation $\equiv$ on $\mathcal{F}_{AB}$: for $f, g \in \mathcal{F}_{AB}$, we set $f \equiv g$ if $f \leq g$ and $g \leq f$, and in this case we say that f and g are *equivalent.* As observed in the introduction, f and g are equivalent if each of f and g can be obtained from the other by permutation of variables, addition of inessential variables, and deletion of inessential variables. Note that $\operatorname{ess} f = \operatorname{ess} g$ whenever $f \equiv g$, and that every nonconstant function is equivalent to a function that depends on all of its variables.

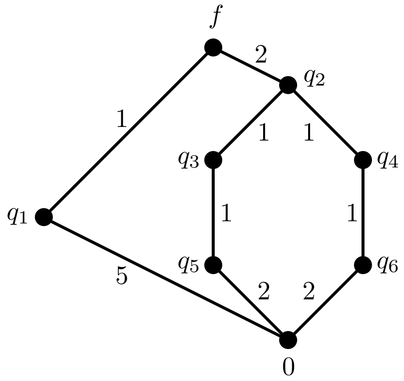

FIGURE 1. The simple minors of the function f given in Example 7.

Remark 1. These and other observations were made for the case $A = B = \{0, 1\}$ in the paper [14], in which it was shown that the poset of equivalence classes of functions induced by simple minor order is equimorphic (i.e., equivalent w.r.t. embeddings) to the poset made of finite subsets of positive integers ordered by inclusion. This result was later extended in various ways in [21], in particular to arbitrary sets A and B.

We say that g is a *strict minor* of f, denoted by $g < f$, if $g \leq f$ but $f \not\equiv g$. If $g < f$ but there is no h such that $g < h < f$, then we say that g is a *lower cover* of f and we denote this fact by $g \prec f$.

Remark 2. It was shown in [2] that the lower covers of any function $f\colon A^n \to B$ have the same essential arity when $A = B = \{0, 1\}$. The proof of this fact given in [2] actually shows that this claim is true whenever $|A| = 2$ and $|B| \geq 2$. However, as the following example shows, this is not the case when $|A| > 2$.

Example 7. Let A be the 5-element field and consider the polynomial function $f\colon A^6 \to A$ defined by:

$$f(x_1, x_2, x_3, x_4, x_5, x_6) := (x_1 - x_2)(x_5 - x_6) + \prod_{\substack{1 \leq i < j \leq 6 \\ (i,j) \neq (5,6)}} (x_i - x_j).$$

It is easy to verify that $\operatorname{ess} f = 6$, $f_{1\leftarrow 2} \equiv 0$, and that f has, up to equivalence, two lower covers, namely,

$$\begin{aligned} f_{1\leftarrow 3} &= (x_1 - x_2) \cdot (x_5 - x_6), \\ f_{5\leftarrow 6} &= \prod_{1 \leq i < j \leq 4} (x_i - x_j) \cdot \prod_{1 \leq i \leq 4} (x_i - x_5)^2. \end{aligned}$$

Figure 1 presents the Hasse diagram of the principal ideal generated by the equivalence class of f in the simple minor poset. The label of each edge $g \prec h$ is the number $\operatorname{ess} h - \operatorname{ess} g$. We use the following notation for simple minors of f:

$$\begin{aligned} q_1 &= \prod_{1 \leq i < j \leq 4} (x_i - x_j) \cdot \prod_{1 \leq i \leq 4} (x_i - x_5)^2, & q_2 &= (x_1 - x_2) \cdot (x_3 - x_4), \\ q_3 &= (x_1 - x_2) \cdot (x_1 - x_3), & q_4 &= (x_1 - x_2) \cdot (x_2 - x_3), \\ q_5 &= (x_1 - x_2)^2, & q_6 &= -(x_1 - x_2)^2. \end{aligned}$$

Remark 3. Even though not every identification minor $f_{i\leftarrow j}$ is a lower cover of f, every lower cover of f is of the form $f_{i\leftarrow j}$. Therefore, we can alternatively define the arity gap of f as

$$\operatorname{gap} f := \min_{g \prec f} (\operatorname{ess} f - \operatorname{ess} g).$$

We say that f is *totally symmetric,* if for all permutations π of $[n]$ the identity $f(a_1, \ldots, a_n) = f(a_{\pi(1)}, \ldots, a_{\pi(n)})$ holds for all $a_1, \ldots, a_n \in A$. Observe that a totally symmetric function depends either on all of its variables or on none of them.

Fact 1. *If $f\colon A^n \to B$ is totally symmetric, then for all $i, j, i', j' \in [n]$ ($i \neq j$, $i' \neq j'$), $f_{i \leftarrow j} \equiv f_{i' \leftarrow j'}$. Therefore, if f is nonconstant, then for all distinct $i, j \in [n]$, $f_{i \leftarrow j}$ is, up to equivalence, the unique lower cover of f.*

Berman and Kisielewicz [1] also introduced the following analogue of oddsupp. Let $\operatorname{supp}\colon \bigcup_{n \geq 1} A^n \to \mathcal{P}(A)$ be the mapping defined by

$$\operatorname{supp}(a_1, \ldots, a_n) := \{a_1, \ldots, a_n\}.$$

A function $f\colon A^n \to B$ is *determined by* supp, if there exists a function $\varphi\colon \mathcal{P}(A) \to B$ such that $f = \varphi \circ \operatorname{supp}|_{A^n}$. As for oddsupp, every function determined by supp is totally symmetric, and such a function either depends on all of its variables or on none of them.

In [13] it was observed that there exist functions that are determined both by supp and oddsupp. For instance, every constant function and every unary function is determined by both supp and oddsupp. In fact, for each $2 \leq n \leq |A|$ we can construct functions $f\colon A^n \to B$ determined by both supp and oddsupp; see [13]. However, such nontrivial examples cannot be found when $n > |A|$.

Proposition 4 ([13])**.** *If $n > |A|$, then $f\colon A^n \to B$ is determined by both* supp *and* oddsupp *if and only if f is a constant function.*

Now it is natural to ask how the properties of being determined by supp and oddsupp are affected when taking strict minors. The answer is given by the following result.

Proposition 5 ([13])**.** *Let A and B be finite nonempty sets, let $k := |A|$, and let $f\colon A^n \to B$ be a nonconstant function.*

(1) *If f is determined by* oddsupp*, then the simple minors of f form a chain*

$$f = f_n \succ f_{n-2} \succ \cdots \succ f_{n-2t-2} \succ f_{n-2t}$$

of length t such that $\operatorname{ess} f_{n-2i} = n - 2i$ for all $i < t$; in this case, we either have $\operatorname{ess} f_{n-2t} = 1$ and $t = \frac{n-1}{2}$ or $\operatorname{ess} f_{n-2t} = 0$ and $\lceil \frac{n-k}{2} \rceil < t \leq \lfloor \frac{n}{2} \rfloor$.

(2) *If f is determined by* supp*, then the simple minors of f form a chain*

$$f = f_n \succ f_{n-1} \succ \cdots \succ f_{n-t+1} \succ f_{n-t}$$

of length t such that $\operatorname{ess} f_{n-i} = n - i$ for all $i < t$; in this case, we either have $\operatorname{ess} f_{n-t} = 1$ and $t = n - 1$, or $\operatorname{ess} f_{n-t} = 0$ and $n - k < t < n$.

As made apparent by Willard [30], the notion of arity gap is tightly related to determinability by supp and oddsupp. The following corollary is an immediate consequence of Proposition 5.

Corollary 3. *Let A and B be finite nonempty sets, let $k := |A|$, and let $f\colon A^n \to B$ be a nonconstant function.*

(1) *If f is determined by* oddsupp *with $n > k$, then* $\operatorname{gap} f = 2$.
(2) *If f is determined by* supp *with $n > k$, then* $\operatorname{gap} f = 1$.

We now recall a noteworthy result about the arity gap.

Lemma 4 (Willard [30])**.** *Let A and B be finite nonempty sets, and let $k := |A|$. Suppose that $f\colon A^n \to B$ depends on all of its variables.*

(1) *If $n > 2$, $\operatorname{gap} f = 1$, f is totally symmetric, and for any distinct $i, j \in [n]$, $f_{i \leftarrow j}$ is equivalent to a totally symmetric function, then f is determined by* supp.
(2) *If f is determined by* supp*, then $f_{i \leftarrow j}$ is equivalent to a function determined by* supp *for any distinct $i, j \in [n]$. Moreover, if $n > k$, then $f_{i \leftarrow j}$ is nonconstant.*
(3) *If $n \geq \max(k, 3) + 2$ and f is not totally symmetric, then there exist distinct $i, j \in [n]$ such that $f_{i \leftarrow j}$ depends on $n - 1$ variables and is not equivalent to a totally symmetric function.*

From Proposition 5, Theorem 4 and Lemma 4, we can obtain the following theorem whose importance is made apparent in the next subsection.

Theorem 15 ([13])**.** *Let A and B be finite nonempty sets, and let $k := |A|$. Suppose that $f \colon A^n \to B$ depends on all of its variables.*

(1) *If $n \geq \max(k, 3) + 1$ and $\operatorname{gap} f = 2$, then for all $g < f$ with $\operatorname{ess} g > k$, it holds that $\operatorname{gap} g = 2$.*
(2) *If $n \geq \max(k, 3) + 2$ and $\operatorname{gap} f = 1$, then there exists a $g \prec f$ such that $\operatorname{gap} g = 1$ and $\operatorname{ess} g = n - 1$.*

6.2. **Parametrized variants of arity gap.** So far we have considered the effect that the identification of *two* essential variables has on the essential arity of functions. Indeed, the arity gap of a function measures the minimum decrease in the essential arity when two essential variables are identified. We shall now discuss the minimum decrease in the essential arity when we identify an arbitrarily large number of essential variables. There are two ways of formalizing such a measure: by sequentially identifying pairs of variables or by simultaneously identifying "blocks" of variables. Despite being related, as we will see these two approaches are rather different.

6.2.1. *Sequential identification.* One approach to formalizing the minimum decrease in the essential arity when several essential variables are identified is to consider the following parametrized version of arity gap which measures the minimum decrease in the essential arity when we take $\ell \geq 0$ steps downwards in the simple minor quasiorder:

$$\operatorname{gap}(f, \ell) := \min_{g \in \downarrow^\ell f} (\operatorname{ess} f - \operatorname{ess} g), \tag{5}$$

where

$$\downarrow^\ell f := \{g \in \mathcal{F}_{AB} \mid \exists f_1, \dots, f_{\ell-1} \colon f \succ f_1 \succ \cdots \succ f_{\ell-1} \succ g\}.$$

Note that $\operatorname{gap}(f, \ell)$ is defined only if there exists a chain of length ℓ below f, and in this case $\ell \leq \operatorname{gap}(f, \ell) \leq \operatorname{ess} f$. In fact, it is not difficult to verify that:

$$\operatorname{gap}(f, \ell) = \min_{g \prec f} (\operatorname{ess} f - \operatorname{ess} g + \operatorname{gap}(g, \ell - 1)). \tag{6}$$

Thus for every function f, we have $\operatorname{gap}(f, 1) = \operatorname{gap} f$ and $\operatorname{gap}(f, 0) = 0$.

Remark 4. We saw in the previous subsection that taking a strict minor of a function f requires the identification of at least one pair of essential variables of f; otherwise, the minors we obtain are equivalent to f. This means that $\operatorname{gap}(f, \ell)$ can be computed by sequentially identifying a pair of essential variables ℓ times in all possible ways, starting from f, and then determining the sequence of ℓ identifications which results in the minimum loss of essential variables.

It is worth stressing the fact that the identification of variables is performed sequentially, and at each step only one pair of essential variables is identified; otherwise, ambiguities could occur since a priori we do not know which essential variables become inessential after a pair is identified.

We have also observed in Remark 3 that not every identification minor $f_{i \leftarrow j}$ is a lower cover of f and, by the alternative definition of arity gap, if $f_{i \leftarrow j}$ is not a lower cover of f, then $\operatorname{gap} f < \operatorname{ess} f - \operatorname{ess} f_{i \leftarrow j}$. Moreover, as observed in Remark 2 it can be the case that f has two lower covers f_1 and f_2 such that $\operatorname{ess} f_1 < \operatorname{ess} f_2$, and again we would conclude that $\operatorname{gap} f < \operatorname{ess} f - \operatorname{ess} f_1$. Hence, one might be led to thinking that in order to compute $\operatorname{gap}(f, \ell)$ it suffices to choose at each recursion step an identification which results in the minimum loss of essential variables. However, this is not true.

Example 8. Consider function f in Example 7. If we choose as our first identification the pair $\{5, 6\}$, then any other identification of essential variables results in the loss of all the remaining essential variables. In other words, any downward path in Figure 1 which starts from f and passes through q_1 has length 2, and along it we first lose 1 and then 5 essential variables. However, the downward paths that start from f and pass through q_2 have length 4, and along them we lose 2, 1, 1, and then 2 essential variables. This shows that, in order to compute $\operatorname{gap}(f, 1)$ as in (5), the minimum value is attained at the lower cover q_1, whereas, for $2 \leq \ell \leq 4$, we need to pass through q_2 for computing $\operatorname{gap}(f, \ell)$. Hence, $\operatorname{gap}(f, 0) = 0$, $\operatorname{gap}(f, 1) = 1$, $\operatorname{gap}(f, 2) = 3$, $\operatorname{gap}(f, 3) = 4$, and $\operatorname{gap}(f, 4) = 6$.

The following result can be obtained inductively by making use of equation (6) and Theorem 15. Essentially, it asserts that, if ℓ is not too large, then we can walk from f down ℓ steps in the simple minor quasiorder in such a way that in each step we lose only one essential variable (resp. only two essential variables) if $\operatorname{gap} f = 1$ (resp. $\operatorname{gap} f = 2$).

Theorem 16 ([13])**.** *Let A and B be finite nonempty sets, and let $k := |A|$. Let $f \colon A^n \to B$, $\operatorname{ess} f = n$.*

(1) *If $\operatorname{gap} f = 1$ and $1 \leq \ell \leq n - \max(k, 3)$, then $\operatorname{gap}(f, \ell) = \ell$.*
(2) *If $\operatorname{gap} f = 2$ and $1 \leq \ell \leq \lceil \frac{n-k}{2} \rceil$, then $\operatorname{gap}(f, \ell) = 2\ell$.*

Surprisingly, it turns out that for almost every integer sequence $0 = n_0 < n_1 < n_2 < \cdots < n_r \leq n$, we can construct a function $f \colon A^n \to B$ whose parametrized arity gap meets every member of the sequence. This statement is made precise in the following theorem.

Theorem 17 ([13])**.** *Let A be a finite set with k elements and let B be a set with at least two elements. Let $2 \leq n \leq k$, $1 \leq r \leq n - 1$, $0 = n_0 < n_1 < n_2 < \cdots < n_r \leq n$ such that $n - 1 \leq n_r \leq n$ and $n_{r-1} \neq n - 1$. Then there exists a function $f \colon A^n \to B$ such that $\operatorname{gap}(f, \ell) = n_\ell$ for every $0 \leq \ell \leq r$.*

This parametrized version of arity gap constitutes a tool for tackling yet another natural problem pertaining to the effect of variable identification on the number of essential variables of a function. Given a function $f \colon A^n \to B$ and an integer $p \geq 1$, what is the smallest number m such that any m successive identifications of essential variables result in the loss of at least p essential variables? Denoting this smallest number by $\operatorname{pag}(f, p)$, we can easily see that $\operatorname{pag}(f, p)$ is the smallest ℓ for which $\operatorname{gap}(f, \ell) \geq p$.

Example 9. Consider the 6-ary function f of Example 7. We can read off of Figure 1 that

$$\begin{aligned} &\operatorname{pag}(f,1) = 1, && \operatorname{pag}(f,2) = 2, && \operatorname{pag}(f,3) = 2, \\ &\operatorname{pag}(f,4) = 3, && \operatorname{pag}(f,5) = 4, && \operatorname{pag}(f,6) = 4. \end{aligned}$$

6.2.2. *Block identification.* Instead of taking successive identifications of pairs of variables, another approach to measuring the minimum decrease in the essential arity when several essential variables are identified is to consider simultaneous identifications of "blocks" of variables. To formalize this measure we follow Willard's [30] framework and view functions of several variables as maps $f \colon A^V \to B$, where $V \subseteq \{x_i : i \in \mathbb{N}\}$. The cardinality of V is called the *arity* of f. In this framework, a function $g \colon A^W \to B$ is a *simple minor* of $f \colon A^V \to B$, if there exists a map $\alpha \colon V \to W$ such that $g(\mathbf{a}) = f(\mathbf{a} \circ \alpha)$ for all $\mathbf{a} \in A^W$.

Let $\operatorname{Eq}(V)$ denote the set of all equivalence relations on V. Given an equivalence relation $\theta \in \operatorname{Eq}(V)$, denote the canonical surjection by $v_\theta \colon V \to V/\theta$. For a function $f \colon A^V \to B$, we define the function $f^\theta \colon A^{V/\theta} \to B$ by $f^\theta(\mathbf{a}) = f(\mathbf{a} \circ v_\theta)$, and we say that f^θ is obtained from f by *block identification of variables* through θ. We informally identify V/θ with any one of its distinct representatives so that f^θ becomes a simple minor of f, and every simple minor of f is then equivalent to f^ψ for some $\psi \in \operatorname{Eq}(V)$. The number of variables identified through θ is

$$e(\theta) := \sum_{X \in V/\theta} (|X| - 1) = |V| - |V/\theta|.$$

Assuming that f depends on all of its variables, i.e., $\operatorname{ess} f = |V|$, we have that $\operatorname{ess} f^\theta \leq |V/\theta| = |V| - e(\theta) = \operatorname{ess} f - e(\theta)$.

Now we can define the analogue of the parametrized arity gap for block identification of variables. For a function $f \colon A^V \to B$ with $\operatorname{ess} f = |V| = n$ and for an integer ℓ such that $0 \leq \ell \leq n - 1$, we define

$$\text{b-gap}(f, \ell) := \min_{\substack{\theta \in \operatorname{Eq}(V) \\ e(\theta) = \ell}} (\operatorname{ess} f - \operatorname{ess} f^\theta).$$

Note that $\text{b-gap}(f,0) = 0$ and $\text{b-gap}(f,1) = \operatorname{gap} f$ for every function f. It is also clear that $\ell \leq \text{b-gap}(f,\ell) \leq n$ for every $0 \leq \ell \leq n-1$, and $\text{b-gap}(f,\ell) \leq \operatorname{gap}(f,\ell)$ for every ℓ for which $\operatorname{gap}(f,\ell)$ is defined.

Let $H(f) := \{\operatorname{ess} f - \operatorname{ess} g : g \leq f\}$. It is clear that

$$\{\text{b-gap}(f,\ell) : 0 \leq \ell \leq n-1\} \subseteq H(f).$$

Proposition 6 ([13]). *Let $f\colon A^V \to B$ be a function such that $\operatorname{ess} f = |V| = n$. Then $\text{b-gap}(f,\ell) = \min\{m \in H(f) : m \geq \ell\}$, for all $0 \leq \ell \leq n-1$.*

Example 10. Consider the 6-ary function f of Example 7. We can read off of Figure 1 that $H(f) = \{0,1,2,3,4,6\}$ and

$$\begin{aligned} &\text{b-gap}(f,1) = 1, && \text{b-gap}(f,2) = 2, && \text{b-gap}(f,3) = 3, \\ &\text{b-gap}(f,4) = 4, && \text{b-gap}(f,5) = 6. \end{aligned}$$

Following the same steps as in the sequential approach, we ended up considering the following problem: Given a function $f\colon A^V \to B$ that depends on all of its variables and an integer $p \geq 1$, what is the smallest number m such that block identification of variables of f through every equivalence relation θ on V with $e(\theta) = m$ results in the loss of at least p essential variables? Let us denote this smallest number by $\text{b-pag}(f,p)$. It is again clear that $\text{b-pag}(f,p)$ is the smallest ℓ for which $\text{b-gap}(f,\ell) \geq p$. In other words, $\text{b-pag}(f,0) = 0$ and $\text{b-pag}(f,p) = \max\{m \in H(f) : m < p\} + 1$ for $1 \leq p \leq n$.

Example 11. Consider the 6-ary function f of Example 7. We can determine from the values of $\text{b-gap}(f,\ell)$ listed in Example 10, or we can easily read off of Figure 1 that

$$\begin{aligned} &\text{b-pag}(f,1) = 1, && \text{b-pag}(f,2) = 2, && \text{b-pag}(f,3) = 3, \\ &\text{b-pag}(f,4) = 4, && \text{b-pag}(f,5) = 5, && \text{b-pag}(f,6) = 5. \end{aligned}$$

7. Open problems and future work

Looking ahead to possible future research topics related to arity gap, we are naturally drawn to two rather distinct problems. On the one hand, the work mentioned in Section 3 and Subsection 5.2 inevitably brings up the question on whether a general classification of algebras can be attained in terms of decomposability as in the case of abelian groups or fields of prescribed characterictic. On the other hand, the work surveyed in Section 4 naturally asks for similar enumeration results for the parametrized versions of arity gap discussed in Subsection 6.2.

Acknowledgment

The first named author is supported by the internal research project F1R-MTH-PUL-12RDO2 of the University of Luxembourg. The third named author acknowledges that the present project is supported by the TÁMOP-4.2.1/B-09/1/KONV-2010-0005 program of National Development Agency of Hungary, by the Hungarian National Foundation for Scientific Research under grants no. K77409 and K83219, by the National Research Fund of Luxembourg, and cofunded under the Marie Curie Actions of the European Commission (FP7-COFUND).

References

[1] J. Berman and A. Kisielewicz, On the number of operations in a clone, *Proc. Amer. Math. Soc.* **122** (1994) 359–369.

[2] M. Bouaziz, M. Couceiro and M. Pouzet, Join-irreducible Boolean functions, *Order* **27** (2010) 261–282.

[3] K. N. Čimev, On some properties of functions, in: B. Csákány and I. Rosenberg (eds.) *Finite Algebra and Multiple-Valued Logic,* Abstracts of lectures of the colloquium on finite algebra and multiple-valued logic (Szeged, 1979), North-Holland, 1981, pp. 38–40.

[4] K. N. Čimev, *Separable Sets of Arguments of Functions,* Studies 180/1986, Computer and Automation Institute, Hungarian Academy of Sciences, Budapest, 1986.

[5] M. Couceiro, On the lattice of equational classes of Boolean functions and its closed intervals, *J. Mult.-Valued Logic Soft Comput.* **18** (2008) 81–104.

[6] M. Couceiro and E. Lehtonen, On the effect of variable identification on the essential arity of functions on finite sets, *Int. J. Found. Comput. Sci.* **18** (2007) 975–986.
[7] M. Couceiro and E. Lehtonen, Generalizations of Świerczkowski's lemma and the arity gap of finite functions, *Discrete Math.* **309** (2009) 5905–5912.
[8] M. Couceiro and E. Lehtonen, The arity gap of polynomial functions over bounded distributive lattices, *40th IEEE International Symposium on Multiple-Valued Logic (ISMVL 2010),* IEEE Computer Society, Los Alamitos, 2010, pp. 113–116.
[9] M. Couceiro, E. Lehtonen and T. Waldhauser, Decompositions of functions based on arity gap, *Discrete Math.* **312** (2012) 238–247.
[10] M. Couceiro, E. Lehtonen and T. Waldhauser, The arity gap of order-preserving functions and extensions of pseudo-Boolean functions, *Discrete Appl. Math.* **160** (2012) 383–390.
[11] M. Couceiro, E. Lehtonen and T. Waldhauser, On the arity gap of polynomial functions, arXiv:1104.0595v2.
[12] M. Couceiro, E. Lehtonen and T. Waldhauser, Additive decomposability of functions over abelian groups, *Internat. J. Algebra Comput.* **23** (2013) 643–662.
[13] M. Couceiro, E. Lehtonen and T. Waldhauser, Parametrized arity gap, *Order* **30** (2013) 557–572.
[14] M. Couceiro and M. Pouzet, On a quasi-ordering on Boolean functions, *Theoret. Comput. Sci.* **396** (2008) 71–87.
[15] R. O. Davies, Two theorems on essential variables, *J. London Math. Soc.* **41** (1966) 333–335.
[16] K. Denecke and J. Koppitz, Essential variables in hypersubstitutions, *Algebra Universalis* **46** (2001) 443–454.
[17] O. Ekin, S. Foldes, P. L. Hammer and L. Hellerstein, Equational characterizations of Boolean function classes, *Discrete Math.* **211** (2000) 27–51.
[18] A. Feigelson and L. Hellerstein, The forbidden projections of unate functions, *Discrete Appl. Math.* **77** (1997) 221–236.
[19] E. Lehtonen, Descending chains and antichains of the unary, linear, and monotone subfunction relations, *Order* **23** (2006) 129–142.
[20] E. Lehtonen and Á. Szendrei, Equivalence of operations with respect to discriminator clones, *Discrete Math.* **309** (2009) 673–685.
[21] E. Lehtonen and Á. Szendrei, Partial orders induced by quasilinear clones, *Contributions to General Algebra* **20**, Proceedings of the Salzburg Conference 2011 (AAA81), Verlag Johannes Heyn, Klagenfurt, 2012, pp. 51–84,
[22] O. B. Lupanov, On a class of schemes of functional elements, *Problemy Kibernetiki* (1962) 61–114 (in Russian).
[23] N. Pippenger, Galois theory for minors of finite functions, *Discrete Math.* **254** (2002) 405–419.
[24] A. Salomaa, On essential variables of functions, especially in the algebra of logic, *Ann. Acad. Sci. Fenn. Ser. A I. Math.* **339** (1963) 3–11.
[25] S. Shtrakov, Essential arity gap of Boolean functions, *Serdica J. Computing* **2** (2008) 249–266.
[26] S. Shtrakov and J. Koppitz, On finite functions with non-trivial arity gap, *Discuss. Math. Gen. Algebra Appl.* **30** (2010) 217–245.
[27] N. A. Solovjev, On the question of the essential dependence of functions of the algebra of logic, *Problemy Kibernetiki* **9** (1963) 333–335 (in Russian).
[28] C. Wang, Boolean minors, *Discrete Math.* **141** (1991) 237–258.
[29] W. Wernick, An enumeration of logical functions, *Bull. Amer. Math. Soc.* **45** (1939) 885–887.
[30] R. Willard, Essential arities of term operations in finite algebras, *Discrete Math.* **149** (1996) 239–259.
[31] I. E. Zverovich, Characterizations of closed classes of Boolean functions in terms of forbidden subfunctions and Post classes, *Discrete Appl. Math.* **149** (2005) 200–218.

(M. Couceiro) Mathematics Research Unit, University of Luxembourg, 6, rue Richard Coudenhove-Kalergi, L-1359 Luxembourg, Luxembourg and LAMSADE – CNRS, Université Paris-Dauphine, Place du Maréchal de Lattre de Tassigny, 75775 Paris Cedex 16, France
E-mail address: miguel.couceiro@dauphine.fr

(E. Lehtonen) Computer Science and Communications Research Unit, University of Luxembourg, 6, rue Richard Coudenhove-Kalergi, L-1359 Luxembourg, Luxembourg and Centro de Álgebra da Universidade de Lisboa, Avenida Professor Gama Pinto 2, 1649-003 Lisboa, Portugal
E-mail address: erkko@campus.ul.pt

(T. Waldhauser) Mathematics Research Unit, University of Luxembourg, 6, rue Richard Coudenhove-Kalergi, L-1359 Luxembourg, Luxembourg and Bolyai Institute, University of Szeged, Aradi vértanúk tere 1, H-6720 Szeged, Hungary
E-mail address: twaldha@math.u-szeged.hu