Maximal Number of Vertices of Polytopes Defined by F-Probabilities

Anton Wallner

Ludwig-Maximilians-University Munich, Germany toni@stat.uni-muenchen.de

Abstract

Every F-probability (= coherent probability) \mathcal{F} on a finite sample space Ω_k with k elements defines a set of classical probabilities in accordance with the interval limits. This set, called "structure" of \mathcal{F} , is a convex polytope having dimension $\leq k-1$. We prove that the maximal number of vertices of structures is exactly k!.

Keywords. Geometry of interval probability, number of vertices of structures/cores/credal sets, combinatorial theory of polyhedra, 0/1-matrices.

1 Introduction

There is a famous subclass of F-probabilities = coherent probabilities (terminologies of Weichselberger [10] and Walley [8] respectively) on finite sample spaces, for which the computing and counting of the vertices of the corresponding structures are algorithmically very easy: the class of 2-monotone capacities (= C-probabilities in Weichselberger's terminology). It is well-known that the structures of 2-monotone capacities on the sample space Ω_k with k elements have at most k! vertices and that this bound is sharp.

Inspired by this and by the visible barycentric representation of structures for k = 1, 2, 3, 4 (where the classes of C- and F-probabilities coincide for k = 1, 2, 3) Weichselberger expressed the following conjecture in the early 1990s, i.e., at the beginning of his theory of *interval probability*:

Weichselberger conjecture (WEC). The structures of all F-probabilities on Ω_k have at most k! vertices.

For more than one decade many efforts were made to find a counterexample or to verify Weichselberger's conjecture.

We here prove him right.

For everyone who is a little bit familiar with coherent (= F-) probabilities and the description of vertices of polyhedra in general, Sections 2, 3, and 4 are mainly boring as regards content, but not as regards the special nomenclature used here for the sake of convenience.

Section 2 formally defines F-probability, C-probability and structure, and Section 3 is concerned with basic facts about polyhedra and their vertices. Section 4 embeds structures of F-probabilities on Ω_k isomorphically into the space \mathbb{R}^k , for convincing that they "are in fact polytopes" which are, moreover, defined by 0/1-matrices, and for preparing that apart from this the F-property of F-probabilities is not needed to obtain the result.

The core, and in fact new part, is Section 5, where the (WEC) is proven.

Section 6 is reserved for concluding remarks.

2 F-probabilities, C-probabilities, and their Structures

Here we report very briefly the main concepts of Weichselberger's theory of interval probability (see, e.g., [10]), in particular the concepts of F- and C-probability and, indispensably, of their structures.

Just as in classical probability theory, the starting point is some fixed measurable space $(\Omega; A)$, where Ω is a non-empty sample space and A is a σ -algebra over Ω . Here we only want to deal with the finite case. Hence we assume that Ω is finite, w.l.o.g. $\Omega = \Omega_k := \{1, \ldots, k\}$ with $k \geq 1$, and $A = \mathcal{P}(\Omega_k)$, i.e., the power set of Ω_k . The members A of $\mathcal{P}(\Omega_k)$ are called events, and, in particular, the singletons $E_i := \{i\}$, $i = 1, \ldots, k$, are called atoms.

2.1. Definition.

(a) A set function $p: \mathcal{P}(\Omega_k) \to [0; 1]$ is called a K-function (or classical probability) on $(\Omega_k; \mathcal{P}(\Omega_k))$,

if it satisfies Kolmogorov's axioms, i.e.,

- (i) $p(A) \geq 0, \forall A \subseteq \Omega_k$;
- (ii) $p(\Omega_k) = 1$ (norm condition);
- (iii) p is additive, i.e., $p(A \cup B) = p(A) + p(B)$, $\forall A, B \subseteq \Omega_k$ with $A \cap B = \emptyset$.

The set of all K-functions on $(\Omega_k; \mathcal{P}(\Omega_k))$ is denoted by \mathcal{K}_k .

(b) A quadruple $\mathcal{F} = (\Omega_k; \mathcal{P}(\Omega_k); L, U)$ is called an F-(probability) field on $(\Omega_k; \mathcal{P}(\Omega_k))$, if L, U: $\mathcal{P}(\Omega_k) \to [0; 1]$ are set functions such that

$$\mathcal{M}(\mathcal{F}) := \{ p \in \mathcal{K}_k | L(A) \le p(A) \le U(A), \ \forall A \subseteq \Omega_k \}$$

is not empty and $L(A) = \min_{p \in \mathcal{M}(\mathcal{F})} p(A),$ $U(A) = \max_{p \in \mathcal{M}(\mathcal{F})} p(A), \forall A \subseteq \Omega_k.$

(c) The set $\mathcal{M}(\mathcal{F})$ according to (b) is named the structure of \mathcal{F} .

Sometimes structures are also called *cores* (see, e.g., [6]) or *credal sets* (see, e.g., [4]).

There are four basic properties of F-fields:

- **2.2. Corollary.** Let $\mathcal{F} = (\Omega_k; \mathcal{P}(\Omega_k); L, U)$ be an *F-field. Then* we have:
- (a) $L(.) \ge 0$.
- (b) $L(\emptyset) = 0$.
- (c) $L(\Omega_k) = 1$ (i.e., L is normed).

(d)
$$U(A) = 1 - L(\Omega_k \backslash A), \forall A \subseteq \Omega_k$$
.

Proof. (a), (b) and (c): straightforward. For (d) use
$$p(A) = 1 - p(\Omega_k \setminus A), \forall A \subseteq \Omega_k, \forall p \in \mathcal{K}_k.$$

It is possible to describe the structure of an F-field only by its lower limit L:

2.3. Corollary. Let $\mathcal{F} = (\Omega_k; \mathcal{P}(\Omega_k); L, U)$ be an *F-field. Then*

$$\mathcal{M}(\mathcal{F}) = \{ p \in \mathcal{K}_k | \ p(A) \ge L(A), \ \forall A \subseteq \Omega_k \}$$
 (1)

already determines the structure of \mathcal{F} .

Proof. Straightforwardly with Corollary 2.2(d).

As a special subclass of F-probabilities we make a note of C-probabilities = 2-monotone capacities:

- **2.4. Definition.** A quadruple $C = (\Omega_k; \mathcal{P}(\Omega_k); L, U)$ is called a C-(probability) field on $(\Omega_k; \mathcal{P}(\Omega_k))$, if the following conditions are satisfied:
- (a) $L(.) \geq 0$.
- (b) $L(\emptyset) = 0$.
- (c) $L(\Omega_k) = 1$.
- (d) $U(A) = 1 L(\Omega_k \setminus A), \forall A \subseteq \Omega_k$.
- (e) L is 2-monotone, i.e., $L(A) + L(B) \le L(A \cap B) + L(A \cup B)$, $\forall A, B \subseteq \Omega_k$.

2.5. Lemma. Every C-field on $(\Omega_k; \mathcal{P}(\Omega_k))$ is an F-field on $(\Omega_k; \mathcal{P}(\Omega_k))$.

Proof. E.g., see [3].

The consideration of F- and C-fields, in particular the geometrical properties of their structures, will be continued in Section 4 — after fixing some basic facts about polyhedra and their vertices.

3 Basic Facts about Polyhedra and their Vertices; the psi-Function

In this section we give the usual definitions of polyhedra, polytopes, and their vertices and, moreover, introduce some notational conventions.

3.1. Conventions.

- (a) For every finite set X let |X| be the cardinality of X.
- (b) The components of $x \in \mathbb{R}^k$ are denoted by $x(1), \ldots, x(k)$.
- (c) For $x, y \in \mathbb{R}^k$ let $\langle x, y \rangle$ be the usual scalar product of x and y, i.e., $\langle x, y \rangle := \sum_{i=1}^k x(i)y(i)$. \square
- **3.2. Definition.** Let $X \subseteq \mathbb{R}^k$.
- (a) X is called a *polyhedron*, if X is a finite intersection of closed affine halfspaces of \mathbb{R}^k , i.e., if there exist $n \in \mathbb{N}$, $a_1, \ldots, a_n \in \mathbb{R}^k$, and $\beta_1, \ldots, \beta_n \in \mathbb{R}$ such that

$$X = \bigcap_{i=1}^{n} \{ x \in \mathbb{R}^k | \langle a_i, x \rangle \ge \beta_i \}.$$
 (2)

- (b) A polyhedron X is called a *polytope*, if X is bounded.
- (c) A point x of a polyhedron X is called a vertex of X, if $\forall y, z \in X \ \forall \lambda \in]0$; 1[$(x = (1-\lambda)y + \lambda z \Longrightarrow x = y = z)$. The set of all vertices of X is denoted by $\mathcal{E}(X)$.

There is a well-known — but generally very ineffective — procedure for finding all the vertices of some given polyhedron. We describe it in the following lemma, point (a).

- **3.3. Lemma.** Let $X \subseteq \mathbb{R}^k$ be a polyhedron of the form (2).
- (a) If $x \in X$, then x is a vertex of X iff there are indices $i_1, \ldots, i_k \in \{1, \ldots, n\}$ such that $\{a_{i_1}, \ldots, a_{i_k}\}$ is a basis of \mathbb{R}^k and $\langle a_{i_j}, x \rangle = \beta_{i_j}$, $\forall j = 1, \ldots, k$.

(b)
$$|\mathcal{E}(X)| \leq \binom{n}{k}$$
.

Proof. For (a) see, e.g., [5], 7.2(b), p. 122; (b) follows from (a). \Box

The estimation in Lemma 3.3(b) has the big disadvantage that it is very weak in general, but it has the advantage that it does not depend on a_1, \ldots, a_n and β_1, \ldots, β_n . Subsequently we steer a middle course: We fix $a_1, \ldots, a_n \in \mathbb{R}^k$ and ask for the maximal number of vertices the corresponding polyhedra can have, if we let β_1, \ldots, β_n vary over all (!) real numbers, i.e., if we allow all parallel translations of the defining hyperplanes.

For approaching this problem in a terminologically convenient manner, we now take a closer look at the definition (2) of a polyhedron X, which will lead to specific notation.

First, it is usual to employ the matrix-vector-style for an equivalent description of the polyhedron X: X is the solution space of the system

$$A \cdot x \ge b, \tag{3}$$

where
$$A:=\begin{pmatrix}a_1\\\vdots\\a_n\end{pmatrix}$$
 and $b:=\begin{pmatrix}\beta_1\\\vdots\\\beta_n\end{pmatrix}$, if we consider

 a_1, \ldots, a_n as row vectors and x as a column vector in \mathbb{R}^k . But subsequently we do not need the description (3). Instead it will be more comfortable to use a functional style which is, moreover, close to the "language" of interval probability: Since w.l.o.g. we can assume that a_1, \ldots, a_n are different and since the order of a_1, \ldots, a_n is irrelevant, the relation $a_i \mapsto \beta_i$, $i = 1, \ldots, n$, is a function, and it is possible to represent X equivalently by

$$\{x \in \mathbb{R}^k | \langle a, x \rangle \ge L(a), \ \forall a \in \mathcal{A}_> \},$$
 (4)

if we put $A_{\geq} := \{a_1, \ldots, a_n\}$ and $L: A_{\geq} \to \mathbb{R}$, $L(a_i) := \beta_i, i = 1, \ldots, n$ (L stands for lower bound).

Secondly, of course it is feasible to incorporate into (4) linear equations $\langle a, x \rangle = L(a)$ by letting $a, -a \in \mathcal{A}_{\geq}$ and writing $\langle a, x \rangle \geq L'(a) \wedge \langle -a, x \rangle \geq L'(-a)$, where L'(a) := L(a) and L'(-a) := -L(a) (w.l.o.g. L(a) = 0, if a = 0, because we can presume that X is not empty). But, in order to avoid boring transformations when dealing with linear equations, we replace (4) by

$$\{x \in \mathbb{R}^k | \langle a, x \rangle \ge L(a), \ \forall a \in \mathcal{A}_\ge \}$$
$$\cap \{x \in \mathbb{R}^k | \langle a, x \rangle = L(a), \ \forall a \in \mathcal{A}_= \}$$

for an appropriate finite set $\mathcal{A}_{=} \subseteq \mathbb{R}^{k}$. Moreover, it is handy to assume $\mathcal{A}_{=} \subseteq \mathcal{A}_{\geq}$.

We summarize:

3.4. Definition. Let $A_{\geq} \subseteq \mathbb{R}^k$ be finite, $L: A_{\geq} \to \mathbb{R}$, and $A_{=} \subseteq A_{>}$. Then we define:

- (a) $\mathcal{M}(L) := \{ x \in \mathbb{R}^k | \langle a, x \rangle \ge L(a), \forall a \in \mathcal{A}_{>} \}.$
- (b) $\mathcal{M}(L, \mathcal{A}_{=}) := \mathcal{M}(L) \cap \{x \in \mathbb{R}^{k} | \langle a, x \rangle = L(a), \forall a \in \mathcal{A}_{=}\}. \square$

3.5. Corollary. Let $A_{\geq} \subseteq \mathbb{R}^k$ be finite and $L: A_{\geq} \to \mathbb{R}$. Then we have:

(a) $\mathcal{M}(L) = \mathcal{M}(L, \emptyset)$.

(b)
$$\mathcal{B}_{=} \subseteq \mathcal{A}_{=} \subseteq \mathcal{A}_{\geq} \Longrightarrow$$

 $\mathcal{E}(\mathcal{M}(L, \mathcal{A}_{=})) \subseteq \mathcal{E}(\mathcal{M}(L, \mathcal{B}_{=})).$

Proof. (a) is trivial, and (b) follows from the fact that $\mathcal{M}(L, \mathcal{A}_{=})$ is a *face* of $\mathcal{M}(L, \mathcal{B}_{=})$, if $\mathcal{B}_{=} \subseteq \mathcal{A}_{=}$ (see, e.g., [5], p. 123ff).

3.6. Definition. We define

$$BAS(A_{>}) := \{ \mathcal{B} \subseteq A_{>} | \mathcal{B} \text{ is a basis of } \mathbb{R}^{k} \}$$

for any
$$A > \subseteq \mathbb{R}^k$$
.

Using this abbreviation and the notations introduced in Definition 3.4 we can rewrite Lemma 3.3(a):

3.7. Corollary. Let $A_{\geq} \subseteq \mathbb{R}^k$ be finite, $L: A_{\geq} \to \mathbb{R}$, and $A_{=} \subseteq A_{>}$. Then we have:

(a)
$$\mathcal{E}(\mathcal{M}(L, \mathcal{A}_{=})) = \bigcup \{ \mathcal{M}(L, \mathcal{A}_{=} \cup \mathcal{B}) | \mathcal{B} \in BAS(\mathcal{A}_{>}) \}.$$

(b)
$$\mathcal{E}(\mathcal{M}(L)) = \bigcup \{ \mathcal{M}(L, \mathcal{B}) | \mathcal{B} \in BAS(\mathcal{A}_{>}) \}.$$

Proof. (a) can be deduced straightforwardly from Lemma 3.3(a), and (b) follows from (a) and Corollary 3.5(a).

Note that in Corollary 3.7 all the sets $\mathcal{M}(L, \mathcal{A}_{=} \cup \mathcal{B})$ and $\mathcal{M}(L, \mathcal{B}), \mathcal{B} \in \text{BAS}(\mathcal{A}_{\geq})$, are empty or singletons (cf. Proposition 5.2).

Now we formally define the maximal number of vertices the polyhedra $\mathcal{M}(L, \mathcal{A}_{=})$ can have, letting L vary over all real-valued functions with domain $\mathcal{A}_{>}$:

3.8. Definition. Let $A_{\geq} \subseteq \mathbb{R}^k$ be finite and $A_{=} \subseteq A_{>}$. Then we define:

(a)
$$\psi(\mathcal{A}_{\geq}, \mathcal{A}_{=}) := \max\{|\mathcal{E}(\mathcal{M}(L, \mathcal{A}_{=}))| \mid L \colon \mathcal{A}_{>} \to \mathbb{R}\}.$$

(b)
$$\psi(\mathcal{A}_{\geq}) := \psi(\mathcal{A}_{\geq}, \emptyset) = \max\{|\mathcal{E}(\mathcal{M}(L))| \mid L \colon \mathcal{A}_{\geq} \to \mathbb{R}\}.$$

We refer to $\psi(A_{\geq}, A_{=})$ and $\psi(A_{\geq})$ as the *psi-function* of $(A_{>}, A_{=})$ or $A_{>}$ respectively.

Therefore, both versions of the psi-function (itself) only depend on the dimension k of \mathbb{R}^k .

3.9. Corollary. Let $A_{\geq} \subseteq \mathbb{R}^k$ be finite. Then we have:

(a)
$$\mathcal{B}_{=} \subseteq \mathcal{A}_{=} \subseteq \mathcal{A}_{\geq} \Longrightarrow \psi(\mathcal{A}_{\geq}, \mathcal{A}_{=}) \leq \psi(\mathcal{A}_{\geq}, \mathcal{B}_{=}).$$

(b)
$$A_{=} \subseteq A_{>} \implies \psi(A_{>}, A_{=}) \le \psi(A_{>}).$$

Proof. (a) is a consequence of Corollary 3.5(b), and (a) implies (b). \Box

4 Making the Goal Smooth

In this section we will see that structures of F-probabilities on $(\Omega_k; \mathcal{P}(\Omega_k))$ are "in fact" polytopes in the space \mathbb{R}^k . Moreover, we will establish a relationship between their maximal vertices and both versions of the psi-function introduced in Section 3, which makes the goal, i.e., (WEC), "smooth" for its proof.

First, we interpret the elements p of \mathcal{K}_k as vectors in \mathbb{R}^k with components $p(E_1) = p(\{1\}), \ldots, p(E_k) = p(\{k\})$, and for brevity we write $p(1), \ldots, p(k)$. All the values p(A), $A \subseteq \Omega_k$, are determined by these components due to the axiom of additivity:

$$p(A) = \sum_{i=1}^{k} 1_A(i)p(i), \quad \forall A \subseteq \Omega_k,$$

where 1_A is the indicator function of A, i.e.,

$$1_A(i) := \begin{cases} 1, & \text{if } i \in A \\ 0, & \text{else.} \end{cases}$$

Indicator functions 1_A are also vectors in \mathbb{R}^k , only having components 0 or 1. Such vectors are called $\theta/1$ -vectors. To simplify notation, we subsequently write A instead of 1_A . In particular, the empty event \emptyset corresponds to $0 \in \mathbb{R}^k$, and Ω_k corresponds to $(1,\ldots,1) \in \mathbb{R}^k$. Analogously, $\mathcal{P}(\Omega_k)$ can be seen as the set of all 0/1-vectors in \mathbb{R}^k , and the atoms E_1,\ldots,E_k form the standard basis of \mathbb{R}^k . Note that E_1,\ldots,E_k are also K-functions (Dirac measures), and the (k-1)-dimensional simplex spanned by those is exactly the set \mathcal{K}_k :

$$\mathcal{K}_{k} = \{ p \in \mathbb{R}^{k} | \langle \Omega_{k}, p \rangle = 1 \land \langle E_{i}, p \rangle \geq 0, \ \forall i = 1, \dots, k \}.$$
 (5)

Now let $\mathcal{F} = (\Omega_k; \mathcal{P}(\Omega_k); L, U)$ be an F-field. If we rewrite equation (1) of its structure by

$$\mathcal{M}(\mathcal{F}) = \{ p \in \mathbb{R}^k | \langle A, p \rangle \ge L(A), \, \forall A \in \mathcal{P}(\Omega_k) \} \cap \mathcal{K}_k$$

and take into account that L(.) is non-negative and that $L(\Omega_k) = 1$ according to Corollary 2.2, (a) and (c), we infer from (5)

$$\mathcal{M}(\mathcal{F}) = \{ x \in \mathbb{R}^k | \langle A, x \rangle \ge L(A), \ \forall A \in \mathcal{P}(\Omega_k) \}$$
$$\cap \{ x \in \mathbb{R}^k | \langle \Omega_k, x \rangle = L(\Omega_k) \}.$$

Using the terminology introduced in Definition 3.4(b) we conclude:

$$\mathcal{M}(\mathcal{F}) = \mathcal{M}(L, \{\Omega_k\}). \tag{6}$$

In particular, $\mathcal{M}(\mathcal{F})$ is a polyhedron in the space \mathbb{R}^k , and, since it is bounded by the simplex \mathcal{K}_k , $\mathcal{M}(\mathcal{F})$ is a polytope with dimension $\leq k-1$. For $k \leq 4$ it is possible to make structures visible using barycentric coordinates: cf. Figures 1–4.

If we would prefer a matrix-vector-style for the description of $\mathcal{M}(\mathcal{F})$, we would have, let's say for k=4, the following system of inequalities:

supplemented by the norm condition

$$\left(\begin{array}{cccc} 1 & 1 & 1 & 1 \end{array} \right) \cdot \left(\begin{array}{c} x(1) \\ x(2) \\ x(3) \\ x(4) \end{array} \right) = L((1,1,1,1)). (7)$$

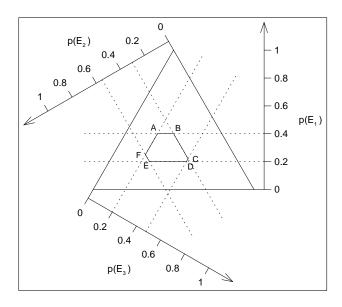


Figure 1: An F-field $\mathcal{F}=(\Omega_3;\mathcal{P}(\Omega_3);L,U)$ using the barycentric representation. The probability intervals P(.)=[L(.);U(.)] are given by

$$\begin{array}{ll} P(\emptyset) = [0;0] & P(E_1 \cup E_2) = [0.52; \, 0.80] \\ P(E_1) = [0.20; \, 0.40] & P(E_1 \cup E_3) = [0.45; \, 0.70] \\ P(E_2) = [0.30; \, 0.55] & P(E_2 \cup E_3) = [0.60; \, 0.80] \\ P(E_3) = [0.20; \, 0.48] & P(\Omega_3) = [1; \, 1]. \end{array}$$

Since k=3, \mathcal{F} is a C-field. Hence it is possible to describe the 3!=6 structure vertices A, B, C, D, E, and F by the characteristic equations (9). For example, consider $p_D=(0.20, 0.32, 0.48) \in \mathcal{K}_3$ given by the point D: We have $p_D(E_1)=0.20=L(E_1), p_D(E_1\cup E_2)=0.52=L(E_1\cup E_2),$ and $p_D(\Omega_3)=1=L(\Omega_3).$

Now let

$$M_k^F$$

be the maximal number of vertices of structures of Ffields on $(\Omega_k; \mathcal{P}(\Omega_k))$. Our goal is to prove

$$M_k^F = k!$$

First we say a word to the statement $M_k^F \geq k!$.

For this, let M_k^C be the maximal number of vertices of structures of C-fields on $(\Omega_k; \mathcal{P}(\Omega_k))$. By Lemma 2.5 we trivially get $M_k^F \geq M_k^C$. But for C-fields the maximal number of vertices, and moreover the description of the vertices itself, is well-known and easy (e.g., see [6], Theorem 3, p. 19, or [1], Proposition 13, p. 277):

Let $C = (\Omega_k; \mathcal{P}(\Omega_k); L, U)$ be a C-field. Then, characteristically, we have $p \in \mathcal{E}(\mathcal{M}(C))$ iff there exists a permutation π of $\{1, \ldots, k\}$ such that $p = p_{\pi}$, where p_{π} is defined by

$$p_{\pi}(E_{\pi(i)}) = L\left(\bigcup_{j=1}^{i} E_{\pi(j)}\right) - L\left(\bigcup_{j=1}^{i-1} E_{\pi(j)}\right)$$
 (8)

for i = 1, ..., k, or — equivalently — by

$$p_{\pi}\left(\bigcup_{j=1}^{i} E_{\pi(j)}\right) = L\left(\bigcup_{j=1}^{i} E_{\pi(j)}\right)$$
(9)

for i = 1, ..., k (cf. Fig. 1 and 2). Hence, on the one

¹This correspondence between vertices and permutations is false for every F-field which is no C-field (cf. Fig. 3 and 4).

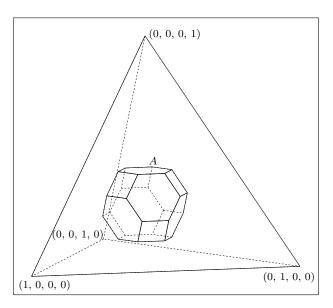


Figure 2: An F-field $\mathcal{F}=(\Omega_4;\,\mathcal{P}(\Omega_4);\,L,\,U)$ using the barycentric representation. \mathcal{F} is a C-field. E.g., we verify (9) and (8) by considering the structure vertex p_A represented by the point A: p_A is determined by $p_A(E_1)=L(E_1),\,p_A(E_1\cup E_2)=L(E_1\cup E_2),\,p_A(E_1\cup E_2\cup E_3)=L(E_1\cup E_2\cup E_3),\,$ and $p_A(\Omega_4)=1=L(\Omega_4)$. Therefore $p_A=(p_A(E_1),\,p_A(E_2),\,p_A(E_3),\,p_A(E_4))=(L(E_1),\,L(E_1\cup E_2)-L(E_1),\,L(E_1\cup E_2\cup E_3)-L(E_1\cup E_2),\,$ $1-L(E_1\cup E_2\cup E_3))$. Here all the 4!=24 structure vertices of $\mathcal F$ can be obtained from p_A by permuting its coordinates. Hence $\mathcal F$ is a uniform C-field (cf. (10)).

hand, we get $M_k^C \leq k!$, but, on the other hand, there are C-fields which have exactly k! structure vertices.

For example, consider a strict convex function f: $[0; 1] \rightarrow [0; 1]$ with f(0) = 0 and f(1) = 1, e.g., take $f(x) := x^2$, define L, U: $\mathcal{P}(\Omega_k) \rightarrow [0; 1]$ by $L(A) := f\left(\frac{|A|}{k}\right)$ and $U(A) := 1 - L(\Omega_k \setminus A)$, $\forall A \subseteq \Omega_k$, and let $\mathcal{C} := (\Omega_k; \mathcal{P}(\Omega_k); L, U)$. Then it is easy to see that L is 2-monotone, and thus \mathcal{C} is a C-field which, in addition, is uniform (cf. Fig. 2), i.e.,

$$L(A) = L(B), \ \forall A, B \in \mathcal{P}(\Omega_k) \text{ with } |A| = |B|. \ (10)$$

It is a straightforward exercise to show that for \mathcal{C} defined like this the p_{π} 's according to (8) are pairwise different. Hence $|\mathcal{E}(\mathcal{M}(\mathcal{C}))| = k!$ and thus $M_k^C = k!$.

We summarize:

4.1. Lemma.
$$M_k^F \geq k!$$
 for all $k \geq 1$.

Now we turn to the Weichselberger conjecture, i.e., to the statement

(WEC) $M_k^F \leq k!$ for all $k \geq 1$.

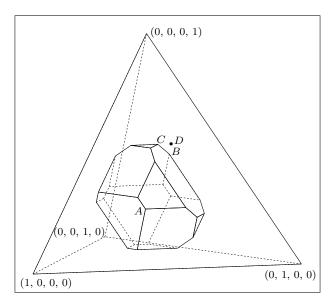


Figure 3: An F-field $\mathcal{F} = (\Omega_4; \mathcal{P}(\Omega_4); L, U)$ using the barycentric representation. \mathcal{F} has 4! = 24 structure vertices, and as examples we consider the vertices p_A , p_B , and p_C corresponding to the points A, B, and C respectively:

- p_A is defined by $p_A(E_3) = L(E_3), p_A(E_3 \cup E_4) = L(E_3 \cup E_4), p_A(E_3 \cup E_4 \cup E_2) = L(E_3 \cup E_4 \cup E_2),$
- p_B is defined by $p_B(E_1) = L(E_1), p_B(E_1 \cup E_2) = L(E_1 \cup E_2),$ $p_B(E_1 \cup E_3) = L(E_1 \cup E_3),$
- p_C is defined by $p_C(E_1 \cup E_2) = L(E_1 \cup E_2), p_C(E_1 \cup E_3) = L(E_1 \cup E_3), p_C(E_1 \cup E_2 \cup E_3) = L(E_1 \cup E_2 \cup E_3),$

supplemented by $p_A(\Omega_4) = p_B(\Omega_4) = p_C(\Omega_4) = 1 = L(\Omega_4)$. Since p_B and p_C do not respect the equations (9), $\mathcal F$ is no C-field. Additionally, consider the point D: The corresponding K-function p_D , defined by $p_D\left(\bigcup_{j=1}^i E_j\right) = L\left(\bigcup_{j=1}^i E_j\right)$, i=1,2,3,4, lies outside the structure of $\mathcal F$.

As mentioned in the Introduction, for k = 1, 2, 3 every F-field is a C-field such that $M_k^F \leq k!$ is true at least for these three values of k. Moreover, in [9] it is shown (with completely different methods) that, for any $k \geq 1$, the bound k! is valid for every *uniform* F-field (in the sense of (10), cf. Fig. 2 and 4).

Here we prove the entire (WEC). For preparing this, we propose two daring hypotheses. The *first one* is:

• "The conditions on the lower bound L of an F-field are, compared with our problem, extremely weak. So we forget them at all."

This means: If we look at (6), we ignore the fact that the L in the term $\mathcal{M}(L, \{\Omega_k\})$ stems from an F-field. We consider the polytopes $\mathcal{M}(L, \{\Omega_k\})$ for any map $L: \mathcal{P}(\Omega_k) \to \mathbb{R}$, hoping that the maximal number of their vertices do not increase in comparison with the smaller class of L's which are lower bounds of F-fields.

In other words: Recalling the definition of the psifunction in 3.8(a) we make the following conjecture:

(WAC1) $\psi(\mathcal{P}(\Omega_k), \{\Omega_k\}) \leq k!$ for all $k \geq 1$.

Hence (WAC1) implies (WEC).

The second hypothesis is:

• "The norm condition (like (7) for k=4) does not fit to the pattern of the remaining system of inequalities. So we forget it."

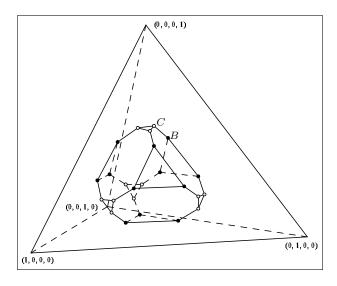


Figure 4: A uniform (cf. (10)) F-field $\mathcal{F} = (\Omega_4; \mathcal{P}(\Omega_4); L, U)$ using the barycentric representation. The 4! = 24 structure vertices of \mathcal{F} can be divided into two parts: $\mathcal{E}_{\bullet}(\mathcal{M}(\mathcal{F}))$ and $\mathcal{E}_{\mathbf{O}}(\mathcal{M}(\mathcal{F}))$ marked by \bullet and \bullet resp. Let $p_B, p_C \in \mathcal{K}_4$ correspond to the points B and C resp. Then p_B and p_C are defined like p_B and p_C in Figure 3. Hence \mathcal{F} is no C-field. By the way, each of the convex hulls, of $\mathcal{E}_{\bullet}(\mathcal{M}(\mathcal{F}))$ and of $\mathcal{E}_{\mathbf{O}}(\mathcal{M}(\mathcal{F}))$, are structures of uniform C-fields, each of them having "only a few" vertices (12), but together "spanning" the whole structure of \mathcal{F} (see [9], Section 4.2, for more details on this subject).

Geometrically this hypothesis says the following: Let, for fixed k, L: $\mathcal{P}(\Omega_k) \to \mathbb{R}$ be a function such that the corresponding polytope $\mathcal{M}(L, \{\Omega_k\})$ has the maximal number of vertices, i.e., $|\mathcal{E}(\mathcal{M}(L, \{\Omega_k\}))| = \psi(\mathcal{P}(\Omega_k), \{\Omega_k\})$. Then the (unbounded) polyhedron $\mathcal{M}(L)$ (without "norm condition") has no additional vertex, although $\mathcal{M}(L, \{\Omega_k\})$ is just a (lower dimensional) face of $\mathcal{M}(L)$ (cf. Corollary 3.5).

Again recalling the definition of the psi-function — now the second version in 3.8(b) —, the formal equivalent of the second hypothesis is the conjecture

(WAC2) $\psi(\mathcal{P}(\Omega_k)) \leq k!$ for all $k \geq 1$.

By Corollary 3.9(b) we have $\psi(\mathcal{P}(\Omega_k), \{\Omega_k\}) \leq \psi(\mathcal{P}(\Omega_k))$, and thus (WAC2) implies (WAC1).

In the next section we prove that indeed (WAC2) is true. But beforehand, we store for later reference:

4.2. Lemma. (WAC2) implies (WEC).
$$\Box$$

5 The Main Theorem

Now we concentrate on proving (WAC2). Let the natural number $k \geq 1$ be fixed. As already partially introduced in Section 4 we here officially (re-)define:

5.1. Definition.

- (a) $a \in \mathbb{R}^k$ is called a $\theta/1$ -vector, if $a(1), \ldots, a(k) \in \{0, 1\}$. The set of all 0/1-vectors in \mathbb{R}^k is denoted by $\mathcal{P}(\Omega_k)$.
- (b) Every element of BAS($\mathcal{P}(\Omega_k)$) (cf. Definition 3.6) is called a 0/1-basis.

Clearly we have to deal with 0/1-vectors and 0/1-bases for proving (WAC2). But as far as possible and no additional work is needed, we will extend our view to all vectors in \mathbb{R}^k and to all bases of \mathbb{R}^k .

We proceed in $four\ steps$, where our $first\ goal$ is Lemma 5.3.

5.2. Proposition. Let $\mathcal{B} = \{b_1, \ldots, b_k\}$ be a basis of \mathbb{R}^k and $x, y \in \mathbb{R}^k$. Then we have:

$$(\forall i = 1, \ldots, k. \langle b_i, x \rangle = \langle b_i, y \rangle) \implies x = y.$$

5.3. Lemma. Let $A_{\geq} \subseteq \mathbb{R}^k$ be finite and $A = \{a_1, \ldots, a_n\} \subseteq A_{\geq}$. Let $B = \{b_1, \ldots, b_k\} \subseteq A_{\geq}$ be a basis of \mathbb{R}^k such that there exist $\alpha_1, \ldots, \alpha_n \geq 0$ and $\beta_1, \ldots, \beta_k > 0$ with

$$\sum_{i=1}^{n} \alpha_i a_i = \sum_{i=1}^{k} \beta_i b_i. \tag{11}$$

Then for all functions $L: \mathcal{A}_{\geq} \to \mathbb{R}$ and all $x, y \in \mathbb{R}^k$ we have (cf. Definition 3.4(b)):

$$x \in \mathcal{M}(L, \mathcal{A}) \land y \in \mathcal{M}(L, \mathcal{B}) \implies x = y.$$

Proof. Let all the premises be given. Then $x \in \mathcal{M}(L, A)$ implies

$$\langle a_i, x \rangle = L(a_i), \quad \forall i = 1, \dots, n,$$
 (12)

$$\langle b_i, x \rangle \geq L(b_i), \quad \forall i = 1, \dots, k,$$
 (13)

whereas $y \in \mathcal{M}(L, \mathcal{B})$ implies

$$\langle b_i, y \rangle = L(b_i), \quad \forall i = 1, \dots, k, \qquad (14)$$

$$\langle a_i, y \rangle \ge L(a_i), \quad \forall i = 1, \dots, n.$$
 (15)

Therefore we compute:

$$\sum_{i=1}^{n} \alpha_{i} L(a_{i}) \stackrel{(12)}{=} \sum_{i=1}^{n} \alpha_{i} \langle a_{i}, x \rangle = \left\langle \sum_{i=1}^{n} \alpha_{i} a_{i}, x \right\rangle$$

$$\stackrel{(11)}{=} \left\langle \sum_{i=1}^{k} \beta_i b_i, x \right\rangle = \sum_{i=1}^{k} \beta_i \langle b_i, x \rangle \stackrel{(13)}{\geq} \sum_{i=1}^{k} \beta_i L(b_i)$$

$$\stackrel{(14)}{=} \sum_{i=1}^{k} \beta_i \langle b_i, y \rangle = \left\langle \sum_{i=1}^{k} \beta_i b_i, y \right\rangle \stackrel{(11)}{=} \left\langle \sum_{i=1}^{n} \alpha_i a_i, y \right\rangle$$

$$= \sum_{i=1}^{n} \alpha_i \langle a_i, y \rangle \stackrel{(15)}{\geq} \sum_{i=1}^{n} \alpha_i L(a_i).$$

Hence, at all positions we have in fact "=", in particular

$$\sum_{i=1}^{k} \beta_i \langle b_i, x \rangle = \sum_{i=1}^{k} \beta_i L(b_i).$$

From (13) and $\beta_1, \ldots, \beta_k > 0$ we conclude $\langle b_i, x \rangle = L(b_i), \forall i = 1, \ldots, k$, hence, by (14), $\langle b_i, x \rangle = \langle b_i, y \rangle$, $\forall i = 1, \ldots, k$. Now Proposition 5.2 proves x = y. \square

The goal of the second step is Lemma 5.8, which can be understood as a condensation and interpretation of the result developed in Lemma 5.3: it's a step into the dual space of \mathbb{R}^k connecting vertices in the primal space with simplices in the dual space.

5.4. Definition.

- (a) For any set $X \subseteq \mathbb{R}^k$ let $\operatorname{int}(X)$ be the *interior of* X (w.r.t. the usual topology on \mathbb{R}^k), i.e., $\operatorname{int}(X) := \{x \in \mathbb{R}^k \mid \exists \epsilon > 0. \ B_{\epsilon}(x) \subseteq X\}$, where $B_{\epsilon}(x) := \{y \in \mathbb{R}^k \mid || \ x y || < \epsilon\}$ and $|| \cdot ||$ is some fixed norm on \mathbb{R}^k .
- (b) For any finite set $\mathcal{A} = \{a_1, \ldots, a_n\} \subseteq \mathbb{R}^k$ let $\operatorname{conv}(\mathcal{A})$ be the convex hull of \mathcal{A} , i.e., $\operatorname{conv}(\mathcal{A}) := \{\sum_{i=1}^n \lambda_i a_i | \lambda_1, \ldots, \lambda_n \geq 0, \sum_{i=1}^n \lambda_i = 1\}.$
- (c) For every basis $\mathcal{B} = \{b_1, \ldots, b_k\}$ of \mathbb{R}^k let $S(\mathcal{B}) := \text{conv}(\{0\} \cup \mathcal{B})$, i.e., $S(\mathcal{B})$ is the (k-dimensional) simplex spanned by $0, b_1, \ldots, b_k$. \square
- **5.5. Corollary.** For every basis $\mathcal{B} = \{b_1, \ldots, b_k\}$ of \mathbb{R}^k we have:

(a)
$$S(\mathcal{B}) = \{\sum_{i=1}^k \lambda_i b_i \mid \lambda_1, \dots, \lambda_k \ge 0, \sum_{i=1}^k \lambda_i \le 1\}.$$

(b)
$$\operatorname{int}(S(\mathcal{B})) = \{\sum_{i=1}^k \lambda_i b_i \mid \lambda_1, \dots, \lambda_k > 0, \sum_{i=1}^k \lambda_i < 1\}. \quad \Box$$

Proof. (a) is trivial, and (b) can be deduced straightforwardly from (a), since \mathcal{B} is a basis of \mathbb{R}^k .

5.6. Definition. Two bases \mathcal{A} and \mathcal{B} of \mathbb{R}^k are called *compatible* if

$$int(S(A)) \cap int(S(B)) \neq \emptyset.$$

Otherwise \mathcal{A} and \mathcal{B} are called *incompatible*. \square

Clearly, compatibility is no equivalence relation on the class of all bases of \mathbb{R}^k , since it is not transitive. But at least we get:

- **5.7. Corollary.** Let \mathcal{A} and \mathcal{B} be bases of \mathbb{R}^k . Then we have:
- (a) A and A are compatible.
- (b) \mathcal{A} and \mathcal{B} compatible $\Longrightarrow \mathcal{B}$ and \mathcal{A} compatible. \square

Proof. (a) is a consequence of Corollary 5.5(b), and (b) is trivial. \Box

5.8. Lemma. Let $A_{\geq} \subseteq \mathbb{R}^k$ be finite and $L \colon A_{\geq} \to \mathbb{R}$. Then there exists an injective map

$$\mathrm{choice}_L \colon \ \mathcal{E}(\mathcal{M}(L)) \to \mathrm{BAS}(\mathcal{A}_{\geq})$$

such that the image $\operatorname{choice}_L(\mathcal{E}(\mathcal{M}(L)))$ is a set of pairwise incompatible bases of \mathbb{R}^k .

Proof. Let $\mathcal{A}_{\geq} \subseteq \mathbb{R}^k$ be finite, and let $L: \mathcal{A}_{\geq} \to \mathbb{R}$. By Corollary 3.7(b), for every $x \in \mathcal{E}(\mathcal{M}(L))$ there exists a $\mathcal{B}(x) \in \text{BAS}(\mathcal{A}_{\geq})$ such that $x \in \mathcal{M}(L, \mathcal{B}(x))$. Define choice_L: $\mathcal{E}(\mathcal{M}(L)) \to \text{BAS}(\mathcal{A}_{\geq})$, $x \mapsto \mathcal{B}(x)$. From Corollary 5.5(b) and Lemma 5.3 we receive

$$\forall x, y \in \mathcal{E}(\mathcal{M}(L))$$

($\mathcal{B}(x)$ and $\mathcal{B}(y)$ compatible $\implies x = y$),

hence $\operatorname{choice}_L(\mathcal{E}(\mathcal{M}(L))) = \{\mathcal{B}(x) | x \in \mathcal{E}(\mathcal{M}(L))\}$ is a set of pairwise incompatible bases of \mathbb{R}^k . By Corollary 5.7(a) choice_L is also injective.

On the one hand, in steps 1 and 2 we were not concerned with the specific properties of 0/1-bases. This alters in $step\ 3$, which has Lemma 5.12 as its goal. On the other hand, in step 3 it will not be necessary to deal with the lower bounds L like above: We will now just consider 0/1-bases without regarding their roles defining vertices of polyhedra.

5.9. Proposition. For every 0/1-basis \mathcal{B} of \mathbb{R}^k we have $S(\mathcal{B}) \subseteq [0; 1]^k$.

Proof. The claim follows from $\{0\} \cup \mathcal{B} \subseteq \{0, 1\}^k$, which implies $S(\mathcal{B}) = conv(\{0\} \cup \mathcal{B}) \subseteq conv(\{0, 1\}^k) = [0; 1]^k$.

5.10. Definition. Let Vol_k be the *k*-dimensional volume, i.e., the Lebesgue measure on the Borel sets of \mathbb{R}^k .

5.11. Lemma. For every 0/1-basis² \mathcal{B} of \mathbb{R}^k we have $\operatorname{Vol}_k(S(\mathcal{B})) \geq \frac{1}{k!}$.

Proof. For every basis $\mathcal{B} = \{b_1, \ldots, b_k\}$ of \mathbb{R}^k there is a well-known formula for the volume of the simplex $S(\mathcal{B})$ (e.g., see [7], p. 374): $\operatorname{Vol}_k(S(\mathcal{B})) = \frac{1}{k!} \cdot |\det \bar{\mathcal{B}}|$, where $\bar{\mathcal{B}} := (b_1 \ldots b_k)$, if we consider b_1, \ldots, b_k as column vectors. Now let \mathcal{B} especially be a 0/1-basis of \mathbb{R}^k . Since then $\bar{\mathcal{B}}$ is a matrix with integer entries, its determinant $d := \det \bar{\mathcal{B}}$ is also an integer. But $d \neq 0$, since \mathcal{B} is a basis of \mathbb{R}^k . Hence $|d| \geq 1$.

5.12. Lemma. Every set of pairwise incompatible 0/1-bases of \mathbb{R}^k has at most k! elements. \square

Proof. Let $\mathcal{B}_1, \ldots, \mathcal{B}_n$ be pairwise incompatible 0/1-bases of \mathbb{R}^k . Then we have:

$$\frac{n}{k!} = \sum_{i=1}^{n} \frac{1}{k!} \stackrel{5.11}{\leq} \sum_{i=1}^{n} \operatorname{Vol}_{k}(S(\mathcal{B}_{i}))$$

$$= \sum_{i=1}^{n} \operatorname{Vol}_{k}(\operatorname{int}(S(\mathcal{B}_{i}))) \stackrel{(*)}{=} \operatorname{Vol}_{k} \left(\bigcup_{i=1}^{n} \operatorname{int}(S(\mathcal{B}_{i}))\right)$$

$$\leq \operatorname{Vol}_{k} \left(\bigcup_{i=1}^{n} S(\mathcal{B}_{i})\right) \stackrel{5.9}{\leq} \operatorname{Vol}_{k} \left([0; 1]^{k}\right) = 1,$$

where (*) follows from the assumption that $\mathcal{B}_1, \ldots, \mathcal{B}_n$ are pairwise incompatible.

Now, as our *fourth and last step* we just have to combine the results of step 2 and step 3:

5.13. Main Theorem.
$$\psi(\mathcal{P}(\Omega_k)) \leq k!$$
.

Proof. Let $L: \mathcal{P}(\Omega_k) \to \mathbb{R}$. According to the definition of the psi-function (3.8(b)) we have to prove that $|\mathcal{E}(\mathcal{M}(L))| \leq k!$. By Lemma 5.8 there exists an injective map

$$\operatorname{choice}_L : \mathcal{E}(\mathcal{M}(L)) \to \operatorname{BAS}(\mathcal{P}(\Omega_k))$$

such that $\operatorname{choice}_L(\mathcal{E}(\mathcal{M}(L)))$ is a set of pairwise incompatible 0/1-bases of \mathbb{R}^k . Hence $|\mathcal{E}(\mathcal{M}(L))| = |\operatorname{choice}_L(\mathcal{E}(\mathcal{M}(L)))| \le k!$, where the inequality is a consequence of Lemma 5.12.

We conclude (cf. the definitions of (WAC2), (WEC), and ${\cal M}_k^F$ in Section 4):

5.14. Corollary.

- (a) (WAC2) is true.
- (b) (WEC) is true.

(c)
$$M_k^F = k!$$
 for all $k \ge 1$.

Proof.

- (a) This is the statement of the Main Theorem.
- (b) Follows from (a) and Lemma 4.2.
- (c) Follows from (b) and Lemma 4.1. \Box

6 Concluding Remarks

How many vertices structures can have, is an important question within the theory of interval probability. For example, it is often necessary to minimize or maximize linear functionals subject to structures, and therefore the complexity of corresponding algorithms could be estimated adequately. We here computed the smallest upper bound for the number of vertices.

Acknowledgements

I thank Thomas Augustin, Peter Gritzmann, Kurt Weichselberger, and three anonymous referees for their valuable comments. Especially I thank Franz Ferschl — as well for structural pieces of advice as for algebraic and geometric stimuli.

References

- [1] A. Chateauneuf and J.-Y. Jaffray. Some characterizations of lower probabilities and other monotone capacities through the use of Möbius inversion. *Mathematical Social Sciences*, 17: 263–283, 1989
- [2] G. Ewald. Combinatorial Convexity and Algebraic Geometry. Springer, New York, 1996.
- [3] P.J. Huber and V. Strassen. Minimax tests and the Neyman-Pearson lemma for capacities. *Annals of Statistics*, 1: 251–263, 1973. Correction: 2: 223–224, 1974.
- [4] I. Levi. *The Enterprise of Knowledge*. MIT Press, Cambridge, Massachusetts, 1980.
- [5] M.W. Padberg. Linear Optimization and Extensions. Springer, Berlin, Heidelberg, New York, 1995.
- [6] S. Shapley. Cores of convex games. *International Journal of Game Theory*, 1: 12–26, 1971.
- [7] K.T. Smith. *Primer of Modern Analysis*. Springer, New York, Berlin, Heidelberg, 1983.
- [8] P. Walley. Statistical Reasoning with Imprecise Probabilities. Chapman and Hall, London, New York, 1991.
- [9] A. Wallner. Beiträge zur Theorie der Intervallwahrscheinlichkeit — Der Blick über Kolmogorov und Choquet hinaus. Dr. Kovač, Hamburg, 2002.
- [10] K. Weichselberger. Elementare Grundbegriffe einer allgemeineren Wahrscheinlichkeitsrechnung I — Intervallwahrscheinlichkeit als umfassendes Konzept. In cooperation with T. Augustin and A. Wallner. Physica, Heidelberg, 2001.
- [11] G.M. Ziegler. Lectures on Polytopes, Springer, New York, Berlin, Heidelberg, 1995.

 $^{^2{\}rm The}$ statement is also true for bases, in which all the vectors have only integer components.