A SHARP NON-CONVEXITY BOUND FOR PARTITION RANGES OF VECTOR MEASURES WITH ATOMS

PIETER C. ALLAART

University of North Texas, Denton, TX 76203

Abstract. A sharp upper bound is given for the degree of non-convexity of the partition range of a finite-dimensional vector measure, in terms of the maximum (one-dimensional) mass of the atoms of that measure. This upper bound improves on a bound of Hill and Tong (1989) by an order of magnitude $\sqrt{n}$. Its proof uses several ideas from graph theory, combinatorics, and convex geometry. Applications are given to optimal-partitioning and fair division problems.

Date: March 4, 1999.

1991 *Mathematics Subject Classification.* Primary 28B05, 51K99, 52B12; Secondary 05C20, 60A10, 90A06.

Key words and phrases. Partition range, optimal-partitioning, convexity theorem, vector measure, vector atom, Hausdorff-distance, digraph, tree.

A Sharp Non-convexity Bound

Author's address:

Pieter C. Allaart

Mathematics Department

University of North Texas

Denton, TX 76203-5118

1. Introduction

It is well known that the range of a finite-dimensional, atomless finite vector measure is convex and compact; this is Lyapounov's celebrated convexity theorem [13] of 1940. Somewhat lesser known is a generalization of Lyapounov's theorem due to Dvoretzky, Wald and Wolfowitz [3], which says that the *partition range*

$$\{(\mu_1(A_1), \mu_2(A_2), \ldots, \mu_n(A_n)) : (A_1, \ldots, A_n) \text{ is a partition}\}$$

of a finite-dimensional, atomless finite vector measure $\vec{\mu}$ is also both convex and compact.

If the vector measure has atoms, then convexity of both the range and the partition range may fail in general, as is best seen by considering a vector measure supported on a finite set. The question of necessary and sufficient conditions for the range (or partition range) to be convex was addressed by Gouweleeuw [6]. A different approach may be based on the following idea: if the sizes of the atoms are small, then the range (and partition range) are very close to being convex. Elton and Hill [4] have proved a bound on the degree of non-convexity of the range of a vector measure, as a function of the maximum atom size. The aim of the present paper is to derive an analogous sharp bound on the degree of non-convexity of the partition range. The main result is that

> *the Hausdorff-distance (with respect to the* sup*-norm) from the partition range of an n-dimensional vector measure to its convex hull is at most $\alpha(n-1)/n$, where α is the size of the largest atom.*

This bound is attained, and improves on an earlier inequality of Hill and Tong [11].

A formal statement of the main result is given in Section 2 (cf. Theorem 2.5), and a proof is presented in Sections 3 and 4. The proof consists of a number of steps: first the problem is reduced to the case of a purely atomic vector measure with at most a finite number of atoms, using an approximation scheme analogous to that used in [4]; then it is shown, using the Shapley-Folkman lemma from convex geometry, that only vector measures with no more than n atoms need to be considered; and finally the theorem is proved for this special case (of a purely atomic vector measure with at most n atoms), using a variety of tools from graph theory, combinatorics and geometry, such as directed graphs and trees, convex polytopes, and supporting hyperplanes.

Section 3 contains the two above-mentioned reduction steps, discusses some fundamental geometric properties of the partition range, and introduces a certain type of directed graphs (to be called *graphs of options*) which play an important role in the proof. The main body of the proof is contained in Section 4.

The reason for choosing the *sup*-norm will become apparent in Section 5, where it will be pointed out how Theorem 2.5 can be used to generalize existing optimal-partitioning inequalitites for atomless measures to measures with atoms. As an example, generalizations will be given of two well-known inequalities of Elton, Hill and Kertz (1986) and Hill (1987). These results turn out to have some interesting implications for the problem of fair division.

2. The main theorem

Throughout this paper, $\mu, \mu_1, \ldots, \mu_n$ will always denote finite, non-negative countably additive measures on a fixed measurable space $(\Omega, \mathcal{F})$. The *vector measure* $\vec{\mu} = (\mu_1, \ldots, \mu_n)$ is defined by $\vec{\mu}(A) := (\mu_1(A), \ldots, \mu_n(A)) \in \mathrm{I\!R}^{\mathrm{n}},\ \mathrm{A} \in \mathcal{F}$.

For a set $B \in \mathcal{F}, \vec{\mu}\lfloor_B$ denotes the restriction of $\vec{\mu}$ to B: $\vec{\mu}\lfloor_B (A) = \vec{\mu}(A \cap B)$. A set $E \in \mathcal{F}$ is called a (scalar) *atom* of μ if $\mu(E) > 0$ and for each $F \subset E, F \in \mathcal{F} : \mu(F) \in \{0, \mu(E)\}$. Similarly, E is a *vector atom* of $\vec{\mu}$ if $\vec{\mu}(E) \neq \vec{0}$ and for each $F \subset E, F \in \mathcal{F} : \vec{\mu}(F) = \vec{\mu}(E)$ or $\vec{\mu}(F) = \vec{0}$. A (vector) measure is *atomless* if it does not have any atoms. A measure (resp. vector measure) is *purely atomic* if is assigns mass 0 (resp. $\vec{0}$) to the complement of the union of its atoms.

Remark 2.1. From the definition of vector atom it can be seen that if E is a vector atom of $\vec{\mu} = (\mu_1, \ldots, \mu_n)$, then

(i) E is a scalar atom of at least one μ_i;

(ii) for each $i \in \{1, \ldots, n\}$, either E is an atom of μ_i, or $\mu_i(E) = 0$.

Conversely, it follows from Lemma 2.4 (iii) in [6] that if E is a scalar atom of μ_i for some i, then E contains a vector atom F of $\vec{\mu}$ with $\vec{\mu}(F) = \vec{\mu}(E)$.

As a consequence, a vector measure is purely atomic if and only if all its component measures are.

A *measurable n-partition* of Ω is an ordered collection $(A_1, \ldots, A_n)$ of subsets of Ω such that $A_i \in \mathcal{F}$ $(i = 1, \ldots, n)$, $A_i \cap A_j = \emptyset$ for all $i \neq j$, and $\bigcup_{i=1}^n A_i = \Omega$. Let Π^n denote the collection of all measurable n-partitions of Ω. Because all partitions considered in this paper are measurable n-partitions, we will simply use the word 'partition'.

Throughout this paper, the following notation will be used: for a partition $\mathbf{A} := (A_i)_{i=1}^n$ and a vector measure $\vec{\mu}$, $\overrightarrow{\mu(\mathbf{A})}$ denotes the vector $(\mu_1(A_1), \ldots, \mu_n(A_n))$ in $\mathbb{R}^n$.

Definition 2.2. $\mathcal{PR}(\vec{\mu}) := \{\overrightarrow{\mu(\mathbf{A})} : \mathbf{A} \in \Pi^n\}$ *is the* partition range *of $\vec{\mu}$.*

Proposition 2.3. [Dvoretzky, Wald and Wolfowitz (1951)]. *If $\vec{\mu}$ is atomless, then $\mathcal{PR}(\vec{\mu})$ is convex and compact.*

In fact the theorem of Dvoretzky, Wald and Wolfowitz is more general: it says that the *matrix range* $\{(\mu_i(A_j))_{i=1,j=1}^{n,k} : (A_i)_{i=1}^n \in \Pi^n\}$ is convex and compact for each $k \in \mathbb{N}$. This paper, however, will focus on the partition range.

The main goal of this paper is to generalize the above convexity result to measures with atoms. In order to do so, the following notation is needed. For a vector $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, let $\|x\|_p$ denote the ℓ_p-norm $(\sum_{i=1}^n |x_i|^p)^{1/p}$ for $1 \leq p < \infty$, and $\max_{1 \leq i \leq n} |x_i|$ for $p = \infty$. For vectors x and y in $\mathbb{R}^n$, let $d_p(x, y) = \|x - y\|_p$ denote the distance between x and y. For a set S in $\mathbb{R}^n$ and a point x in $\mathbb{R}^n$, let $d_p(x, S) = \inf_{y \in S} d_p(x, y)$ denote the distance from x to S, and $D_p(S)$ the Hausdorff distance from S to its convex hull $\mathrm{co}(S)$:

$$D_p(S) := \sup_{x \in \mathrm{co}(S)} d_p(x, S)$$

Definition 2.4. *For $\alpha \geq 0$, $\mathcal{P}_n(\alpha)$ is the collection of all n-dimensional finite vector measures $\vec{\mu}$ for which $\|\vec{\mu}(E)\|_\infty \leq \alpha$ for each atom E of $\vec{\mu}$.*

The following theorem is the main result of this paper. It generalizes the convexity part of Proposition 2.3.

Theorem 2.5. *If $\vec{\mu} \in \mathcal{P}_n(\alpha)$, then*

$$D_\infty(\mathcal{PR}(\vec{\mu})) \leq \frac{n-1}{n}\alpha,$$

and this bound is attained.

Example 2.6. *(sharpness of Theorem 2.5) Let $\mu_i = \alpha\delta_{\{0\}}, i = 1, \ldots, n$, where δ denotes Dirac measure. Then $\mathcal{PR}(\vec{\mu}) = \{\alpha e_i : i = 1, \ldots, n\}$, where e_i denotes the $i-$th unit vector in $\mathbb{R}^n$ with 1 in the $i-$th position and zeros elsewhere. It follows that $\mathrm{co}(\mathcal{PR}(\vec{\mu})) = \{x \in \mathbb{R}^n_+ : \sum_{i=1}^n x_i = \alpha\}$. In particular, $y = (\alpha/n, \ldots, \alpha/n) \in \mathrm{co}(\mathcal{PR}(\vec{\mu}))$, and for any $x \in \mathcal{PR}(\vec{\mu}), \|x - y\|_\infty = \alpha(n-1)/n$.*

The following immediate consequence of Theorem 2.5 improves on an earlier result of Hill and Tong ([11], Theorem 3.2).

Corollary 2.7. *If $\vec{\mu} \in \mathcal{P}_n(\alpha)$, then*

$$D_2(\mathcal{PR}(\vec{\mu})) \leq \frac{n-1}{\sqrt{n}}\alpha.$$

Example 2.8. *The bound in Corollary 2.7 is of the correct order of magnitude in n: let $\mu_i = \alpha\delta_{\{i\}}, i = 1, \ldots, n$; then $\mathcal{PR}(\vec{\mu}) = \{0, \alpha\}^n$, hence $\mathrm{co}(\mathcal{PR}(\vec{\mu})) = [0, \alpha]^n$. In particular, $y = (\alpha/2, \ldots, \alpha/2) \in \mathrm{co}(\mathcal{PR}(\vec{\mu}))$, and for any $x \in \mathcal{PR}(\vec{\mu})$, $\|x - y\|_2 = \alpha\sqrt{n}/2$.*

3. Preliminaries

3.1. **Digraphs and out-trees.** An important role in the proof of Theorem 2.5 is played by directed graphs or *digraphs*. A *digraph* G consists of a finite set $\mathcal{V} = \mathcal{V}(G)$ of *vertices* and a collection $\mathcal{E} = \mathcal{E}(G)$ of ordered pairs of distinct vertices called *arcs*. An arc $e = (u, v)$ is said to be *directed from u to v*. The vertex u is called the *initial endpoint* of e, denoted $\mathrm{ini}(e)$, and v is called the *terminal endpoint* of e, denoted $\mathrm{ter}(e)$. If v is a vertex of G, we will often write $v \in G$ instead of $v \in \mathcal{V}(G)$ for brevity. Similarly, if e is an arc of G we write $e \in G$.

A *subgraph* of G is a digraph G' such that $\mathcal{V}(G') \subset \mathcal{V}(G)$ and $\mathcal{E}(G') \subset \mathcal{E}(G)$. If G' is a subgraph of G, then we say that G *contains* G', and denote $G' \subset G$.

The *indegree* of a vertex v is the number of arcs directed *to* it. Similarly, the *outdegree* of v is the number of arcs directed *from* it.

A *path* P is an alternating sequence $(v_0, e_1, v_1, e_2, \ldots, e_m, v_m)$ of distinct vertices and arcs so that $e_i = (v_{i-1}, v_i)$ for all $i = 1, \ldots, m$. The path P is said to be directed from v_0 to v_m. The vertex v_0 is called the *initial endpoint* of P, denoted $\mathrm{ini}(P)$, and v_m is called the *terminal endpoint* of P, denoted $\mathrm{ter}(P)$. If a digraph G contains a path from u to v, then v is said to be *reachable* (in G) from u. Every vertex is reachable from itself via the path consisting of just that vertex.

A digraph T is called an *out-tree* if exactly one vertex v_0 has indegree 0 and all other vertices have indegree 1, and each vertex of T is reachable from v_0. The vertex v_0 is called the *root* of T. (The converse structure, with all arcs pointing in the opposite direction, is called an *in-tree*). An out-tree T is called *maximal in* G if T is a subgraph of G and T contains all vertices of G that are reachable in G from the root of T. (See Figure 1 for an example.)

The following two observations will be quite useful. Their verification is left to the interested reader.

(OT1) In each out-tree T there is a unique path from the root to every other vertex.

(OT2) Every out-tree $T \subset G$ can be extended to a maximal out-tree in G.

Lemma 3.1. *Let T be an out-tree, let v_0 be its root, and let G be a digraph containing v_0. Then there exists a maximal out-tree T' in G with root v_0, such that each path P with initial vertex v_0 which is both in T and in G, is also in T'.*

Proof. Let T_0 be the smallest digraph that contains all such paths; obviously T_0 is a subgraph of G. Then for every vertex $v \neq v_0$ of T_0 there is a path from v_0 to v in T_0, by the minimality of T_0. Moreover, v_0 has indegree 0 (again by the minimality of T_0), and all other vertices of T_0 have indegree at most 1, since T_0 is a subgraph of T, and T is an out-tree. On the other hand, the indegree of these vertices is at

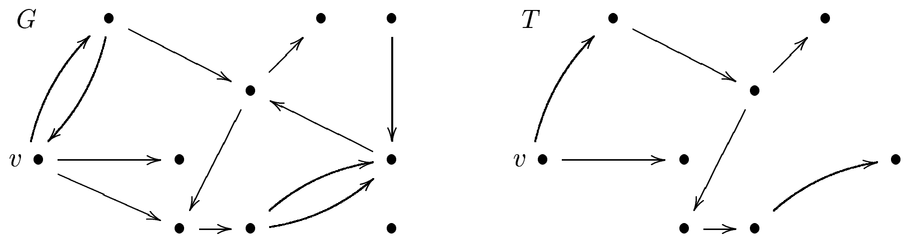

FIGURE 1. A sample digraph G, and a maximal out-tree T in G with root v.

least one since they are reachable from v_0. Hence T_0 is an out-tree with root v_0, which by (OT2) can be extended to a maximal out-tree in G. $\square$

For an extensive discussion of directed graphs and their applications the reader is referred to Harary *et al.* [8].

3.2. **Reduction to the purely-atomic case.** The purpose of this subsection is to reduce the problem to the case of a purely atomic vector measure with at most a finite number of atoms. This will be done using an approximation machinery analogous to that used by Elton and Hill [4]. The proofs of the following lemmas are omitted since they are comletely analogous to the proofs that appear in [4], §3.

Lemma 3.2. *For each $\vec{\mu}$, each $p \in [1, \infty]$, and each $\varepsilon > 0$, there exists a measurable partition $\{B_i\}_{i=1}^N$ of Ω such that for each $\mathbf{A} \in \Pi^n$ there exists a partition $(I_j)_{j=1}^n$ of $\{1, \dots, N\}$ satisfying*

$$\left\| \overrightarrow{\mu(\mathbf{A})} - \left(\mu_1(\bigcup_{i \in I_1} B_i), \dots, \mu_n(\bigcup_{i \in I_n} B_i) \right) \right\|_p < \varepsilon.$$

Lemma 3.3. *For each $\vec{\mu} \in \mathcal{P}_n(\alpha)$ and each $B \in \mathcal{F}$, there exists a measurable partition $\{B_j\}_{j=1}^k$ of B such that $\mu_i(B_j) \leq \alpha$ for all $j \leq k$ and $i \leq n$.*

Proposition 3.4. *For each $\vec{\mu} \in \mathcal{P}_n(\alpha)$, each $p \in [1, \infty]$, and each $\varepsilon > 0$, there is a purely atomic vector measure $\vec{\mu}_0 \in \mathcal{P}_n(\alpha)$ with finitely many atoms, such that*

$$D_p(\mathcal{PR}(\vec{\mu})) \leq D_p(\mathcal{PR}(\vec{\mu}_0)) + \varepsilon.$$

Proof. Follows immediately by Lemma 3.2 and a repeated application of Lemma 3.3. $\square$

3.3. **Geometric properties of the partition range.** In view of Proposition 3.4, we will assume for the remainder of this section that $\vec{\mu}$ is purely atomic with at most a finite number of atoms. Thus $V = \mathcal{PR}(\vec{\mu})$ is a finite set in $\mathbb{R}^n$ in view of Remark 2.1, so its convex hull $C = \text{co}(\mathcal{PR}(\vec{\mu}))$ is a convex polytope. Hence C can be written as a finite intersection $\bigcap_{i=1}^N H_i$ of halfspaces of the form $H_i = \{x : \langle p^{(i)}, x\rangle \leq c_i\}$, where $p^{(i)} \in \mathbb{R}^n$ with $p^{(i)} \neq 0$, and $\langle p^{(i)}, x\rangle := \sum_{j=1}^n p_j^{(i)} x_j$. Each hyperplane $L_i = \{x : \langle p^{(i)}, x\rangle = c_i\}$ that has a non-empty intersection with C is a *supporting hyperplane* of C, and the vector $p^{(i)}$ is an *outward normal* of L_i. The intersections of the supporting hyperplanes with C are called the *faces* of C. By an outward normal of a face F will be meant an outward normal of a supporting hyperplane containing F. The union of all the faces of C is called the boundary of C, denoted $\text{Bd}(C)$. It is clear that each face of C is both convex and compact. The *dimension* of a face is the affine dimension of the smallest affine subspace containing that face. A face is called *maximal* if it is not contained in any other face.

We will assume that $\mu_i(\Omega) > 0$ for all $i \leq n$ (this is no real restriction: e.g. if $\mu_n(\Omega) = 0$, then consider the $(n-1)$-dimensional vector measure $(\mu_1, \ldots, \mu_{n-1})$ instead of $\vec{\mu}$). Then V contains the n points $\mu_i(\Omega)e_i,\ i = 1, \ldots, n$ (where e_i is the i-th standard unit vector in $\mathbb{R}^n$), so the affine dimension of V is at least $n-1$. It is easily seen that each maximal face of C then has dimension $n-1$, and hence has a unique outward normal.

Lemma 3.5. *Let F be a maximal face of C, and let $p = (p_1, \ldots, p_n)$ be the unique outward normal of F. Then either $p_i \geq 0$ for all $i \leq n$, or $p_i \leq 0$ for all $i \leq n$.*

Proof. By contradiction. Suppose there exist i and j, $i \neq j$, such that $p_i > 0$ and $p_j < 0$. Then p is not the j-th standard unit vector, so F does not lie in the hyperplane $\{x : x_j = 0\}$, and hence there is a point $x \in F \cap V$ with $x_j > 0$. Let $(A_l)_{l=1}^n$ be a partition such that $x_l = \mu_l(A_l), l = 1, \ldots, n$. Define the partition $\mathbf{A}' = (A_l')_{l=1}^n$ by $A_i' = A_i \cup A_j$, $A_j' = \emptyset$, and $A_l' = A_l$ for $l \notin \{i, j\}$. Let $x' = \overrightarrow{\mu(\mathbf{A}')}$. Then $x' \in V$, and

$$\langle p, x'\rangle = \sum_{l \notin \{i,j\}} p_l x_l + p_i(x_i + \mu_i(A_j)) \geq \sum_{l \neq j} p_l x_l > \langle p, x\rangle$$

where the first inequality follows since $p_i > 0$, and the second since $p_j < 0$ and $x_j > 0$. But this contradicts the fact that p is an outward normal of F. $\square$

Lemma 3.6 (The Key Equation). *Let F be a face of C with outward normal p. Then for any two partitions $(A_l)_{l=1}^n$ and $(B_l)_{l=1}^n$ in Π_F, indices i and j $(i \neq j)$ and subset $E \subset A_i \cap B_j$,*

$$p_i \mu_i(E) = p_j \mu_j(E) \tag{1}$$

Proof. There exists a constant $c \in \mathbb{R}$ such that

$$\langle p, y \rangle = c \quad \text{for all } y \in F, \quad \text{and} \quad \langle p, y \rangle \leq c \quad \text{for all } y \in C. \tag{2}$$

Consider the partition $(A'_l)_{l=1}^n$ given by $A'_i = A_i \backslash E, A'_j = A_j \cup E$ and $A'_l = A_l$ for $l \notin \{i, j\}$. Then by (2),

$$p_1 \mu_1(A_1) + \cdots + p_n \mu_n(A_n) = c;$$

$$p_1 \mu_1(A'_1) + \cdots + p_n \mu_n(A'_n) \leq c.$$

Subtracting these two equations yields $p_i \mu_i(E) \geq p_j \mu_j(E)$. Similarly, by considering $(B_l)_{l=1}^n$ and a suitable partition $(B'_l)_{l=1}^n$, we can derive $p_i \mu_i(E) \leq p_j \mu_j(E)$. $\square$

The last result in this subsection, which concerns a further reduction to vector measures which have no more than n atoms, is based on the following Carathéodory-type result (see [1]).

Let the vector sum $V_1 \oplus V_2$ of two sets V_1 and V_2 be defined by $V_1 \oplus V_2 = \{v_1 + v_2 : v_1 \in V_1, v_2 \in V_2\}$.

Lemma 3.7. (Shapley and Folkman) *Let $V_1, \ldots, V_k$ be nonempty subsets of $\mathbb{R}^{\mathrm{n}}$. Then for each $y \in \mathrm{co}(V_1) \oplus \cdots \oplus \mathrm{co}(V_k)$ there exists a representation $y = x_1 + \cdots + x_k$, with $x_i \in \mathrm{co}(V_i)$ for all i, but $x_i \notin V_i$ for at most n indices i.*

Proposition 3.8. *Let the atoms of $\vec{\mu}$ be $E_1, \ldots, E_k$. Then for each $p \in [1, \infty]$,*

$$D_p(\mathcal{PR}(\vec{\mu})) \leq \max\{D_p(\mathcal{PR}(\vec{\mu}\mid_{(\cup_{i \in I} E_i)})) \mid I \subset \{1, \ldots, k\}, |I| \leq n\}.$$

Proof. Let $V_i := \mathcal{PR}(\vec{\mu}\downharpoonright_{E_i}), i = 1, \ldots, k$. Then $\mathcal{PR}(\vec{\mu}) = V_1 \oplus \cdots \oplus V_k$ and hence $\mathrm{co}(\mathcal{PR}(\vec{\mu})) = \mathrm{co}(V_1) \oplus \cdots \oplus \mathrm{co}(V_k)$ (see, e.g. [7], p. 316). Let $y \in \mathrm{co}(\mathcal{PR}(\vec{\mu}))$. By Lemma 3.7 there is a representation $y = x_1 + \cdots + x_k$, with $x_i \in \mathrm{co}(V_i)$ for all i, but $x_i \notin V_i$ for at most n indices i. Let I denote the set of indices i for which $x_i \notin V_i$. Then $d_p(y, V_1 \oplus \cdots \oplus V_k) \leq d_p(\sum_{i \in I} x_i, \oplus_{i \in I} \mathrm{co}(V_i)) \leq D_p(\oplus_{i \in I} V_i)$. Since $\oplus_{i \in I} V_i = \mathcal{PR}(\vec{\mu}\downharpoonright_{(\cup_{i \in I} E_i)})$, this completes the proof. □

3.4. **Graphs of options.** Throughout this subsection, let F denote an arbitrary, but fixed face of $\mathrm{co}(\mathcal{PR}(\vec{\mu}))$.

Definition 3.9. *$\Pi_F = \Pi_F^n$ is the collection of all measurable partitions $\mathbf{A} = (A_l)_{l=1}^n$ such that $\overrightarrow{\mu(\mathbf{A})} \in F$.*

Definition 3.10. *For a partition $\mathbf{A} = (A_l)_{l=1}^n \in \Pi_F$, the* graph of options $G(\mathbf{A}; F)$ *of $\mathbf{A}$ with respect to F is the digraph with vertex set $\{1, 2, \ldots, n\}$ and an arc from i to j ($i \neq j$) for each atom $E \subset A_i$ for which $\mathbf{A}' = (A'_l)_{l=1}^n \in \Pi_F$, where $A'_i = A_i \backslash E, A'_j = A_j \cup E$ and $A'_l = A_l, l \notin \{i, j\}$.*

To distinguish between two or more arcs from i to j, each arc is labeled with the corresponding atom. Thus to each arc e corresponds an ordered triplet (i, j, E), and vice-versa. If $e = (i, j, E)$, then we shall denote $\mathrm{at}(e) = E$.

The intuitive interpretation of an arc (i, j, E) is that if E is moved from A_i to A_j, then the resulting partition is still in Π_F.

Definition 3.11. *$\mathcal{E}(F)$ is the set of arcs that occur in $G(\mathbf{A}; F)$ for at least one partition $\mathbf{A} \in \Pi_F$.*

Lemma 3.12. *Let $e \in \mathcal{E}(F)$ and $\mathbf{A} = (A_l)_{l=1}^n \in \Pi_F$. Then $e \in G(\mathbf{A}; F)$ if and only if* $\mathrm{at}(e) \subset A_{\mathrm{ini}(e)}$.

Proof. Suppose first that $\mathrm{at}(e) \subset A_{\mathrm{ini}(e)}$. Let $e = (i, j, E)$, so $E \subset A_i$. Let L be a supporting hyperplane of $\mathrm{co}(\mathcal{PR}(\vec{\mu}))$ that contains F, and let $p = (p_1, \ldots, p_n)$ be an outward normal of L. Then by Definition 3.11, $e \in G(\mathbf{B}; F)$ for some partition $\mathbf{B} \in \Pi_F$, so by Definition 3.10 and Lemma 3.6 it follows that $p_i \mu_i(E) = p_j \mu_j(E)$. Hence with $\mathbf{A}' = (A'_l)_{l=1}^n$ defined as in Definition 3.10 it follows that the point

$x' := \overrightarrow{\mu(\mathbf{A}')}$ lies in L. Since also $x' \in \mathcal{PR}(\vec{\mu})$, it follows that $x' \in F$. Hence $e \in G(\mathbf{A}; F)$. The converse is trivial. $\square$

Definition 3.13. *Two not necessarily distinct arcs e and e' are* related *if* $\mathrm{ini}(e) = \mathrm{ini}(e')$ *and* $\mathrm{at}(e) = \mathrm{at}(e')$.

Lemma 3.14. *Let e and e' be two related arcs in $\mathcal{E}(F)$, and let G be a graph of options with respect to F. Then $e \in G$ if and only if $e' \in G$.*

Proof. Immediate from Definition 3.13 and Lemma 3.12. $\square$

Definition 3.15. *For a partition $\mathbf{A}$ in Π_F and a path Q in $G(\mathbf{A}; F)$, define a new partition $M_Q(\mathbf{A})$ as follows. Let $e_1, \ldots, e_k$ denote the arcs of Q, ordered so that* $\mathrm{ini}(e_r) = \mathrm{ter}(e_{r+1}), r = 1, \ldots, k-1$ *(so e_k is the initial, and e_1 the terminal arc of Q). For brevity, write $i_0 = \mathrm{ter}(e_1)$, and for $r = 1, \ldots, k$, $i_r = \mathrm{ini}(e_r)$ and $E_r = \mathrm{at}(e_r)$. Now define $M_Q(\mathbf{A}) := \mathbf{A}'$, where $\mathbf{A}' = (A'_l)_{l=1}^n$ is given by $A'_{i_0} = A_{i_0} \cup E_1, A'_{i_r} = (A_{i_r} \cup E_{r+1}) \backslash E_r$ for $r = 1, \ldots, k-1, A'_{i_k} = A_{i_k} \backslash E_k$, and $A'_l = A_l$ for all $l \notin \{i_0, i_1, \ldots, i_k\}$.*

Intuitively, $M_Q(\mathbf{A})$ is the partition arising from A by 'executing' all arcs of Q, i.e. by moving $\mathrm{at}(e)$ from $A_{\mathrm{ini}(e)}$ to $A_{\mathrm{ter}(e)}$ for each $e \in Q$.

Example 3.16. *Suppose that Q consists of the two arcs $(1, 2, E)$ and $(2, 4, \hat{E})$. Then $M_Q(\mathbf{A})$ is the partition resulting from moving E from A_1 to A_2, and moving $\hat{E}$ from A_2 to A_4; more precisely, $M_Q(\mathbf{A}) = (A'_l)_{l=1}^n$, with $A'_1 = A_1 \backslash E, A'_2 = (A_2 \cup E) \backslash \hat{E}, A'_4 = A_4 \cup \hat{E}$, and $A'_l = A_l$ for all $l \notin \{1, 2, 4\}$.*

Lemma 3.17. *Let $\mathbf{A} \in \Pi_F$, let Q be a path in $G(\mathbf{A}; F)$, and $e \in G(\mathbf{A}; F)$.*

(i) $M_Q(\mathbf{A}) \in \Pi_F$.

(ii) If $e \in Q$, then $e \notin G(M_Q(\mathbf{A}); F)$.

(iii) If $\mathbf{A}' = M_Q(\mathbf{A})$, and E is an atom such that $E \subset A_i \cap A'_j$ $(i \neq j)$, then $(i, j, E) \in Q$.

(iv) $e \in G(M_Q(\mathbf{A}); F)$ iff e is not related to any arc of Q.

Proof. Let $\mathbf{A}' = M_Q(\mathbf{A})$, and let $x' = \overrightarrow{\mu(\mathbf{A}')}$. Clearly $x' \in \mathcal{PR}(\vec{\mu})$. An argument similar to that in the proof of Lemma 3.12 shows that x' is also in the supporting hyperplane that contains F. Hence $x' \in F$. This proves (i). Note that (iv) implies (ii), while (iii) follows at once from the definition of $M_Q(\mathbf{A})$ above. It therefore remains to prove (iv).

Let $i = \mathrm{ini}(e)$, and $E = \mathrm{at}(e)$. Since $e \in G(\mathbf{A}; F)$, Lemma 3.12 implies that $E \subset A_i$. Suppose first that e is related to e_r for some $r \in \{1, \ldots, k\}$. Then $i = i_r$ and $E = E_r$, hence $A'_i = A'_{i_r} = (A_{i_r} \cup E_{r+1}) \backslash E_r$ if $r \leq k-1$, or else (if $r = k$) $A'_i = A'_{i_k} = A_{i_k} \backslash E_k$. In both cases it follows that $E \not\subset A'_i$, so by Lemma 3.12 $e \notin G(M_Q(\mathbf{A}); F)$.

Conversely, if $e \notin G(M_Q(\mathbf{A}); F)$, then $E \not\subset A'_i$ by Lemma 3.12, so $E \subset A'_j$ for some $j \neq i$. Hence (iii) implies $(i, j, E) \in Q$. But (i, j, E) is related to e. □

4. Proof of the main theorem

Note that if $n = 1$, then $\mathcal{PR}(\vec{\mu}) = \mathcal{PR}(\mu_1) = \{\mu_1(\Omega)\}$ and the statement of the theorem is trivially true. Therefore it can and will be assumed from here on that $n \geq 2$. In addition it will be assumed that $\alpha = 1$; the general case $\alpha > 0$ then follows easily by rescaling, and the case $\alpha = 0$ follows by continuity.

First we prove a 'face-wise' version of Theorem 2.5.

Proposition 4.1. *Let $\vec{\mu} \in \mathcal{P}_n(1)$ be purely atomic with finitely many atoms, and let $V = \mathcal{PR}(\vec{\mu})$. Then for every face F of* $\mathrm{co}(V)$*,*

$$D_\infty(V \cap F) \leq \frac{n-1}{n}.$$

In order to prove Proposition 4.1, the following lemmas are needed. In Lemmas 4.2-4.4, F denotes a face of $\mathrm{co}(V)$ and y denotes a point in F.

Lemma 4.2. *Let* $\mathbf{A}$ *be a partition in* Π_F*, and let* $x = \overrightarrow{\mu(\mathbf{A})}$*. If*

$$x_1 > y_1 + \frac{n-1}{n}, \tag{3}$$

then every maximal out-tree with root 1 *in* $G(\mathbf{A}; F)$ *contains an* executable arc*, i.e. an arc* e *such that*

$$x_{\mathrm{ter}(e)} + \mu_{\mathrm{ter}(e)}(\mathrm{at}(e)) \leq y_{\mathrm{ter}(e)} + \frac{n-1}{n}. \tag{4}$$

Proof. Let $\hat{F}$ be a maximal face that contains F, and let $p = (p_1, \ldots, p_n)$ be an outward normal of $\hat{F}$. Then p is also an outward normal of F, and by Lemma 3.5 either $p_i \geq 0$ for all i, or $p_i \leq 0$ for all i. Assume without loss of generality that $p_i \geq 0$. (The proof of the other case is similar, the only difference being the directions of the inequalities which involve a factor p_j).

Let T be a maximal out-tree with root 1 in $G := G(\mathbf{A}; F)$. First we claim that T has at least one vertex other than 1. To see this, note that by (3) and since $y \in \mathrm{co}(F \cap V)$, there is a partition $\mathbf{A}' \in \Pi_F$ for which $\mu_1(A'_1) < y_1 < x_1 = \mu_1(A_1)$. Hence there is an atom E such that $E \subset A_1$ but $E \not\subset A'_1$, so $E \subset A'_i$ for some $i \neq 1$, say $E \subset A'_2$. Then the Key Equation (Lemma 3.6) implies that $(A_1 \backslash E, A_2 \cup E, A_3, \ldots, A_n) \in \Pi_F$, which by Definition 3.10 means that $(1, 2, E) \in G$, i.e. 2 is reachable from 1 in G. Since T is maximal, this means that $2 \in T$.

Next, let I denote the vertex set of T. Then for any $j \in I \backslash \{1\}$ there is a unique arc $e_j \in T$ with $j = \mathrm{ter}(e_j)$. Let $E_j = \mathrm{at}(e_j)$, and denote $v_j = \mu_j(E_j)$. The proof will be complete once it has been shown that for some $j \in I \backslash \{1\}$,

$$x_j + v_j \leq y_j + \frac{n-1}{n}, \tag{5}$$

which will yield (4) for $e = e_j$.

The key to the proof of (5) is the following inequality.

$$\sum_{j \in I \backslash \{1\}} p_j v_j \leq \sum_{j \in I \backslash \{i\}} p_j \quad \text{for each } i \in I. \tag{6}$$

For $i = 1$, (6) follows immediately since $v_j \leq 1$ for all j. For $i \neq 1$, let $P = (1 = i_0, e_{i_1}, i_1, \ldots, i_{m-1}, e_{i_m}, i_m = i)$ be the unique path in T from 1 to i. Then by the Key Equation,

$$p_{i_r} v_{i_r} = p_{i_r} \mu_{i_r}(E_{i_r}) = p_{i_{r-1}} \mu_{i_{r-1}}(E_{i_r}) \leq p_{i_{r-1}}, \quad r = 1, \ldots, m. \tag{7}$$

Hence

$$\sum_{j\in I\setminus\{1\}} p_j v_j = \sum_{j\in\{i_1,\ldots,i_m\}} p_j v_j + \sum_{j\notin\{i_0,\ldots,i_m\}} p_j v_j = \sum_{r=1}^{m} p_{i_r} v_{i_r} + \sum_{j\notin\{i_0,\ldots,i_m\}} p_j v_j$$
$$\le \sum_{r=1}^{m} p_{i_{r-1}} + \sum_{j\notin\{i_0,\ldots,i_m\}} p_j = \sum_{j\in I\setminus\{i\}} p_j,$$

where the first and last equalities follow since $i_0 = 1$ and $i_m = i$, respectively, and the inequality follows by (7) and since $v_j \le 1$. This completes the proof of (6).

Applying (6) for each i in turn now yields

$$\begin{aligned}\sum_{j\in I\setminus\{1\}} p_j v_j &= \frac{1}{|I|}\cdot\sum_{i\in I}\sum_{j\in I\setminus\{1\}} p_j v_j \le \frac{1}{|I|}\cdot\sum_{i\in I}\sum_{j\in I\setminus\{i\}} p_j = \frac{1}{|I|}\cdot\sum_{j\in I}\sum_{i\in I\setminus\{j\}} p_j \\ &= \frac{|I|-1}{|I|}\cdot\sum_{j\in I} p_j \le \frac{n-1}{n}\sum_{j\in I} p_j.\end{aligned} \tag{8}$$

Next it will be shown that

$$\sum_{j\notin I} p_j x_j = \max_{z\in F\cap V}\sum_{j\notin I} p_j z_j \ge \sum_{j\notin I} p_j y_j. \tag{9}$$

To prove (9), let $z \in F\cap V$. Then $z = (\mu_1(B_1),\ldots,\mu_n(B_n))$ for some partition $(B_j)_{j=1}^n$. Since T is maximal, G does not contain any arcs from I to $I^c = \{1,\ldots,n\}\setminus I$. By definition of G this implies that

$$\bigcup_{j\notin I} B_j \subset \bigcup_{j\notin I} A_j. \tag{10}$$

Now

$$\sum_{j\notin I} p_j\mu_j(A_j) \ge \sum_{j\notin I}\sum_{i\notin I} p_j\mu_j(A_j\cap B_i) = \sum_{j\notin I}\sum_{i\notin I} p_i\mu_i(A_j\cap B_i)$$
$$= \sum_{i\notin I}\sum_{j\notin I} p_i\mu_i(A_j\cap B_i) = \sum_{i\notin I} p_i\mu_i\left(B_i\cap\left(\bigcup_{j\notin I} A_j\right)\right) = \sum_{i\notin I} p_i\mu_i(B_i),$$

where the inequality follows since $(B_i)_{i=1}^n$ is a partition, the first equality follows by the Key Equation, the third equality follows since $(A_j)_{j=1}^n$ is a partition, and the last equality follows by (10). This settles the first part of (9). The second part then follows since $y \in \mathrm{co}(F\cap V)$ and any linear functional takes its maximum over a convex compact set in one of the extreme points.

Next, observe that since both x and y are in F, it follows that $\sum_{j=1}^{n} p_j(x_j - y_j) = 0$, so (9) implies $\sum_{j\in I} p_j(x_j - y_j) \le 0$, which added to (8) yields

$$\sum_{j\in I} p_j(x_j - y_j) + \sum_{j\in I\setminus\{1\}} p_j v_j \le \frac{n-1}{n} \sum_{j\in I} p_j. \tag{11}$$

Multiplying (3) by p_1 and subtracting the result from (11) gives

$$\sum_{j\in I\setminus\{1\}} p_j \left(x_j + v_j - y_j - \frac{n-1}{n} \right) \le 0,$$

so at least one term of the above sum must be non-positive. This yields (5), and thereby completes the proof of Lemma 4.2. □

Lemma 4.3. *Let $\mathbf{A} \in \Pi_F$, let $x = \overrightarrow{\mu(\mathbf{A})}$, and let T be a maximal out-tree with root 1 in $G(\mathbf{A}; F)$. If $x_1 > y_1 + (n-1)/n$, then there exists a path $Q \subset T$ containing at least one arc such that*

$$|x'_i - y_i| \le \max\left\{ |x_i - y_i|, \frac{n-1}{n} \right\}, \quad i = 1, \ldots, n, \tag{12}$$

where $x' = \overrightarrow{\mu(M_Q(\mathbf{A}))}$.

Proof. By Lemma 4.2, T contains at least one executable arc, i.e. an arc e satisfying (4). Let $e \in T$ be an executable arc with minimum distance in T to 1. Let P denote the unique path in T from 1 to $\mathrm{ter}(e)$, and let the arcs of P be denoted $f_1, \ldots, f_m$, ordered so that f_{r+1} precedes f_r in P for $r = 1, \ldots, m-1$. In particular, $f_1 = e, \mathrm{ini}(f_m) = 1$, and $\mathrm{ini}(f_r) = \mathrm{ter}(f_{r+1})$ for $r = 1, \ldots, m-1$.

Denote $i_0 = \mathrm{ter}(f_1)$, $i_r = \mathrm{ini}(f_r)$ $(r = 1, \ldots, m)$, and $E_r = \mathrm{at}(f_r)$ $(r = 1, \ldots, m)$ (See Figure 2). Now define

$$m^* := \min\left\{ r \ge 1 : x_{i_r} - \mu_{i_r}(E_r) \ge y_{i_r} - \frac{n-1}{n} \right\}.$$

Note that $m^* \le m$ since, by assumption, $x_1 > y_1 + (n-1)/n$ and $n \ge 2$. Define Q to be the path consisting of the arcs $f_1, f_2, \ldots, f_{m^*}$. Next define $\mathbf{A}' = (A'_i)_{i=1}^n$ by

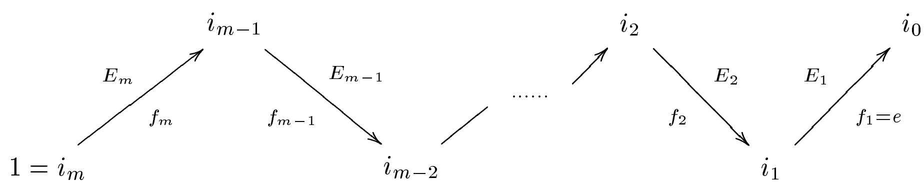

FIGURE 2. Vertices, arcs and atoms of the path P that is used to generate the next partition in the sequence $(\mathbf{A}^k)_{k\in\mathbb{N}}$.

$\mathbf{A}' = M_Q(\mathbf{A})$. More specifically:

$$
\begin{aligned}
&A'_{i_0} = A_{i_0} \cup E_1;\\
&A'_{i_r} = (A_{i_r} \cup E_{r+1})\backslash E_r,\ r = 1, \ldots, m^* - 1;\\
&A'_{i_{m^*}} = A_{i_{m^*}}\backslash E_{m^*};\\
&A'_i = A_i \text{ for } i \notin \{i_0, \ldots, i_{m^*}\}.
\end{aligned} \tag{13}
$$

Let $x' = \overrightarrow{\mu(\mathbf{A}')}$. It follows from (13) that

$$
x'_i = \begin{cases} x_i + \mu_i(E_1), & i = i_0;\\ x_i + \mu_i(E_{r+1}) - \mu_i(E_r), & i = i_r, 1 \le r \le m^* - 1;\\ x_i - \mu_i(E_{m^*}), & i = i_m^*;\\ x_i, & i \notin \{i_0, \ldots, i_{m^*}\}. \end{cases} \tag{14}
$$

For $i \notin \{i_0, \ldots, i_{m^*}\}$, (12) is obvious since $x'_i = x_i$. For i_0 we have $x'_{i_0} \ge x_{i_0}$ since $\mu_{i_0}(E_1) \ge 0$, whereas by the executability of f_1 (cf. (4)),

$$x'_{i_0} = x_{i_0} + \mu_{i_0}(E_1) \le y_{i_0} + \frac{n-1}{n}.$$

Hence (12) holds for $i = i_0$. For $r = 1, \ldots, m^* - 1$, the definition of m^* implies

$$x_{i_r} - \mu_{i_r}(E_r) < y_{i_r} - \frac{n-1}{n},$$

while since f_{r+1} is not executable,

$$x_{i_r} + \mu_{i_r}(E_{r+1}) > y_{i_r} + \frac{n-1}{n}.$$

Using (14) it follows that

$$y_{i_r} - \frac{1}{n} \le y_{i_r} + \frac{n-1}{n} - \mu_{i_r}(E_r) < x'_{i_r} < y_{i_r} - \frac{n-1}{n} + \mu_{i_r}(E_{r+1}) \le y_{i_r} + \frac{1}{n},$$

which yields (12) for $i = i_r$ since $1/n \leq (n-1)/n$.

For i_{m^*}, finally, it follows immediately from (14) that $x'_{i_{m^*}} \leq x_{i_{m^*}}$. On the other hand, by definition of m^*,

$$x'_{i_{m^*}} = x_{i_{m^*}} - \mu_{i_{m^*}}(E_{m^*}) \geq y_{i_{m^*}} - \frac{n-1}{n}.$$

Together, these two facts yield (12) for $i = i_{m*}$.

We have now proved (12) for $i = i_0, i = i_r (1 \leq r \leq m^* - 1), i = i_{m^*}$ and $i \notin \{i_0, \ldots, i_{m^*}\}$. Since this enumeration exhausts all possible values of i, the proof is complete. $\square$

Lemma 4.4. *Let x be a point in $V \cap F$ with $x_1 \geq y_1 - (n-1)/n$. Then there exist $\tau \in \mathbb{N} \cup \{\infty\}$, and sequences $(x^k)_{k=1}^{\tau}$, $(\mathbf{A}^k)_{k=1}^{\tau}$, $(G^k)_{k=1}^{\tau}$, $(T^k)_{k=1}^{\tau}$, and $(Q^k)_{k=1}^{\tau-1}$ with the properties that for all $l \leq \tau$,*

(15) $x^l = \overrightarrow{\mu(\mathbf{A}^l)}$,

(16) $x^l \in F$,

(17) $\|x^l - x^{l+1}\|_\infty \leq 1$ *(for $l < \tau$),*

(18) $|x_i^{l+1} - y_i| \leq \max\{|x_i^l - y_i|, (n-1)/n\}$ *for all i (for $l < \tau$),*

(19) $G^l = G(\mathbf{A}^l; F)$,

(20) *T^l is a maximal out-tree with root 1 in G^l,*

(21) *Each path with root 1 which is in $T^l \cap G^{l+1}$ is also in T^{l+1} (for $l < \tau$),*

(22) *Q^l is a path in T^l containing at least one arc (for $l < \tau$),*

(23) $\mathbf{A}^{l+1} = M_{Q^l}(\mathbf{A}^l)$ *(for $l < \tau$).*

Moreover,

(24) $$\tau := \inf\left\{k \in \mathbb{N} : \mathrm{x}_1^{\mathrm{k}} \leq \mathrm{y}_1 + \frac{\mathrm{n}-1}{\mathrm{n}}\right\}.$$

Proof. Let $x^1 = x$, let $\mathbf{A}^1 = (A_i^1)_{i=1}^n$ be a partition with $\overrightarrow{\mu(\mathbf{A}^1)} = x^1$, let $G^1 = G(\mathbf{A}^1; F)$, and let T^1 be any maximal out-tree with root 1 in G^1.

Next, let $k \in \mathbb{N}$, and suppose that we have constructed $(x^l)_{l=1}^k$, $(\mathbf{A}^l)_{l=1}^k$, $(G^l)_{l=1}^k$, and $(T^l)_{l=1}^k$ satisfying (15), (16), (19) and (20) for $l = 1, \ldots, k$, and (17), (18) and (21)-(23) for $l = 1, \ldots, k-1$. Consider the following two cases:

Case 1, $x_1^k \leq y_1 + (n-1)/n$. Then we can set $\tau := k$ and we are done.

Case 2,

$$x_1^k > y_1 + \frac{n-1}{n}. \tag{25}$$

Then Lemma 4.3 implies the existence of a non-empty path $Q \subset T$ such that (12) holds if $x' = \overrightarrow{\mu(M_Q(\mathbf{A}))}$. Define $Q^k := Q$, $\mathbf{A}^{k+1} = M_Q(\mathbf{A}^k)$, $G^{k+1} := G(\mathbf{A}^{k+1}; F)$, and $x^{k+1} := x'$.

Finally, let T^{k+1} be a maximal out-tree in G^{k+1} that satisfies (21) for $l = k$. (Such a tree exists in view of Lemma 3.1).

We have thus constructed a point x^{k+1}, a partition $\mathbf{A}^{k+1}$, a digraph G^{k+1}, an out-tree T^{k+1} and a path Q^k such that (15), (16), (19) and (20) are satisfied for $l = k+1$, and (17), (18) and (21)-(23) are satisfied for $l = k$ (where (17) is a direct consequence of (14)). By the principle of induction, the proof is complete. $\square$

Lemma 4.5. *Suppose there exist $\tau \in \mathbb{N} \cup \{\infty\}$, and sequences $(x^k)_{k=1}^{\tau}$, $(\mathbf{A}^k)_{k=1}^{\tau}$, $(G^k)_{k=1}^{\tau}$, $(T^k)_{k=1}^{\tau}$, and $(Q^k)_{k=1}^{\tau-1}$ such that (19)-(23) hold for all $l \leq \tau$. Then there exists a sequence of paths $(P^k)_{k=1}^{\tau-1}$ such that for all $k < \tau$*

$$P^k \subset T^k, \tag{26}$$

$$P^k \not\subset G^l \quad \textit{for all } l > k. \tag{27}$$

Proof. For each $k < \tau$, define

e^k : the initial arc of Q^k, i.e. e^k is the arc of Q^k with $\mathrm{ini}(e^k) = \mathrm{ini}(Q^k)$;

P^k : the unique path in T^k from 1 to $\mathrm{ter}(e^k)$.

Note that in particular, $e^k \in P^k$ and $e^k \in Q^k$. The first important fact is that for all $k < \tau$,

$$e^k \notin G^{k+1}, \tag{28}$$

as follows easily by Lemma 3.17 (ii) since $e^k \in Q^k$ and $\mathbf{A}^{k+1} = M_{Q^k}(\mathbf{A}^k)$ by (23). Let ρ_k denote the graph-theoretic distance between 1 and $\mathrm{ini}(e^k)$ in T^k. The proof of Lemma 4.5 will use induction on ρ_k.

First suppose that $\rho_k = 0$. It will be shown that

$$e^k \notin G^l,\ l > k. \tag{29}$$

This will be done by induction on l. Note first that $e^k \notin G^{k+1}$ by (28). Suppose now that $e^k \notin G^l$ for some $l > k$. Since $\rho_k = 0$, it follows that $\mathrm{ini}(e^k) = 1$, so Lemma 3.12 implies $\mathrm{at}(e^k) \not\subset A_1^l$. Hence $\mathrm{at}(e^k) \subset A_i^l$ for some $i \neq 1$. Since T^l has root 1, there is no arc in T^l from i to 1, so in particular, there is no such arc in Q^l. By Lemma 3.17 (iii) it follows that $\mathrm{at}(e^k) \not\subset A_1^{l+1} = A_{\mathrm{ini}(e^k)}^{l+1}$, which by lemma 3.12 and (19) implies $e^k \notin G^{l+1}$. This completes the proof of (29).

Since $e^k \in P^k$, (29) implies that $P^k \not\subset G^l$ for $l > k$, which proves (27) in case $\rho_k = 0$.

Now let $m \in \mathbb{N}$, and suppose (27) holds for all $k < \tau$ with $\rho_k \leq m-1$. It will now be shown that (27) holds for all $k < \tau$ with $\rho_k \leq m$, which will complete the proof of Lemma 4.5 by induction.

Fix any $k < \tau$ so that $\rho_k \leq m$, and define

$$k' := \inf\{l \geq k+1 : e^l \text{ is related to some arc } e \in P^k \backslash e^k\},$$

where $\inf(\emptyset) \equiv \infty$. Since $e^k \in P^k$, the proof will be complete once it has been shown that

$$P^k \not\subset G^l \text{ for all } l > k', \tag{30}$$

$$e^k \notin G^l \text{ for } l = k+1, \ldots, k'. \tag{31}$$

Here and in the sequel, $\{k+1, \ldots, k'\}$ is to be read as $\{k+1, k+2, \ldots\}$ in case $k' = \infty$.

First we establish the following fact:

$$P^k \backslash e^k \subset T^l,\ l = k, \ldots, k'. \tag{32}$$

To see (32), note first that $P^k \backslash e^k \subset P^k \subset T^k$ by definition of P^k. Now suppose that $P^k \backslash e^k \subset T^l$ for some $l \in \{k, \ldots, k'-1\}$, and let e be an arc of $P^k \backslash e^k$.

Suppose, by way of contradiction, that e is related to some arc of Q^l; then $\mathrm{ini}(e) \in Q^l$. Since $P^k \backslash e^k \subset T^l$ by the induction hypothesis, $P^k \backslash e^k$ contains the unique path R in T^l from 1 to $\mathrm{ini}(e)$. Since e^l is the initial arc of Q^l, and since $Q^l \subset T^l$, we have either $e^l \in R$, or $\mathrm{ini}(e^l) = \mathrm{ini}(e)$. In both cases e^l is related to an arc of $P^k \backslash e^k$. But this is a contradiction since $l < k'$.

Thus e is not related to any arc of Q^l, so Lemma 3.17 (iv) implies $e \in G^{l+1}$. Since e was arbitrary, it follows that $P^k \backslash e^k \subset G^{l+1}$. Hence $P^k \backslash e^k \subset T^{l+1}$ in view of (21). This proves (32) by induction.

In particular, (32) implies

$$P^k \backslash e^k \subset T^{k'}. \tag{33}$$

By definition of k', $e^{k'}$ is related to some arc $e \in P^k \backslash e^k$, hence $\mathrm{ini}(e^{k'}) = \mathrm{ini}(e) \in P^k \backslash e^k$. Since $T^{k'}$ is a tree that contains both $P^k \backslash e^k$ and $P^{k'}$, the paths $P^k \backslash e^k$ and $P^{k'}$ must overlap up to the common vertex $\mathrm{ini}(e^{k'})$, i.e.

$$P^{k'} \backslash e^{k'} \subset P^k \backslash e^k. \tag{34}$$

Moreover, $\mathrm{ini}(e^{k'}) = \mathrm{ini}(e) \neq \mathrm{ini}(e^k)$ and therefore, since $\mathrm{ini}(e^k)$ and $\mathrm{ini}(e^{k'})$ lie on the same path $P^k \backslash e^k$, it follows from (34) that $\rho_{k'} \leq \rho_k - 1 \leq m - 1$, and the induction hypothesis implies

$$P^{k'} \not\subset G^l \text{ for all } l > k'. \tag{35}$$

Now for any $l > k'$ there are two possibilities: either $P^{k'} \backslash e^{k'} \not\subset G^l$, which by (34) implies $P^k \not\subset G^l$; or $e^{k'} \notin G^l$, which by Lemma 3.14 implies $e \notin G^l$ (recall that e and $e^{k'}$ are related), so certainly $P^k \not\subset G^l$ because $e \in P^k$. This completes the proof of (30).

The proof of (31) is similar to that of (29), but this time the argument is slightly more subtle. We will again use induction on l. Note first that $e^k \notin G^{k+1}$ by (28). Suppose $e^k \notin G^l$ for some $l \in \{k+1, \ldots, k'-1\}$. For brevity, write $E := \mathrm{at}(e^k)$ and $j := \mathrm{ini}(e^k)$. By the induction hypothesis and Lemma 3.12 we have $E \not\subset A_j^l$, so $E \subset A_i^l$ for some $i \neq j$, since $(A_i^l)_{i=1}^n$ is a partition. By (32), T^l contains the arc $e \in P^k \backslash e^k$ with $\mathrm{ter}(e) = j$. Since both e and e^k are in G^k, and $\mathrm{ini}(e) \neq \mathrm{ter}(e) = j =$

$\mathrm{ini}(e^k)$, it follows by a double application of Lemma 3.12 that $\mathrm{at}(e) \neq \mathrm{at}(e^k) = E$. Therefore, $e \neq (i,j,E)$. Since T^l is an out-tree, j has indegree 1 in T^l, i.e. e is the only arc in T^l directed to j. This implies that $(i,j,E) \notin T^l$. In particular, $(i,j,E) \notin Q^l$, so by Lemma 3.17 (iii), $E \not\subset A_j^{l+1}$. A final application of Lemma 3.12 yields $e^k \notin G^{l+1}$, which completes the proof of (31). $\square$

Proof of Proposition 4.1. Fix a face F of $\mathrm{co}(V)$ and a point $y \in F$. It will be shown that there exists a point $x \in V \cap F$ satisfying

$$|x_i - y_i| \leq \frac{n-1}{n}, \qquad i = 1, \ldots, n.$$

This will be done inductively. Let $m \in \{1, \ldots, n\}$, and suppose there exists a point $x \in V \cap F$ such that

$$|x_i - y_i| \leq \frac{n-1}{n} \qquad \text{for all } i \leq m-1. \tag{36}$$

(Note that if $m = 1$ then (36) is vacuous). We will now construct a finite sequence of points $(x^k)_{k=1}^{\tau}$ in $V \cap F$ such that

$$|x_i^\tau - y_i| \leq \max\left\{|x_i - y_i|, \frac{n-1}{n}\right\} \qquad \forall i \leq n, \quad \text{and} \tag{37}$$

$$|x_m^\tau - y_m| \leq \frac{n-1}{n}. \tag{38}$$

Then the point $x' = x^\tau$ will satisfy

$$|x_i' - y_i| \leq \frac{n-1}{n} \qquad \text{for all } i \leq m, \tag{39}$$

and the proof will be complete by induction.

Suppose that $|x_m - y_m| > (n-1)/n$ (otherwise $x' = x$ satisfies (39)). By reordering the coordinates of x if necessary, we may assume that $m = 1$. Then either

$$x_1 > y_1 + \frac{n-1}{n}, \tag{40}$$

or $x_1 < y_1 - (n-1)/n$. Only the first case will be considered here; it is left as an exercise for the interested reader to verify that the second case can be treated completely analogously.

Suppose that (40) holds. By Lemma 4.4 there exist $\tau \in \mathbb{N} \cup \{\infty\}$, and sequences $(x^k)_{k=1}^{\tau}$, $(\mathbf{A}^k)_{k=1}^{\tau}$, $(G^k)_{k=1}^{\tau}$, $(T^k)_{k=1}^{\tau}$, and $(Q^k)_{k=1}^{\tau-1}$ such that (15)-(23) hold for all $l \le \tau$. We shall now show that $\tau < \infty$.

Notice that Lemma 4.5 implies the existence of a sequence $(P^k)_{k=1}^{\tau-1}$ satisfying (26) and (27) for all $k < \tau$. Since $T^k \subset G^k$, (26) implies that $P^k \subset G^k$, hence using (27) we see that $G^k \neq G^l$ for all $k < \tau$ and $l > k$. Hence all G^k's $(k < \tau)$ are different. This means that all partitions $\mathbf{A}^k (k < \tau)$ are different. Since there are only finitely many different partitions of n atoms, this means that $\tau < \infty$.

It now follows from (24) that $x_1^\tau \le y_1 + (n-1)/n$. On the other hand, (24) also implies that $x_1^{\tau-1} > y_1 + (n-1)/n$, so using (17) and the fact that $n \ge 2$ if follows that $x_1^\tau \ge x_1^{\tau-1} - 1 > y_1 - 1/n \ge y_1 - (n-1)/n$. This proves (38) (recall that $m = 1$), while (37) follows immediately from (18) using induction. □

Proof of Theorem 2.5. First assume that $\alpha = 1$. Fix $\vec{\mu} \in \mathcal{P}_n(1)$ and $\varepsilon > 0$. By Proposition 3.4, there is a purely atomic measure $\vec{\mu}_0 \in \mathcal{P}_n(1)$ with finitely many atoms such that

$$D_\infty(\mathcal{PR}(\vec{\mu})) \le D_\infty(\mathcal{PR}(\vec{\mu}_0)) + \varepsilon, \tag{41}$$

and using Proposition 3.8 we can assume w.l.o.g. that the number of atoms of $\vec{\mu}_0$ is at most n. For brevity, write $V := \mathcal{PR}(\vec{\mu}_0)$. Fix a point $y \in \mathrm{co}(V)$, and consider the following two cases.

Case 1,

$$y_i \le \frac{n-1}{n} \quad \text{for some } i.$$

There is a point $y' \in \mathrm{Bd}(\mathrm{co}(V))$ with $y'_i \le y_i$ and $y'_j = y_j$ for all $j \neq i$. Let F be a face of $\mathrm{co}(V)$ containing y'. By Proposition 4.1 there is a point $x \in V \cap F$ such that $\|x - y'\|_\infty \le (n-1)/n$. Now for $j \neq i$ we have $|x_j - y_j| = |x_j - y'_j| \le (n-1)/n$; but also $|x_i - y_i| \le (n-1)/n$, since

$$-\frac{n-1}{n} \le y'_i - x_i \le y_i - x_i \le y_i \le \frac{n-1}{n}.$$

Hence $\|x - y\|_\infty \le (n-1)/n$.

Case 2,

$$y_i > \frac{n-1}{n} \quad \text{for all } i. \tag{42}$$

Let y' be a point in $\mathrm{Bd}(\mathrm{co}(V))$ with $y'_j \geq y_j$ for all j, and let F be a face of $\mathrm{co}(V)$ containing y'. Again Proposition 4.1 ensures the existence of a point $x \in V \cap F$ with $\|x - y'\|_\infty \leq (n-1)/n$. Thus

$$x_j \geq y'_j - \frac{n-1}{n} \geq y_j - \frac{n-1}{n} > 0 \quad \text{for all } j \in \{1, \ldots, n\}. \tag{43}$$

Let $(A_j)_{j=1}^n$ be a partition such that $x_j = \mu_{0,j}(A_j), j = 1, \ldots, n$. Then (43) implies that each A_j $(j = 1, \ldots, n)$ is non-empty. But since there are at most n atoms, this means that each A_j must contain exactly one atom, and therefore $x_j = \mu_{0,j}(A_j) \leq 1$ for all j. Hence

$$-\frac{n-1}{n} \leq x_j - y_j < \frac{1}{n} \leq \frac{n-1}{n} \quad \text{for all } j,$$

where the first inequality follows by (43), the second by (42) and $x_j \leq 1$, and the last since $n \geq 2$. Hence $\|x - y\|_\infty \leq (n-1)/n$.

It appears that in both cases there is a point $x \in V \cap F$ with $\|x - y\|_\infty \leq (n-1)/n$. Thus we have proved that

$$D_\infty(\mathcal{PR}(\vec{\mu}_0)) \leq \frac{n-1}{n}.$$

Together with (41), this completes the proof of Theorem 2.5 for $\alpha = 1$, since ε was arbitrary. The general case $\alpha > 0$ now follows easily by rescaling, and the case $\alpha = 0$ follows by continuity. $\square$

5. Applications to optimal-partitioning

The objective of this section is to show how Theorem 2.5 can be used to obtain optimal-partitioning inequalities for measures with atoms. The idea is that many well-known partitioning inequalities for atomless measures (e.g. [9, 11, 12]) are proved using the convexity theorem of Dvoretzky, Wald and Wolfowitz (cf. Proposition 2.3), so using Theorem 2.5 instead yields analogous inequalities for the more general (atomic) case. To illustrate this we will generalize three well-known partitioning inequalities to measures with atoms. The first result is an extension of classical 'cake-cutting' results; the other two are generalizations of inequalities of Elton, Hill and Kertz [5], and Hill [9], and have interesting consequences for the existence of fair divisions.

The overall framework is a measurable space $(\Omega, \mathcal{F})$, together with probability measures $\mu_1, \ldots, \mu_n$. In the classical 'cake-cutting' problem (see, e.g. Dubins and Spanier [2]), where $\mu_1, \ldots, \mu_n$ are assumed to be atomless, the existence of a measurable partition $(A_i)_{i=1}^n$ can be shown such that

$$\min_{i \le n} \mu_i(A_i) \ge \frac{1}{n}. \tag{44}$$

A partition satisfying (44) is usually called a 'fair division'.

If the measures $\mu_1, \ldots, \mu_n$ have atoms, then fair divisions need not exist in general, but an application of Theorem 2.5 gives the following *approximate* fair-division result:

Corollary 5.1. *If $\mu_i(E) \le \alpha$ for each i and each atom E of any μ_j, then there exists a measurable partition $(A_i)_{i=1}^n$ of Ω such that*

$$\min_{i \le n} \mu_i(A_i) \ge \frac{1}{n} - \frac{n-1}{n}\alpha.$$

Proof. It is easily seen that $\mathcal{PR}(\vec{\mu})$ contains the n unit vectors $e_1, \ldots, e_n$. Hence $\mathrm{co}(\mathcal{PR}(\vec{\mu}))$ contains the point $y = (1/n, \ldots, 1/n)$, so by Theorem 2.5 there is an $x \in \mathcal{PR}(\vec{\mu})$ such that $\|x - y\|_\infty \le \alpha(n-1)/n$. This means that each coordinate of x is at least $1/n - \alpha(n-1)/n$. $\square$

Note that a stronger lower bound for probability measures was given by Hill [10], though the bound of Corollary 5.1 coincides with Hill's bound for certain values of α. The advantage of our approach is that it can be used for a much larger class of partitioning problems (for example, it does not require the measures to be probability measures).

Without further assumptions, the constant $1/n$ in (44) is best possible, but if $\mu_i \neq \mu_j$ for some $i \neq j$, then there exists a partition satisfying (44) with *strict* inequality; a result by Dubins and Spanier. Quantitative generalizations of this result were proved by Elton, Hill and Kertz [4] and Hill [9], who gave sharp lower bounds for the optimal-partitioning constant

$$C = \sup\{\min_{i \leq n} \mu_i(A_i) | (A_i)_{i=1}^n \text{ is a measurable partition of } \Omega\}$$

in terms of the total masses of the supremum, resp. infimum of the measures. The next result generalizes their inequalities to measures with atoms. First define

$$\bigvee_{i=1}^n \mu_i : \text{ the smallest measure dominating each } \mu_i \ (i = 1, \ldots, n),$$

$$\bigwedge_{i=1}^n \mu_i : \text{ the largest measure dominated by each } \mu_i \ (i = 1, \ldots, n),$$

and let $M := \left(\bigvee_{i=1}^n \mu_i\right)(\Omega)$, and $m := \left(\bigwedge_{i=1}^n \mu_i\right)(\Omega)$.

Theorem 5.2. *If $\mu_i(E) \leq \alpha$ for each i and each atom E of any μ_j, then*

(i) $C \geq (n - M + 1)^{-1} - \alpha(n-1)/n$,

(ii) $C \geq (n + m - 1)^{-1} - \alpha(n-1)/n$.

Proof. The proof of (i) proceeds as in Legut [12], using Theorem 2.5 where [12] applies the convexity theorem. In a similar way, the proof of (ii) proceeds as in Hill [9]. □

As a consequence of Theorem 5.2, we get the following two sufficient conditions for the existence of a fair division in the sense of (44).

Corollary 5.3. *Suppose that either of the following holds for $i = 1, \ldots, n$.*

(i) $\mu_i(E) \leq (M-1)(n-1)^{-1}(n-M+1)^{-1}$ *for each atom E of μ_i;*

(ii) $\mu_i(E) \leq (m+1)(n-1)^{-1}(n+m-1)^{-1}$ *for each atom E of μ_i.*

Then there exists a measurable partition $(A_i)_{i=1}^n$ satisfying (44).

Proof. Immediate from Theorem 5.2. □

Example 5.4. *If $n = 3$ and $M = 2$, then if all the atoms of $\mu_1, \ldots, \mu_n$ have mass $1/4$ or less there exists a fair division in the sense of* (44).

Acknowledgements. The author is grateful to Prof. Ted Hill for his critical comments on earlier versions of this paper; to the referee, whose comments led to an improvement in the presentation; and to Prof. Jan van Mill for the opportunity to present these ideas in his Topology Seminar.

References

1. Artstein, Z. (1980). Discrete and continuous bang-bang and facial spaces or: look for the extreme points. *SIAM J. Review* **22** No. 2, 172-185.
2. Dubins, L.E. and Spanier, E.H. (1961). How to cut a cake fairly. *Amer. Math. Monthly* **68**, 1-17.
3. Dvoretzky, A., Wald, A., and Wolfowitz, J. (1951). Relations among certain ranges of vector measures. *Pacific J. Math.* **1**, 59-74.
4. Elton, J. and Hill, T.P. (1987). A generalization of Lyapounov's convexity theorem to measures with atoms. *Proc. Amer. Math. Soc.* **296**, 297-304.
5. Elton, J., Hill, T.P. and Kertz, R.P. (1986). Optimal-partitioning inequalities for nonatomic probability measures. *Trans. Amer. Math. Soc.* **296**, 703-725.
6. Gouweleeuw, J.M. (1995). A characterization of measures with convex range. *Proc. London Math. Soc.* (3) **70**, 336-362.
7. Grünbaum, B. (1967). *Convex Polytopes.* Wiley (Interscience), New York.
8. Harary, F., Norman, R.Z. and Cartwright, D. (1965). *Structural models: an introduction to the theory of directed graphs.* Wiley, New York.
9. Hill, T. P. (1987). A sharp partitioning-inequality for non-atomic probability measures based on the mass of the infimum of the measures. *Prob. Th. and Rel. Fields* **75**, 143-147.
10. Hill, T. P. (1987). Partitioning general probability measures. *Ann. Prob.* **15**, 804-813.
11. Hill, T. P. and Tong, Y. L. (1989). Optimal-partitioning inequalities in classification and multi-hypotheses testing. *Ann. Stat.* **17**, 1325-1334.
12. Legut, J. (1988). Inequalities for α-optimal partitioning of a measurable space. *Proc. Amer. Math. Soc.* **104**, 1249-1251.
13. Lyapounov, A. (1940). Sur les fonctions-vecteurs complètement additives. *Bull. Acad. Sci. URSS* **4**, 465-478.