DTU Library

Relative volume of comparable pairs under semigroup majorization

Cunden, Fabio Deelan; Czartowski, Jakub; Gramegna, Giovanni; de Oliveira Junior, A.

Published in:
Letters in Mathematical Physics

Link to article, DOI:
10.1007/s11005-025-01968-3

Publication date:
2025

Document Version
Publisher's PDF, also known as Version of record

Letters in Mathematical Physics (2025) 115:79
https://doi.org/10.1007/s11005-025-01968-3

Relative volume of comparable pairs under semigroup majorization

Fabio Deelan Cunden[1,2] · Jakub Czartowski[3,4,5] · Giovanni Gramegna[2,6] · A. de Oliveira Junior[7]

Received: 16 December 2024 / Revised: 6 June 2025 / Accepted: 18 June 2025 /
Published online: 14 July 2025

Abstract

Any semigroup $\mathcal{S}$ of stochastic matrices induces a semigroup majorization relation $\prec^{\mathcal{S}}$ on the set Δ_{n-1} of probability n-vectors. Pick X, Y at random in Δ_{n-1}: what is the probability that X and Y are comparable under $\prec^{\mathcal{S}}$? We review recent asymptotic ($n \to \infty$) results and conjectures in the case of *majorization* relation (when $\mathcal{S}$ is the set of doubly stochastic matrices), discuss natural generalisations, and prove a new asymptotic result in the case of majorization, and new exact finite-n formulae in the case of *UT-majorization* relation, i.e. when $\mathcal{S}$ is the set of upper-triangular stochastic matrices.

Keywords Semigroup majorization · Stochastic matrices · Probability simplex · Resource theories · Asymptotic comparability

Mathematics Subject Classification 60C05 · 05A20 · 15B51 · 52B11 · 81P45

1 Introduction

The subject of combinatorial and algebraic aspects of partially ordered sets has received much attention in recent decades in quantum physics, with an explosion of works in the field, since the (re)discovery of the interplay between certain resource theories

The research of FDC and GG was supported by Gruppo Nazionale di Fisica Matematica GNFM-INdAM and by Istituto Nazionale di Fisica Nucleare INFN through the project QUANTUM. FDC acknowledges the support from PRIN2022 project 2022TEB52W-PE1- ‘The charm of integrability: from nonlinear waves to random matrices’ as FDC acknowledges the support from PRIN2022 project 2022TEB52W-PE1- ‘The charm of integrability: from nonlinear waves to random matrices’, and from PNRR MUR project CN00000013 ‘Italian National Centre on HPC, Big Data and Quantum Computing’. AOJ acknowledges financial support from VILLUM FONDEN through a research Grant (40864) and EU Horizon Europe project QSNP (Grant Agreement No. 101114043). JCz is supported by the start-up grant of the Nanyang Assistant Professorship at the Nanyang Technological University in Singapore, awarded to Nelly Ng.

Extended author information available on the last page of the article

and order relations, most notably between entanglement theory and majorization [27]. See [8, 10, 14, 32] and references therein. The basic idea of a (quantum) resource theory is to study information processing under a restricted set of operations, called *free*. Thus, a resource theory leads to a partial ordering of states, ranking them from harder to easier to prepare by means of free operations. This notion of partial order, introduced in physics by Uhlmann [4], has deep mathematical roots dating back to Schur, Hardy, Littlewood, Pólya, and others.

Consider a finite system whose states can be described by n-dimensional probability vectors $x = (x_1, \ldots, x_n)$, $x_i \geq 0$, $\sum_i x_i = 1$. These states belong to the probability simplex Δ_{n-1} referred to as the **state space**. Stochastic matrices represent the most general linear transformations between states in Δ_{n-1}. Formally, a resource theory [9] can be constructed by declaring that the set of **free operations** between states in Δ_{n-1} is a given proper semigroup $\mathcal{S}$ of stochastic matrices. The set of free operations $\mathcal{S}$ leads to an order relation in the state space: we write $x \prec^{\mathcal{S}} y$ whenever $x = yP$ for some $P \in \mathcal{S}$, that is if the state y can be converted into x by means of a free operation. In this paper, we adopt the right (row) stochastic matrix convention, where stochastic matrices act on the right on row probability vectors. This is standard in probability theory, where stochastic matrices are also called Markov matrices (see Section 2). The standard majorization relation [23] is obtained by choosing $\mathcal{S}$ as the semigroup of doubly stochastic matrices. An equivalent characterisation of majorization (and other semigroup majorization) is in terms of monotonicity of family of functions that deserve the name of 'generalised entropies' [31]. In some situations, the family of generalised entropies can be generated by nonnegative linear combinations of a finite family of 'monotone' functions.

Given two states x, y, for any free operations $P \in \mathcal{S}$, one can certainly decompose yP, as a convex combination $px + (1-p)z$ for some $z \in \Delta_{n-1}$ (with p that depends on P). The trivial case $p = 1$ corresponds to the case $yP = x$, i.e. $x \prec^{\mathcal{S}} y$. For general x, y, it is natural to ask what is the largest value of p, denoted with $\Pi_{\mathcal{S}}(x, y)$, that can be achieved in a transformation $yP = px + (1-p)z$, as P runs over all free transformations. In the context of resource theories, $\Pi_{\mathcal{S}}$ can be thought of as the *maximal success probability of conversion* of y into x [10, 28, 39, 40]. This happens when a quantum measurement is used in the conversion protocol, which allows one to extract x from the convex combination $px + (1-p)z$ conditionally on certain outcomes of the measurement, whose occurrence probability sum to p. However, it is worth mentioning that measurements may not be free operations in a given quantum resource theory [3].

It is clear that in a resource theory, a state y is 'useful' (or 'more useful than others') if from y it is possible to reach 'many' other states by free operations. For example, in entanglement theory maximally entangled states of a bipartite system can be converted to *every* other state using local operations and classical communication only, and this is why such special states (like Bell states), sometimes referred to as 'golden state', are key ingredients in applications. But such golden states are very special.

We therefore address the following question: given a fixed resource theory, how likely is it that two 'generic' states can be connected via a free operation in $\mathcal{S}$? More precisely, pick X and Y at random according to a (natural) probability measure on the state space Δ_{n-1}: what is the probability that $X \prec^S Y$? Our initial intuition is that in the limit of large dimensionality $n \to \infty$, typical pairs of states are so uncorrelated, that they cannot be converted to each other via free operations and hence are incomparable. This *'comparability is atypical'* statement was first put forward by Nielsen [26] in quantum theory and was only recently proved in [7] and in [18] for the resource theories of quantum coherence and quantum entanglement, respectively. Given that exact conversion between generic states is typically impossible to achieve by means of free operations, one can consider transformations that succeed with some probability p (not necessarily equal to 1) at transforming the initial state into the desired final state. If we allow nondeterministic transformations (i.e. some probability of failure), what is the typical maximal probability of successful transformation $\Pi_{\mathcal{S}}(X, Y)$ for random states X and Y?

1.1 Majorization

Let $\prec$ be the majorization relation. This is the semigroup majorization corresponding to the set of doubly stochastic matrices (precise definitions are given later). Let also $\Pi(x, y)$ be the largest value $p \in [0, 1]$ such that $px + (1-p)z \prec y$ for some $z \in \Delta_{n-1}$.

Theorem A *Pick X, Y independently and uniformly at random in Δ_{n-1}. Then,*

$$\lim_{n\to\infty} P\,(X \prec Y) = 0, \tag{1.1}$$

and $\Pi(X, Y)$ converges in probability of 1: for all $\epsilon > 0$,

$$\lim_{n\to\infty} P(\Pi(X, Y) < 1 - \epsilon) = 0. \tag{1.2}$$

The limit (1.1) is a restatement of a claim proved in [7], while the convergence (1.2) is a new result (Theorem 4.2). Here is an informal rephrasing: the proportion of pairs of states (x, y) such that x is majorized by y tends to 0 for large n; nevertheless, in any ϵ-neighbourhood of x, if n is large, there is a vector x_ϵ that is majorized by y. The result may, at first, feel paradoxical. A way to reconcile these facts is to observe that (1.1)-(1.2) show that exchanging limits does not work here: for fixed $\epsilon > 0$, the $n \to \infty$ limit yields 1, but if one first takes $\epsilon \to 0$ and then $n \to \infty$, then the probability becomes zero.

1.2 Upper-triangular majorization

Let $\prec^{\mathrm{UT}}$ be the semigroup majorization relation corresponding to the semigroup $\mathcal{S}^{\mathrm{UT}}$ of upper-triangular stochastic matrices (UT-majorization). Let also $\Pi_{\mathrm{UT}}(x, y)$ be the largest value $p \in [0, 1]$ such that $px + (1-p)z \prec^{UT} y$ for some $z \in \Delta_{n-1}$. In physics, the semigroup $\mathcal{S}^{\mathrm{UT}}$ arises in the zero-temperature limit of thermal operations [25].

Theorem B *Pick X, Y independently and uniformly at random in Δ_{n-1}. Then,*

$$P\left(X \prec^{UT} Y\right) = \frac{1}{n}. \tag{1.3}$$

Moreover, the distribution function of $\Pi_{UT}(X, Y)$ is

$$P(\Pi_{UT}(X, Y) \leq t) = \begin{cases} 0 & \text{if } t \leq 0, \\ \left(1 - \frac{1}{n}\right) t & \text{if } 0 < t < 1, \\ 1 & \text{if } t \geq 1. \end{cases} \tag{1.4}$$

A more precise statement of (1.3) (for more general measures μ_n on Δ_{n-1}) and (1.4) is given in Theorems 3.3 and 4.4 later on. Equation (1.3) implies the claimed 'comparability is atypical' statement: the probability that two random states can be connected by a upper-triangular stochastic matrix vanishes in the limit $n \to \infty$. Equation (1.4) addresses the question of 'nondeterministic convertibility'.

We stress that Theorem B is an exact finite-n result, and in this sense the model of comparability of random vectors with respect to UT-majorization can be considered 'exactly solvable': explicit calculations are possible and might give insights to understand the asymptotics of other, non-solvable models.

1.3 Outline of the paper

In the next section, we give some background on partial orders, semigroup majorization, and orderings induced by function sets. In Section 3, we formulate the problem, recall known asymptotic results for majorization (Theorems 3.1 and 3.2) and state the first main theorem on UT-majorization, Theorem 3.3, which gives an exact universal formula for the probability that two random states X and Y are comparable with respect to UT-majorization. In Section 4, we consider the maximal success probability of conversion of random states X and Y. In the case of majorization and UT-majorization relations, we first present and prove an explicit formula for $\Pi_{\mathcal{S}}(\cdot, \cdot)$ involving ratios of monotone functions (Propositions 4.1 and 4.3). In the majorization relation, one expects the existence of a limit in distribution of the random variable $\Pi_{\mathcal{S}}(X, Y)$. We prove that this is indeed the case and we find the limit distribution to be concentrated at 1 (Theorem 4.2). For UT-majorization, we find the exact finite-n distribution function (Theorem 4.4), which, in fact, has a nontrivial limit. The final section contains the proofs of the main results. We included some technical results in the appendices A (Bolshev's recursion) and B (stick-breaking representation of the Dirichlet distribution).

2 Background

We begin by recalling some terminology. The reader is referred to the monograph of Marshall, Olkin, and Arnold [23] for further details.

2.1 Order relations

A **partial order** is a pair $(\mathcal{A}, \sqsubseteq)$, where $\mathcal{A}$ is a set and $\sqsubseteq$ a binary relation on $\mathcal{A}$ satisfying

(1) **reflexivity:** for all $x \in \mathcal{A}$, $x \sqsubseteq x$.
(2) **transitivity:** for all $x, y, z \in \mathcal{A}$, if $x \sqsubseteq y$ and $y \sqsubseteq z$, then $x \sqsubseteq z$.
(3) **antisymmetry:** for all $x, y \in \mathcal{A}$, if $x \sqsubseteq y$ and $y \sqsubseteq x$, then $x = y$.

A **preorder** is a pair $(\mathcal{A}, \sqsubseteq)$ satisfying the first two conditions. If $x \sqsubseteq y$ we say that 'x is less than y'. If it exists, a **bottom element** of a partial order is an element $\perp \in \mathcal{A}$ satisfying $\perp \sqsubseteq x$ for all $x \in \mathcal{A}$. If it exists, a **top element** of a partial order, an element $\top \in \mathcal{A}$ satisfying $x \sqsubseteq \top$ for all $x \in \mathcal{A}$. If neither $x \sqsubseteq y$ nor $y \sqsubseteq x$, we say that x and y are **incomparable**. A **total order** is a partial order $(\mathcal{A}, \sqsubseteq)$ in which for every pair of elements x, y either $x \sqsubseteq y$ or $y \sqsubseteq x$ holds. The **dual** (or converse) relation $(\mathcal{A}, \sqsubseteq^*)$ is defined by setting $x \sqsubseteq^* y$ if and only if $y \sqsubseteq x$.

For a given $y \in \mathcal{A}$, the set of all **antecedents** of y in the relation $\sqsubseteq$ is denoted by $\gamma(y, \sqsubseteq)$. Thus,

$$\gamma(y, \sqsubseteq) = \{x \in \mathcal{A} \colon x \sqsubseteq y\}. \tag{2.1}$$

The **graph** of the relation $\sqsubseteq$ on $\mathcal{A}$ will be denoted by $\Gamma(\mathcal{A}, \sqsubseteq)$. Thus,

$$\Gamma(\mathcal{A}, \sqsubseteq) = \{(x, y) \in \mathcal{A} \times \mathcal{A} \colon x \in \gamma(y, \sqsubseteq)\} = \{(x, y) \in \mathcal{A} \times \mathcal{A} \colon x \sqsubseteq y\}. \tag{2.2}$$

We will simply write Γ when $(\mathcal{A}, \sqsubseteq)$ is clear from the context.

2.2 Semigroup majorization

For $x \in \mathbb{R}^n$, we denote by $x^\downarrow$ the permutation of x with the components arranged in nonincreasing order, i.e. $x_1^\downarrow \geq x_2^\downarrow \geq \cdots \geq x_n^\downarrow$. For $x, y \in \mathbb{R}^n$, we say that x is *majorized* by y (or y majorizes x) and write $x \prec y$ if

$$\begin{cases} \displaystyle\sum_{i=1}^{k} x_i^\downarrow \leq \sum_{i=1}^{k} y_i^\downarrow, \quad \text{for } k = 1, \ldots, n-1 \\ \displaystyle\sum_{i=1}^{n} x_i^\downarrow = \sum_{i=1}^{n} y_i^\downarrow. \end{cases} \tag{2.3}$$

If the equality in (2.3) is replaced by a corresponding inequality, we say that x is *weakly (sub)majorized* by y and write $x \prec_w y$, i.e. for $x, y \in \mathbb{R}^n$:

$$x \prec_w y \quad \text{if} \quad \sum_{i=1}^{k} x_i^{\downarrow} \leq \sum_{i=1}^{k} y_i^{\downarrow}, \quad k = 1, \ldots, n. \tag{2.4}$$

It can be seen [21] that $x \prec y$ if and only if, for $j = 1, \ldots, n$,

$$\begin{aligned} &\mathrm{conv}\{x_{i_1} + \cdots + x_{i_j} : 1 \leq i_1 < \cdots < i_j \leq n\} \\ &\quad \subset \mathrm{conv}\{y_{i_1} + \cdots + y_{i_j} : 1 \leq i_1 < \cdots < i_j \leq n\}, \end{aligned} \tag{2.5}$$

where $\mathrm{conv}(A)$ denotes the convex hull of a set A (the smallest convex set containing A).

By a celebrated result by Hardy, Polya and Littlewood [15], $x \prec y$ if and only if $x = yD$ for some doubly stochastic matrix D. This relation is a preorder because the set of $n \times n$ doubly stochastic matrices contains the identity and is closed under multiplication. Other preorders are obtained if the set of doubly stochastic matrices are replaced by another semigroup $\mathcal{S}$ of matrices with identity [23].

Let $\mathcal{A}$ be a subset of $\mathbb{R}^n$, and let $\mathcal{S}$ be a semigroup of linear transformations (matrices) mapping $\mathcal{A}$ to $\mathcal{A}$. A vector $x \in \mathcal{A}$ is said to be **semigroup majorized** by y, written $x \prec^{\mathcal{S}} y$, if $x = yS$ for some $S \in \mathcal{S}$. From the semigroup property of $\mathcal{S}$, it follows that $\left(\mathcal{A}, \prec^{\mathcal{S}}\right)$ is a preorder, but not a partial order in general. See [19, 36, 37] for further details.

Recall that a square matrix P is row (sub)stochastic if all entries are nonnegative and the row sums are (less than or) equal to 1. If we denote the row vector $e = (1, \ldots, 1)$, we have that a square matrix P with nonnegative entries is row (sub)stochastic if $Pe' = e'$ ($Pe' \leq e'$), where e' denotes the transposed vector of e and $\leq$ is to be understood component-wise.

We define the following types of matrices with nonnegative entries (mapping $\mathcal{A} = \mathbb{R}_+^n$ to itself):

- **Doubly stochastic matrices (DSM)**:
  row stochastic matrices D with column sums equal to 1, i.e. $eD = e$.
- **Weakly doubly stochastic matrices (wDSM)**:
  row substochastic matrices D with column sums at most 1, i.e. $eD \leq e$.
- **Upper-triangular stochastic matrices (UTSM)**:
  row stochastic matrices $S = (s_{ij})$ with $s_{ij} = 0$ whenever $i > j$.
- **Weakly upper triangular stochastic matrices (wUTSM)**:
  row substochastic matrices $S = (s_{ij})$ with $s_{ij} = 0$ whenever $i > j$.

The sets of $n \times n$ DSMs, wDSM, UTSMs, and wUTSMs are all semigroups with identity. We denote the semigroup majorization of DSMs, wDSM, UTSMs, and wUTSMs by $\prec$, $\prec_w$, $\prec^{\mathrm{UT}}$, $\prec_w^{\mathrm{UT}}$ for short, and we say that x is majorized, weakly majorized, UT-majorized, weakly UT-majorized by y, whenever $x \prec y$, $x \prec_w y$, $x \prec^{\mathrm{UT}} y$, or $x \prec_w^{\mathrm{UT}} y$, respectively.

Birkhoff's theorem [5] states that the set of doubly stochastic matrices is the convex hull of all permutations matrices. As a consequence, $x \prec y$ if and only if x lies in the convex hull of the $n!$ permutations of y (this is a restatement of (2.5)). The set of $n \times n$ doubly substochastic matrices is the convex hull of the set of $n \times n$ matrices that have at most one unit in each row and each column, and all other entries are zero [22]. Therefore, $x \prec_w y$ if and only if x is in the convex hull of the points $(\eta_1 y_{\pi(1)}, \dots, \eta_n y_{\pi(n)})$, where π is a permutation and each η_i is 0 or 1.

Let $\mathcal{S}^{\mathrm{UT}}$ and $\mathcal{S}_w^{\mathrm{UT}}$ be the sets of UTSMs and wUTSMs, respectively. (All entries are nonnegative and the row sums are equal to 1.) It is easy to verify that both $\mathcal{S}^{\mathrm{UT}}$ and $\mathcal{S}_w^{\mathrm{UT}}$ are convex sets. The extremal points of $\mathcal{S}^{\mathrm{UT}}$ are upper-triangular matrices with a single entry 1 in each row and all other entries zero. The extremal points of $\mathcal{S}_w^{\mathrm{UT}}$ are upper-triangular matrices with at most one entry 1 in each row and all other entries zero.

To summarise, the antecedents of $y \in \Delta_{n-1}$ in $\mathcal{A} = \mathbb{R}_+^n$ for the preorders $\prec$, $\prec_w$, $\prec^{\mathrm{UT}}$ and $\prec_w^{\mathrm{UT}}$ are

$$\gamma(y, \prec) = \operatorname{conv}\{\big(y_{\pi(1)}, \dots, y_{\pi(n)}\big) : \pi \in S_n\}, \tag{2.6}$$

$$\gamma(y, \prec_w) = \operatorname{conv}\{\big(\eta_1 y_{\pi(1)}, \dots, \eta_n y_{\pi(n)}\big) : \pi \in S_n, \eta_i \in \{0, 1\}\}, \tag{2.7}$$

$$\gamma(y, \prec^{\mathrm{UT}}) = \operatorname{conv}\left\{\left(\sum_{j_1 \in f^{-1}(1)} y_{j_1}, \dots, \sum_{j_n \in f^{-1}(n)} y_{j_n}\right) : f : [n] \to [n] \text{ s.t. } f(i) \geq i \text{ for all } i\right\}. \tag{2.8}$$

$$\gamma(y, \prec_w^{\mathrm{UT}}) = \operatorname{conv}\left\{\left(\sum_{j_1 \in f^{-1}(1)} \eta_{j_1} y_{j_1}, \dots, \sum_{j_n \in f^{-1}(n)} \eta_{j_n} y_{j_n}\right) : f : [n] \to [n] \text{ s.t. } f(i) \geq i \text{ for all } i, \text{ and } \eta \in \{0, 1\}^n\right\}. \tag{2.9}$$

Example 2.1 For $n = 3$, let $y = (y_1, y_2, y_3)$. Then,

$$\begin{aligned}\gamma(y, \prec) &= \operatorname{conv}\{(y_1, y_2, y_3), (y_2, y_1, y_3), (y_2, y_3, y_1), (y_1, y_3, y_2), \\ &\qquad (y_3, y_1, y_2), (y_3, y_2, y_1)\}, \qquad (2.10)\\ \gamma(y, \prec^{\mathrm{UT}}) &= \operatorname{conv}\{(y_1, y_2, y_3), (y_1, 0, y_2 + y_3), (0, y_1 + y_2, y_3), \\ &\qquad (0, y_2, y_1 + y_3), (0, y_1, y_2 + y_3), (0, 0, y_1 + y_2 + y_3)\}. \qquad (2.11)\end{aligned}$$

See Fig. 1 for examples of the antecedent sets.

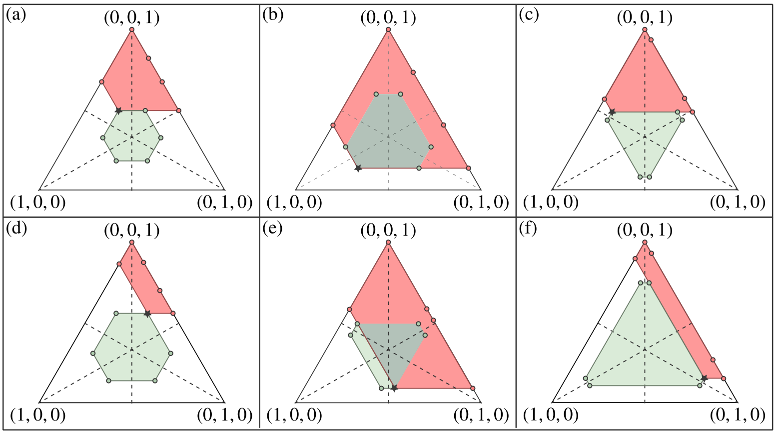

Fig. 1 Antecedent sets $\gamma(y, \prec)$ (in green) and $\gamma(y, \prec^{\mathrm{UT}})$ (in red) for each $y \in \Delta_2$ (depicted as a black star), where y is **a** $y = (0.33, 0.18, 0.49)$, **b** $y = (0.60, 0.26, 0.14)$, **c** $y = (0.43, 0.08, 0.49)$, **d** $y = (0.14, 0.31, 0.55)$, **e** $y = (0.42, 0.49, 0.09)$, and **f** $y = (0.10, 0.74, 0.16)$. Points obtained by applying extreme operations with respect to majorization and UT-majorization are depicted by green and red circles, respectively. Note that these are not necessarily extreme points of the corresponding antecedent sets

Remark 1 The number of extremal points of $\gamma(y, \prec)$ is generically $n!$ (the number of permutations of an n-set); the number of extremal points of $\gamma(y, \prec_w)$ is generically $\sum_{k=0}^{n} k!\binom{n}{k}^2$, the number of partial permutations of an n-set [29]. These are upper bounds for the number of extremal points of $\gamma(y, \prec^{UT})$ and $\gamma(y, \prec_w^{UT})$, respectively. See Fig. 1.

It can be verified [30] that UT-majorization is equivalent to unordered majorization, which is defined by the same inequalities (2.3) where the components of the vectors are unordered. To avoid further notation, we still denote unordered majorization with $x \prec^{\mathrm{UT}} y$, so that we have:

$$x \prec^{\mathrm{UT}} y \quad \text{iff} \quad \begin{cases} \displaystyle\sum_{i=1}^{k} x_i \le \sum_{i=1}^{k} y_i, \quad k = 1, \ldots, n-1 \\ \displaystyle\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i. \end{cases} \tag{2.12}$$

Note that if $x = yM$, with $M \in \mathcal{S}^{\mathrm{UT}}$, then $x_i = \sum_{j=1}^{n} y_j M_{ji} = \sum_{j=1}^{i} y_j M_{ji}$ since $M_{ji} = 0$ for $j > i$. Moreover, $\sum_{j=i}^{n} M_{ij} = 1$. Note that $M_{nn} = 1$ and $e_n = e_n M$ for all $M \in \mathcal{S}^{\mathrm{UT}}$.

Remark 2 *(Physics)* Another prominent example of semigroup majorization is obtained from the semigroup $\mathcal{S}^q$ of q**-stochastic matrices** defined as follows. Let q be a nonnegative vector. A row stochastic matrix S is said to be q-stochastic if

$qS = q$. For $x, y \in \mathbb{R}^n$ we say that x is q-majorized by y, and write $x \prec_{\mathcal{S}^q} y$ if $x = yS$ for some $S \in \mathcal{S}^q$. In q-majorization, $q = \perp$ is always the unique bottom element, while e_k is a top element if and only if q_k is minimal in q. See [36, Theorem 3]. For $q = e$, we get the set of DSMs. Notice that $\mathcal{S}^{\mathrm{UT}}$ is a proper subsemigroup of $\mathcal{S}^{e_n}$, $e_n = (0, \ldots, 0, 1)$.

If we write $q = \left(e^{-\beta E_1}, \ldots, e^{-\beta E_n}\right)$ for some sequence of energy levels $E_1 > \cdots > E_n \in \mathbb{R}$ and inverse temperature $\beta > 0$, then q-majorization is equivalent to **thermomajorization** [16, 33], as $\mathcal{S}^q$ is the semigroup that fixes the 'Gibbs state' q. At high temperature $\beta \to 0$, thermomajorization reduces to standard majorization; at zero temperature $\beta \to \infty$, thermomajorization relation collapses to UT-majorization. Note that in physics it is customary to order the energy levels in increasing order $E_1 < \cdots < E_n$. To make a correspondence with the existing literature [16, 25], it is enough, by taking the transpose, to consider column-stochastic (rather than row stochastic) matrices acting on column probability vectors. Note that transposing an upper-triangular matrix yields a lower-triangular matrix.

Remark 3 *(Combinatorics)* In the realm of discrete combinatorics, the counterpart of majorization is called **dominance order** for integer partitions or integer compositions. There exists a natural extension due to Narayana [24] that can be used in the continuous setting to provide yet another generalisation of majorization. Let $s \in \mathbb{R}$. For $x, y \in \mathbb{R}^n$, the vector x is said to be s*-dominated* by y, and we write $x \prec^s y$ if

$$\sum_{i=1}^{k} x_i^{\downarrow} - \sum_{i=1}^{k} y_i^{\downarrow} \leq s, \text{ for } k = 1, \ldots, n. \tag{2.13}$$

When $s = 0$, the s-dominance order is the weak majorization order.

2.3 Orderings induced by function sets

Here is another convenient description of partial orders induced by functions that can be thought of as generalised entropies. Let Φ be a set of real-valued functions ϕ defined on a subset $\mathcal{A}$ of $\mathbb{R}^n$. For $x, y \in \mathcal{A}$, we write

$$x \prec_{\Phi} y \quad \text{iff} \quad \phi(x) \leq \phi(y), \quad \text{for all } \phi \in \Phi. \tag{2.14}$$

If (2.14) holds, then it holds also for all linear combinations of functions in Φ with positive coefficients. These linear combinations form a cone that is said to be generated by Φ.

For instance, let $\Phi := \{\phi_1, \ldots, \phi_{n+1}\}$, where

$$\phi_k(x) = \sum_{i=1}^{k} x_i^{\downarrow}, \quad k = 1, \ldots, n, \qquad \phi_{n+1}(x) = -\phi_n(x). \tag{2.15}$$

Then, $\prec_{\Phi}$ is equivalent to $\prec$ according to definition (2.3). The cone generated by $\Phi_w := \{\phi_1, \ldots, \phi_n\}$, with ϕ_k in (2.15) induces the weak majorization preorder on $\mathbb{R}^n_+$.

Similarly, the sets $\Phi_w^{\mathrm{UT}} := \{\phi_1^{\mathrm{UT}}, \ldots, \phi_n^{\mathrm{UT}}\}$, and $\Phi^{\mathrm{UT}} := \Phi_w^{\mathrm{UT}} \cup \{\phi_{n+1}^{\mathrm{UT}}\}$, where

$$\phi_k^{\mathrm{UT}}(x) = \sum_{i=1}^{k} x_i, \quad k = 1, \ldots, n, \qquad \phi_{n+1}^{\mathrm{UT}}(x) = -\phi_n^{\mathrm{UT}}(x), \tag{2.16}$$

generate the cones inducing weak UT-majorization and UT majorization, respectively.

3 Probability that generic states are comparable

Let $(\mathcal{A}, \sqsubseteq)$ be a preorder with $\mathcal{A} \subset \mathbb{R}^n$, and let μ be a probability measure on $\mathcal{A}$. If Γ is measurable, we define the **probability** of Γ as

$$P(\Gamma) = \mu{\otimes}\mu\,(\Gamma) = \mu{\otimes}\mu\,(\{(x, y) \in \mathcal{A} \times \mathcal{A} \colon x \sqsubseteq y\})\,. \tag{3.1}$$

The above quantity can be interpreted as follows. Pick X, Y independently at random according to the probability measure μ. Then, $P(\Gamma) = P(X \sqsubseteq Y)$ is the probability that X is less than Y. By symmetry, this is also equal to the probability $P(Y \sqsubseteq X)$, since X, Y are independent and identically distributed.

In this paper, we consider the case of partial orders on the simplex of discrete probability vectors.

Let $n \in \mathbb{N}$ and consider the $(n-1)$-dimensional simplex

$$\Delta_{n-1} := \left\{ x \in \mathbb{R}^n : x_i \geq 0, \sum_{i=1}^{n} x_i = 1 \right\} = \mathrm{conv}\{e_1, \ldots, e_n\}, \tag{3.2}$$

where $e_1, \ldots, e_n$ stand for the unit vectors of the standard orthonormal basis of $\mathbb{R}^n$.

With this notation, consider the poset $(\mathcal{A}, \sqsubseteq) = (\Delta_{n-1}, \prec)$, with $\prec$ the preorder of majorization. Note that the bottom element is $\perp = \frac{1}{n}(e_1 + \cdots + e_n) = (1/n, \ldots, 1/n)$, while the top elements are $e_1, \ldots, e_n$. In [7], Cunden, Facchi, Florio and Gramegna, proved the following asymptotic result.

Theorem 3.1 ([7]) *Let $\Gamma_n = \Gamma(\Delta_{n-1}, \prec)$, and μ_n the uniform probability measure on the simplex Δ_{n-1}. Write $P(\Gamma_n) = \mu_n \otimes \mu_n(\Gamma_n)$. Then,*

$$\lim_{n\to\infty} P(\Gamma_n) = 0. \tag{3.3}$$

This statement is Eq. (1.1) of Theorem A. The proof of Theorem 3.1 in [7] uses Kolmogorov's 0–1 law, and hence, it is fundamentally non-quantitative. Recently, Jain, Kwan and Michelen [18] proved the following quantitative (though presumably not optimal) estimate.

Theorem 3.2 ([18]) *With the same notation of Theorem 3.1, there is a constant $C > 0$ such that*

$$P(\Gamma_n) \leq C\,(\log n/n)^{1/16}\,. \tag{3.4}$$

The same authors [18] proved the analogue of Theorem 3.1 when the measure μ_n on Δ_{n-1} comes from the fixed-trace Wishart measure of random matrix theory, thus settling a conjecture put forward by Nielsen [26] and numerically investigated in [6] in the context of entanglement theory.

Consider now the poset $(\mathcal{A}, \sqsubseteq) = \left(\Delta_{n-1}, \prec^{\mathrm{UT}}\right)$; in this case the bottom element is $\bot = e_n = (0, 0, \ldots, 1)$, while the top element is $\top = e_1 = (1, 0, \ldots, 0)$. We state the following simple theorem on the exact probability of Γ_n for UT-majorization.

Theorem 3.3 *Consider the graph $\Gamma_n = \Gamma\left(\Delta_{n-1}, \prec^{UT}\right)$ and let μ_n be an absolutely continuous and permutation symmetric probability measure on Δ_{n-1}. Then, for all $n \geq 2$,*

$$P(\Gamma_n) = \frac{1}{n}. \tag{3.5}$$

When μ_n is the uniform measure on the simplex, Theorem (3.3) provides Eq. (1.3) of Theorem B.

4 Maximal success probability of conversion

Note that for $x, y \in \Delta_{n-1}$, $x \prec y$ if and only if $x \prec_w y$, since the normalisation condition automatically guarantees the equality constraint in Eq. (2.3).

For $x, y \in \Delta_{n-1}$, it can happen that x is not majorized by y, but it is always possible to find $p \in [0, 1]$ and $z \in \Delta_{n-1}$ such that

$$px + (1-p)z \prec y. \tag{4.1}$$

Let

$$\Pi : \Delta_{n-1} \times \Delta_{n-1} \to [0, 1], \tag{4.2}$$

with $\Pi(x, y)$ defined as the largest $p \in [0, 1]$ such that (4.1) holds for some $z \in \Delta_{n-1}$. Equivalently, $\Pi(x, y)$ is the largest p such that $yD = px + (1-p)z$ for some DSM D and some vector $z \in \Delta_{n-1}$.

The function Π can be expressed in terms of the monotones $\phi_1, \ldots, \phi_n$ as defined in (2.15).

Proposition 4.1

$$\Pi(x, y) = \min_{1 \le k \le n} \frac{\phi_k(y)}{\phi_k(x)}. \tag{4.3}$$

where ϕ_k are defined in (2.15).

Proof Let $p^* = \min_{1 \le k \le n} \frac{\phi_k(y)}{\phi_k(x)}$. We first show that

$$p^* = \max\{p \in [0, 1] : px \prec_w y\}. \tag{4.4}$$

Let $k^* \in \{1, \dots, n\}$ such that $p^* = \frac{\phi_{k^*}(y)}{\phi_{k^*}(x)}$. Then,

$$\phi_k(p^* x) = p^* \phi_k(x) = \left(\frac{\phi_{k^*}(y)}{\phi_{k^*}(x)}\right) \phi_k(x) \le \left(\frac{\phi_k(y)}{\phi_k(x)}\right) \phi_k(x) = \phi_k(y), \text{ for all } k = 1, \dots, n.$$

Hence, $p^* x \prec_w y$. Now suppose that there is $p' > p^*$ such that $p'x \prec_w y$. Then, $\phi_k(p'x) > \phi_k(p^* x)$ for all $k = 1, \dots, n$. Choosing $k = k^*$,

$$\phi_{k^*}(p'x) > \phi_{k^*}(p^* x) = p^* \phi_{k^*}(x) = \phi_{k^*}(y).$$

we get a contradiction. Hence, p^* is the largest value $p \in [0, 1]$ such that $px \prec_w y$, i.e. such that $yP = px$ for some wDSM P.

We can now prove the claim (4.3). The case $p^* = 1$ is trivial. Assume $p^* < 1$. Reasoning as above, we first show that $\Pi(x, y) \le p^*$. Suppose by contradiction that there is $p' > p^*$ such that $p'x + (1 - p')z \prec y$, for some $z \in \Delta_{n-1}$. Then, $\phi_k(y) \ge \phi_k(p'x + (1 - p')z) \ge \phi_k(p'x) > \phi_k(p^* x)$ for all $k = 1, \dots, n$ and, choosing $k = k^*$, we get a contradiction.

It remains to show that there exists a DSM D such that $yD = p^* x + (1 - p^*)z$ with $z \in \Delta_{n-1}$. By (4.4), we know that $yP = p^* x$ with P wDSM. A theorem attributed to von Neumann (see [23, P. 37]) ensures that there exists a DSM D such that $P_{ij} \le D_{ij}$ for all $i, j = 1, \dots, n$. Note that $P' := D - P$ is also a wDSM. Then,

$$yD = yP + y(D - P) = p^* x + (1 - p^*)z, \tag{4.5}$$

where $z = \frac{1}{1 - p^*} y(D - P) \in \Delta_{n-1}$. □

Figure 2 illustrates Proposition 1 by showing how the optimal value p^* of p in the expression $px + (1 - p)z \prec y$ corresponds to the boundary of the antecedent set of y under weak majorization (top panel) and weak UT-majorization (bottom panel). In both cases, the point $p^* x$ (shown in pink) lies on the boundary of the shaded region, confirming that it is the largest value for which px alone remains weakly majorized (UT-majorized) by y. See (4.4).

Note that

$$1/n \le \Pi(x, y) \le 1. \tag{4.6}$$

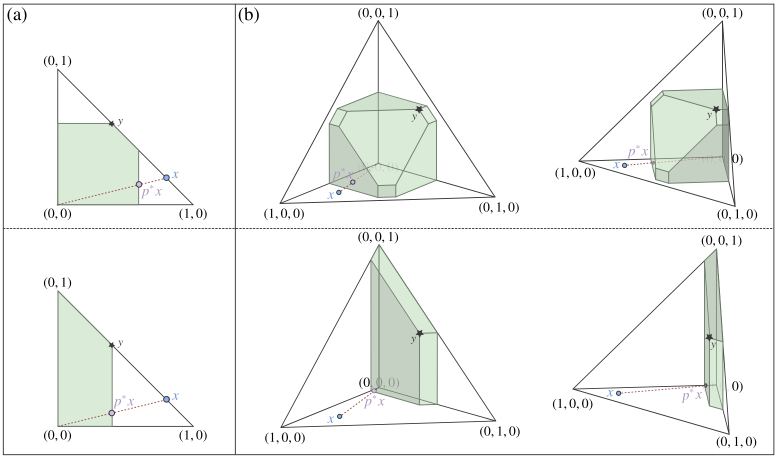

Fig. 2 The antecedent set of y is shown for both the weak (top panel) and weak UT (bottom panel) majorizations for **a** $x = \left(\frac{8}{10}, \frac{2}{10}\right)$ (blue circle) and $y = \left(\frac{2}{5}, \frac{3}{5}\right)$ (black star), with p^*x (pink circle) lying on the boundary of the antecedent set, and **b** $x = \left(\frac{7}{10}, \frac{1}{4}, \frac{1}{20}\right)$ (blue circle) and $y = \left(\frac{1}{12}, \frac{5}{12}, \frac{1}{2}\right)$ (black star), again with p^*x (pink circle) lying on the boundary of the antecedent set

The minimum is achieved for $x = e_1 = (1, 0, \ldots, 0)$ and $y = (1/n, \ldots, 1/n)$. On the other hand, $\Pi(x, y) = 1$ if and only if $x \prec y$, and so

$$\Gamma(\Delta_{n-1}, \prec) = \{(x, y) \in \Delta_{n-1} \times \Delta_{n-1} \colon \Pi(x, y) = 1\}. \tag{4.7}$$

Define

$$P(\Pi(X, Y) \leq t) = \mu \otimes \mu\left(\{(x, y) \in \Delta_{n-1} \times \Delta_{n-1} \colon \Pi(x, y) \leq t\}\right). \tag{4.8}$$

The next result is a restatement of Eq. (1.2) of Theorem A.

Theorem 4.2 *Let μ_n be the uniform probability measure on the simplex Δ_{n-1}. Then,*

$$\lim_{n\to\infty} \Pi(X, Y) = 1, \tag{4.9}$$

in probability.

The proof is based on formula (4.3). The idea is that if X is uniform in the simplex, then for each fixed $\epsilon > 0$, $\phi_k(X)$ is essentially equal to $(1 \pm \epsilon)$ times its mean, aside from an event with probability that is small even when we sum over $k = 1, \ldots, n$. This then shows that $\phi_k(X)/\phi_k(Y) = (1 \pm \epsilon)$ for all k with probability tending to 1. Since this holds for each $\epsilon > 0$, it shows $\Pi(X, Y)$ converges to 1 in probability. The rate of convergence (i.e. how small ϵ can be taken it we allow it to decrease to 0 as $n \to \infty$) seems quite slow; from the proof is appears that ϵ must be at least $1/\text{poly}(\log n)$. It is therefore tricky to see evidence of the convergence (4.9) in a simulation without taking n to be fairly large. See Fig. 3).

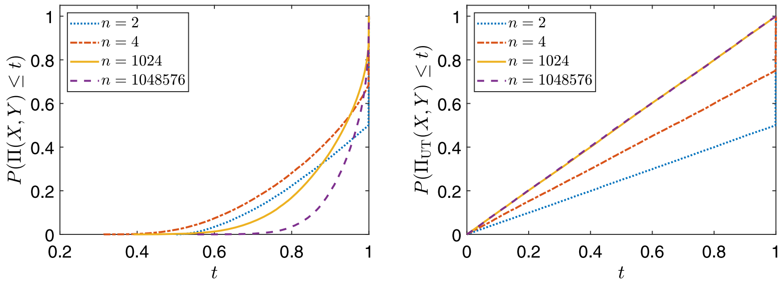

Fig. 3 Empirical cumulative distribution of **a** $\Pi(X, Y)$ and **b** $\Pi_{\mathrm{UT}}(X, Y)$, obtained from 10^5 samples of uniformly random pairs (x_i, y_i) in the simplex Δ_{n-1}. One can see that the distribution of $\Pi(X, Y)$ concentrates at 1 as $n \to \infty$, at a slow rate

Remark 4 A function similar to $\Pi(\cdot, \cdot)$ can be considered for the converse relation $\succ$, defined analogously to (4.4), but in terms of the sums of the smallest components of x and y. Such a function appears in the resource theory of entanglement, and its statistical behaviour was studied in [6, 18].

Mutatis mutandis, the same question can be posed for UT-majorization. Let

$$\Pi_{\mathrm{UT}} : \Delta_{n-1} \times \Delta_{n-1} \to [0, 1], \tag{4.10}$$

defined as

$$\Pi_{\mathrm{UT}}(x, y) := \max \left\{ p \in [0, 1] \colon px + (1-p)z \prec^{\mathrm{UT}} y, \quad \text{for some } z \in \Delta_{n-1} \right\}. \tag{4.11}$$

We have the following simple representation whose proof mimicks the one of Proposition 4.1.

Proposition 4.3

$$\Pi_{UT}(x, y) = \min_{1 \le k \le n} \frac{\phi_k^{UT}(y)}{\phi_k^{UT}(x)}. \tag{4.12}$$

where ϕ_k^{UT} are defined in (2.16).

In this case, $0 \le \Pi_{\mathrm{UT}}(x, y) \le 1$. The minimum $\Pi_{\mathrm{UT}}(x, y) = 0$ is achieved, for instance, when $y_1 = 0$ but $x_1 > 0$, while $\Pi_{\mathrm{UT}}(x, y) = 1$ if and only if $x \prec^{\mathrm{UT}} y$, so that

$$\Gamma(\Delta_{n-1}, \prec^{\mathrm{UT}}) = \{(x, y) \in \Delta_{n-1} \times \Delta_{n-1} \colon \Pi_{\mathrm{UT}}(x, y) = 1\}. \tag{4.13}$$

Define

$$P\left(\Pi_{\mathrm{UT}}(X, Y) \le t\right) = \mu \otimes \mu \left(\{(x, y) \in \Delta_{n-1} \times \Delta_{n-1} \colon \Pi_{\mathrm{UT}}(x, y) \le t\}\right). \tag{4.14}$$

When X, Y are uniformly distributed, $\Pi_{\mathrm{UT}}(X, Y)$ has a (surprisingly simple!) explicit distribution.

Theorem 4.4 *Let μ_n be the uniform measure on Δ_{n-1}. Then, for all $n \geq 2$,*

$$P\left(\Pi_{UT}(X, Y) \leq t\right) = \begin{cases} 0 & \text{if } t \leq 0 \\ \left(1 - \frac{1}{n}\right) t & \text{if } 0 < t < 1 \\ 1 & \text{if } t \geq 1 \end{cases} \tag{4.15}$$

for all $t \in \mathbb{R}$. In particular, as $n \to \infty$, $\Pi_{UT}(X, Y)$ converges in distribution to a uniform random variable in $[0, 1]$.

The theorem above corresponds to Eq. (1.4) of Theorem B. When μ_n is the uniform measure, Theorem 3.3 is a Corollary of Theorem 4.4.

Remark 5 Equation (4.15) can be expressed in purely geometrical terms. The volume of the $2(n-1)$-dimensional polytope

$$\left\{(x, y) \in \Delta_{n-1} \times \Delta_{n-1} \colon t(x_1 + \cdots + x_k) \leq y_1 + \cdots + y_k \text{ for all } 1 \leq k \leq n\right\} \tag{4.16}$$

is the right-hand side of (4.15) times the volume of $\Delta_{n-1} \times \Delta_{n-1}$ (that is $(n-1)! \times (n-1)!$). A related polytope was studied in detail by Stanley and Pitman in [34], and their paper was a great source of information in our investigations.

Note that (4.15) is *not* a universal (distribution independent) result. We have, for instance, considered the family of permutation-invariant Dirichlet distribution $\mu_n = \mathrm{Dir}(\alpha, \ldots, \alpha)$, with density on the unit simplex given by

$$f_{X_1,\ldots,X_n}(x) = \frac{\Gamma(n\alpha)}{\Gamma(\alpha)^n} \prod_{i=1}^{n} x_i^{\alpha-1} \mathbf{1}_{\Delta_{n-1}}(x). \tag{4.17}$$

The above family includes the uniform distribution on Δ_{n-1} for $\alpha = 1$. One can verify by explicit computation for small n, that the distribution of $\Pi_{\mathrm{UT}}(X, Y)$ when $\alpha \neq 1$ is not given by (4.15). For example, when $n = 3$ and $\alpha = 2$,

$$P\left(\Pi_{\mathrm{UT}}(X, Y) \leq t\right) = \frac{10}{7} t^2 - \frac{10}{7} t^3 + \frac{10}{9} t^4 - \frac{4}{9} t^5, \quad 0 < t < 1.$$

We did not manage to unveil a closed form for general α and n. However, using the known 'stick-breaking' representation of Dirichlet-distributed random variables (see Appendix B) one can get the following simple bound.

Proposition 4.5 *Let $\mu_n = \mathrm{Dir}(\alpha, \ldots, \alpha)$ on Δ_{n-1}. Then, for all $n \geq 2$,*

$$P\left(\Pi_{UT}(X, Y) \leq t\right) \leq \left(1 - \frac{1}{\alpha n}\right) t. \tag{4.18}$$

5 Proofs

In this section, we will use some standard notation and terminology. For two random variables X and Y, we write $X \stackrel{d}{=} Y$ to denote **identity in distribution**. We say that $X_1, \ldots, X_n$ are **independent and identically distributed (i.i.d.)** random variables, if they are jointly independent and $X_i \stackrel{d}{=} X_j$ for all i, j. We say that $X_1, \ldots, X_n$ are **exchangeable**, if their joint law is invariant under permutations, $(X_1, \ldots, X_n) \stackrel{d}{=} (X_{\sigma(1)}, \ldots, X_{\sigma(n)})$, for all $\sigma \in S_n$. Finally, Z is an **exponential** random variable **with rate** 1, if it has density $f_Z(z) = e^{-z}$ for $z \geq 0$, and zero otherwise.

5.1 Proof of theorem 3.3

Pick X, Y independently according to μ. Since UT-majorization is equivalent to unordered majorization, we want to compute

$$P\left(X \preceq^{\mathrm{UT}} Y\right) = P\left(\sum_{i=1}^{k}(Y_i - X_i) \geq 0, \text{for all } k = 1, \ldots n\right). \tag{5.1}$$

Let $Z := Y - X$ and define

$$S_0 = 0, \quad S_k = S_{k-1} + Z_k, \quad \text{for } k = 1, \ldots, n. \tag{5.2}$$

Note that $(S_k)_{0 \leq k \leq n}$ is a random bridge starting and ending at 0, i.e. $S_0 = S_n = 0$. Hence, the probability (5.1) can be seen as the *persistence probability* above level 0 of the random bridge $(S_k)_k$:

$$P\left(X \preceq^{\mathrm{UT}} Y\right) = P\left(S_k \geq 0, \text{for all } k = 1, \ldots, n\right). \tag{5.3}$$

The steps $Z_1, \ldots, Z_n$ are exchangeable, and we have $P(S_k = 0) = 0$ for all $k = 1, \ldots n-1$, since μ is absolutely continuous. Hence, by Sparre–Andersen theorem [2, Theorem 2], denoting by N_n the number of sums $S_1, \ldots, S_n$ which are > 0, we have

$$P(N_n = m) = \frac{1}{n},$$

for all $m = 1, \ldots, n$. In particular $P(N_n = n) = P(S_k \geq 0, \text{for allk=1,...,n}) = 1/n$.

5.2 Preparatory propositions

The starting point is that if X is uniformly distributed on the simplex; then, its components and the decreasing rearrangement of its components have a closed-form description in terms of independent exponential random variables (see, for instance, [11] or [18, Theorems 3.1 and 3.2]). These classical results provide convenient representations of the monotones $\phi_k(X)$ and $\phi_k^{\mathrm{UT}}(X)$.

Proposition 5.1 *Let $X = (X_1, \ldots, X_n)$ be a uniform point on the simplex Δ_{n-1}, and $Z_1, \ldots, Z_n$ be i.i.d. exponential random variables with rate* 1.

(1) Denote $\phi_k(X) := \sum_{i=1}^k X_i^{\downarrow}$ for $k = 1, \ldots, n$. Then,

$$(\phi_1(X), \phi_2(X), \ldots, \phi_n(X)) \stackrel{d}{=} \frac{1}{V_n}(V_1, V_2, \ldots, V_n), \tag{5.4}$$

where $V_k = a_{k,1}Z_1 + a_{k,2}Z_2 + \cdots + a_{k,n}Z_n$, with

$$a_{k,i} = \begin{cases} 1 & \text{for } i \leq k \\ \frac{k}{i} & \text{for } i > k \end{cases}. \tag{5.5}$$

(2) Denote $\phi_k^{UT}(X) := \sum_{i=1}^k X_i$ for $k = 1, \ldots, n$. Then,

$$\left(\phi_1^{UT}(X), \phi_2^{UT}(X), \ldots, \phi_n^{UT}(X)\right) \stackrel{d}{=} \frac{1}{W_n}(W_1, W_2, \ldots, W_n), \tag{5.6}$$

where $W_k = b_{k,1}Z_1 + b_{k,2}Z_2 + \cdots + b_{k,n}Z_n$, with

$$b_{k,i} = \begin{cases} 1 & \text{for } i \leq k \\ 0 & \text{for } i > k \end{cases}. \tag{5.7}$$

A more convenient representation for the monotones $\phi_k^{\mathrm{UT}}(X)$ is provided by a connection between partial sums of a uniform point on the simplex and the order statistics of independent uniform random variables. If $U_1, \ldots, U_{n-1}$ are $n-1$ independent random variables, we denote the corresponding order statistics with $U_{(n-1,1)}, \ldots U_{(n-1,n-1)}$, where $U_{(n-1,i)}$ is the $i-$th smallest among the U_j's, that is the $n-1$ random variables are rearranged in nondecreasing order: $U_{(n-1,1)} \leq \cdots \leq U_{(n-1,n-1)}$.

Then, using classical representation theorems for order statistics (see, e.g. [1, 17]), we have the following.

Proposition 5.2 *Let $X = (X_1, \ldots, X_n)$ be a uniform point on the simplex Δ_{n-1}, and $U_1, \ldots, U_{n-1}$ be i.i.d. uniform random variables on the interval* $[0, 1]$. *Denote $\phi_k^{UT}(X) := \sum_{i=1}^k X_i$ for $k = 1, \ldots, n$. Then,*

$$\left(\phi_1^{UT}(X), \phi_2^{UT}(X), \ldots, \phi_n^{UT}(X)\right) \stackrel{d}{=} \left(U_{(n-1,1)}, \ldots, U_{(n-1,n-1)}, 1\right). \tag{5.8}$$

Since $X_j = \phi_j^{UT}(X) - \phi_{j-1}^{UT}(X)$ (with $\phi_0^{UT}(X) := 0$), this is also equivalent to say that X has the same distribution of the sequence of spacings between order statistics of $n-1$ uniform random variables.

Remark 6 If $U_1, \ldots, U_{n-1}$ are i.i.d. uniform random variables on $[0, 1]$, $(U_{(n-1,1)}, \ldots, U_{(n-1,n-1)})$ is a uniform point over the simplex $\mathcal{E}_{n-1} := \{u \in \mathbb{R}^{n-1} : 0 \le u_1 \le u_2 \le \cdots \le u_{n-1} \le 1\}$.

5.3 Proof of theorem 4.2

The technical ingredient to prove Theorem 4.2 is the following bound.

Proposition 5.3 (Bernstein's inequality (see, for example, [38, Theorem 2.8.2])) *Let $Z_1, \ldots, Z_n$ be i.i.d. exponential random variables with rate 1, and let $V = a_1 Z_1 + \cdots + a_n Z_n$ for some coefficients $a_1, \ldots, a_n \in \mathbb{R}$. There is an absolute constant $c > 0$ such that*

$$P\left(|V - EV| \ge t\right) \le \exp\left(-c \min\left(\frac{t^2}{a_1^2 + \cdots + a_n^2}, \frac{t}{\max|a_i|}\right)\right), \tag{5.9}$$

for every $t \ge 0$.

Corollary 5.4 *Let V_k defined as in (5.4)-(5.5). Then, for all $k \le n$,*

$$P\left(|V_k - EV_k| \ge \epsilon E V_k\right) \le \exp\left(-c \min\left(\epsilon^2 A_k, \epsilon B_k\right)\right), \tag{5.10}$$

where

$$A_k := \frac{\left(\sum_{i=1}^n a_{k,i}\right)^2}{\sum_{i=1}^n a_{k,i}^2}, \quad B_k := \frac{\sum_{i=1}^n a_{k,i}}{\max_{1 \le i \le n} a_{k,i}}. \tag{5.11}$$

We denote by $H_n = \sum_{i=1}^n \frac{1}{i}$ the n-th harmonic number, and $H_{n,m} = \sum_{i=1}^n \frac{1}{i^m}$ the n-th harmonic number of order m.

Lemma 5.5 *For all $k = 1, \ldots, n$,*

$$\sum_{i=1}^n a_{k,i} = k\left(1 + H_n - H_k\right),$$
$$\sum_{i=1}^n a_{k,i}^2 = k\left(1 + k(H_{n,2} - H_{k,2})\right), \quad \max_{1 \le i \le n} a_{k,i} = 1. \tag{5.12}$$

Lemma 5.6 *We have*

$$A_k \ge \frac{1}{2}\left(\log\frac{n}{2}\right)^2 + \alpha_n (k-1), \tag{5.13}$$
$$B_k \ge \log\frac{n}{2} + \beta_n (k-1) \tag{5.14}$$

where

$$\alpha_n = \frac{n}{n-1}\left(1 + \frac{\left(1+\log\left(\frac{n+1}{2}\right)\right)^2}{2n-1}\right), \qquad \beta_n = 1 - \frac{\log\left(\frac{n+1}{2}\right)}{n-1}. \tag{5.15}$$

Proof We start from the following bounds,

$$H_n - H_k = \sum_{i=k+1}^{n} \frac{1}{i} \geq \int_{k+1}^{n+1} \frac{dx}{x} = \log\frac{n+1}{k+1}, \tag{5.16}$$

$$H_{n,2} - H_{k,2} = \sum_{i=k+1}^{n} \frac{1}{i^2} \leq \int_{k}^{n} \frac{dx}{x^2} = \frac{n-k}{nk}. \tag{5.17}$$

Therefore,

$$A_k = \frac{k\,(1+H_n-H_k)^2}{\left(1+k(H_{n,2}-H_{k,2})\right)} \geq \frac{nk}{(2n-k)}\left(1+\log\frac{n+1}{k+1}\right)^2. \tag{5.18}$$

The function $x \mapsto g(x) := \frac{nx}{(2n-x)}\left(1+\log\frac{n+1}{x+1}\right)^2$ is concave for $x \in [1,n]$. Hence, it can be bounded from below by the linear interpolation $g(1) + \frac{g(n)-g(1)}{n-1}(x-1)$. This gives

$$A_k \geq \frac{n\left(1+\log\left(\frac{n+1}{2}\right)\right)^2}{2n-1} + \frac{\left(n + \frac{n\left(1+\log\left(\frac{n+1}{2}\right)\right)^2}{2n-1}\right)}{n-1}(k-1), \tag{5.19}$$

and hence (5.13). The proof of (5.14) for B_k is similar. From (5.16),

$$B_k = k\,(1+H_n-H_k) \geq k\left(1+\log\frac{n+1}{k+1}\right). \tag{5.20}$$

The map $x \mapsto h(x) := x\left(1+\log\frac{n+1}{x+1}\right)$ is concave for $x \in [1,n]$. We can therefore write

$$B_k \geq \log\left(\frac{n+1}{2}\right) + 1 + \frac{\left(n-\log\left(\frac{n+1}{2}\right)-1\right)}{n-1}(k-1), \tag{5.21}$$

and hence (5.14). $\square$

Proof of Theorem 4.2 We want to show that $P\left(\Pi(X,Y) < 1-\epsilon\right) \to 0$, as $n \to \infty$. Let $\epsilon \in (0,1)$. We have

$$P\left(\Pi(X,Y) < 1-\epsilon\right) = P\left(\min_{1\leq k\leq n}\frac{\phi_k(Y)}{\phi_k(X)} < 1-\epsilon\right)$$

$$= P\left(\frac{\phi_k(Y)}{\phi_k(X)} < 1-\epsilon,\ \text{ for some } k \le n\right)$$
$$\le \sum_{k=1}^{n} P\left(\frac{\phi_k(Y)}{\phi_k(X)} < 1-\epsilon\right),$$

by the union bound. Using the representation of Proposition 5.1, we want to bound

$$P\left(\frac{\phi_k(Y)}{\phi_k(X)} < 1-\epsilon\right) = P\left(\frac{V_k(Y)}{V_n(Y)} < (1-\epsilon)\frac{V_k(X)}{V_n(X)}\right) \le \delta_{n,k}, \tag{5.22}$$

so that,

$$\sum_{k=1}^{n} \delta_{n,k} \to 0, \quad \text{as } n \to \infty. \tag{5.23}$$

By Corollary 5.4, $|V_n - n| \le n\epsilon$ with probability at least $1 - e^{-c\epsilon^2 n}$. Hence,

$$\begin{aligned}
&P\left(\frac{V_k(Y)}{V_n(Y)} < (1-\epsilon)\frac{V_k(X)}{V_n(X)}\right)\\
&\le P\Bigg(V_k(Y) < (1-\epsilon)\frac{V_n(Y)}{V_n(X)}V_k(X), \quad (V_n(X), V_n(Y))\\
&\quad \in ((1-\epsilon)n, (1+\epsilon)n)^2\Bigg) + 2e^{-c\epsilon^2 n}\\
&\le P\left(V_k(Y) < (1+\epsilon)V_k(X)\right) + 2e^{-c\epsilon^2 n}\\
&\le P\left(V_k(Y) < (1+\epsilon)V_k(X), \quad V_k(X) \in ((1-\epsilon)EV_k(X), (1+\epsilon)EV_k(X))\right)\\
&\quad + P\left(V_k(Y) < (1+\epsilon)V_k(X), \quad V_k(X) \notin ((1-\epsilon)EV_k(X), (1+\epsilon)EV_k(X))\right) + 2e^{-c\epsilon^2 n}\\
&\le P\left(V_k(Y) < (1+\epsilon)^2 EV_k(X)\right) + e^{-c\min\left(\epsilon^2 A_k, \epsilon B_k\right)} + 2e^{-c\epsilon^2 n}\\
&\le P\left(|V_k(Y) - EV_k(Y)| < 3\epsilon EV_k(Y)\right) + e^{-c\min\left(\epsilon^2 A_k, \epsilon B_k\right)} + 2e^{-c\epsilon^2 n}\\
&\le 4e^{-c\min\left(\epsilon^2 A_k, \epsilon B_k\right)},
\end{aligned}$$

where in the last line we used $A_k \leqslant n$.

Therefore, we have the bound (5.22) with

$$\delta_{n,k} := 4e^{-c\epsilon^2 A_k} + 4e^{-c\epsilon B_k}. \tag{5.24}$$

It remains to show (5.23). For this, we use Lemma 5.6:

$$\sum_{k=1}^{n} e^{-c\epsilon^2 A_k} \le e^{-c(\epsilon\log(n/2))^2/2} \sum_{k=0}^{n-1} e^{-c\epsilon^2 \alpha_n k} = e^{-c(\epsilon\log(n/2))^2/2} \frac{1 - e^{-c\epsilon^2 n\alpha_n}}{1 - e^{-c\epsilon^2 \alpha_n}} \to 0,$$
$$\sum_{k=1}^{n} e^{-c\epsilon B_k} \le e^{-c\epsilon\log(n/2)} \sum_{k=0}^{n-1} e^{-c\epsilon\beta_n k} = e^{-c\epsilon\log(n/2)} \frac{1 - e^{-c\epsilon n\beta_n}}{1 - e^{-c\epsilon\beta_n}} \to 0,$$

as $n \to \infty$. □

5.4 Proof of theorem 4.4 and proposition 4.5

The connection with order statistics given in Proposition 5.2 is useful for our purposes since a large body of literature has been devoted to the study of order statistics of independent random variables (see, e.g. [1, 12, 13, 35]). Of special interest for us are the probabilities

$$P\left(a_j \le U_{(m,j)} \le b_j \text{ for all } 1 \le j \le m\right), \tag{5.25}$$

where $0 \le a_1 \le \cdots \le a_m \le 1$ and $0 \le b_1 \le \cdots \le b_m \le 1$.

From now on, we specialise (5.25) to the case $b_j = 1$ for all $1 \le j \le m$, and we use the notation

$$P_m(a_1, \ldots, a_m) := P(U_{(m,j)} > a_j \text{ for all } 1 \le j \le m). \tag{5.26}$$

The key ingredient to prove Theorem 4.4 is a recursive relation for (5.26) discovered by L. N. Bolshev shortly before his death and later published in [20].

Proposition 5.7 (Bolshev's recursion, [35] §9.3) *Let* $0 \le a_1 \le \cdots \le a_m \le 1$. *Then,*

$$P_m(a_1, \ldots, a_m) = 1 - \sum_{k=1}^{m} \binom{m}{k} a_k^k P_{m-k}(a_{k+1}, \ldots, a_m), \tag{5.27}$$

with $P_0 \equiv 1$.

A proof of this proposition can be found in [20] and [35] in terms of empirical distribution functions. For the reader's convenience, we include a self-contained proof in Appendix A.

Proof of Theorem 4.4 Let $F_n(t) := P(\Pi_{\mathrm{UT}}(X, Y) \le t)$ be the distribution function of $\Pi_{\mathrm{UT}}(X, Y)$. By the normalisation of $X, Y \in \Delta_n$,

$$0 \le \Pi_{\mathrm{UT}}(X, Y) \le \frac{Y_1 + \cdots + Y_n}{X_1 + \cdots + X_n} = 1, \tag{5.28}$$

almost surely, which implies $F_n(t) = 0$ for $t < 0$, and $F_n(t) = 1$ for $t \ge 1$. Let $0 \le t < 1$, and denote with $S_k := X_1 + \cdots + X_k$ and $T_k := Y_1 + \cdots + Y_k$ for $k = 1, \ldots, n-1$. Then,

$$F_n(t) = 1 - P(\Pi_{\mathrm{UT}}(X, Y) > t) = 1 - P\left(\frac{T_1}{S_1} > t, \ldots, \frac{T_{n-1}}{S_{n-1}} > t\right), \tag{5.29}$$

where we dropped the inequality $T_n/S_n > t$, since it is redundant for $t < 1$. Using Proposition 5.2 and the notation (5.26), we can write

$$1 - F_n(t) = P\left(T_1 > tS_1, \ldots, T_{n-1} > tS_{n-1}\right) = E_S[P_{n-1}(tS_1, \ldots, tS_{n-1})], \tag{5.30}$$

where E_S denotes the average with respect to $S_1, S_2, \ldots, S_{n-1}$. Inserting Bolshev's recursion (5.27) into (5.30), one has

$$F_n(t) = \sum_{k=1}^{n-1} \binom{n-1}{k} t^k E_S \left[S_k^k P_{n-1-k}(tS_{k+1}, \ldots, tS_{n-1}) \right]. \tag{5.31}$$

Isolating the first term in (5.31) and relabelling the summation index on the remaining terms, one obtains

$$\begin{aligned} F_n(t) = (n-1)tE_S \left[S_1 P_{n-2}(tS_2, \ldots, tS_{n-1}) \right] \\ + \sum_{k=1}^{n-2} \binom{n-1}{k+1} t^{k+1} E_S \left[S_{k+1}^{k+1} P_{n-k}(tS_{k+2}, \ldots, tS_{n-1}) \right], \end{aligned} \tag{5.32}$$

We now use Bolshev's recursion again on the first term of (5.32), obtaining:

$$\begin{aligned} F_n(t) &= (n-1)tE_S S_1 - (n-1)t \sum_{k=1}^{n-2} \binom{n-2}{k} t^k E_S \\ &\quad \left[S_1 S_{k+1}^k P_{n-2-k}(tS_{k+2}, \ldots, tS_{n-1}) \right] \\ &\quad + \sum_{k=1}^{n-2} \binom{n-1}{k+1} t^{k+1} E_S \left[S_{k+1}^{k+1} P_{n-k-2}(tS_{k+2}, \ldots, tS_{n-1}) \right] \\ &= (n-1)tE_S S_1 + \sum_{k=1}^{n-2} \binom{n-1}{k+1} t^{k+1} E_S \\ &\quad \left[(S_{k+1} - (k+1)S_1) S_{k+1}^k P_{n-k-2}(tS_{k+2}, \ldots, tS_{n-1}) \right]. \end{aligned} \tag{5.33}$$

The average in the k-th term of the summation is given by:

$$\begin{aligned} &E_S \left[(S_{k+1} - (k+1)S_1) S_{k+1}^k P_{n-k-2}(tS_{k+2}, \ldots, tS_{n-1}) \right] \\ &= n! \int_0^1 \mathrm{d}s_{n-1} \int_0^{s_{n-1}} \mathrm{d}s_{n-2} \ldots \\ &\quad \int_0^{s_2} \mathrm{d}s_1 [s_{k+1} - (k+1)s_1] s_{k+1}^k P_{n-k-2}(ts_{k+2}, \ldots, ts_{n-1}), \end{aligned} \tag{5.34}$$

where the integration over $s_1, \ldots s_{k+1}$ can be carried out explicitly, yielding

$$\int_0^{s_{k+2}} \mathrm{d}s_{k+1} \ldots \int_0^{s_2} \mathrm{d}s_1 [s_{k+1} - (k+1)s_1] s_{k+1}^k$$

$$= \int_0^{s_{k+2}} \mathrm{d}s_{k+1} \left(s_{k+1} \frac{s_{k+1}^k}{k!} - (k+1) \frac{s_{k+1}^{k+1}}{(k+1)!} \right) = 0.$$

Then, all terms of the summation in (5.33) vanish. Finally, noting that $E_S S_1 = (n-1)!/n! = 1/n$, we conclude the proof

$$F_n(t) = (n-1) t E_S S_1 = \frac{n-1}{n} t = \left(1 - \frac{1}{n}\right) t.$$

□

Proof of Proposition 4.5 Let α be a positive integer. As a special case of the stick-breaking representation in Proposition B.1, we have that if $U_1 \dots, U_{n\alpha-1}$ are independent random variables uniformly distributed over the interval [0, 1], then the vector $(X_1, \dots, X_n)$ with components

$$\begin{aligned} X_1 &= U_{(n\alpha-1,\alpha)} \\ X_2 &= U_{(n\alpha-1,2\alpha)} - U_{(n\alpha-1,\alpha)} \\ &\vdots \\ X_{n-1} &= U_{(n\alpha-1,(n-1)\alpha)} - U_{(n\alpha-1,(n-2)\alpha)} \\ X_n &= 1 - U_{(n\alpha-1,(n-1)\alpha)}, \end{aligned}$$

is distributed according to $\mathrm{Dir}(\alpha, \dots, \alpha)$. Therefore, using the same notation as in the proof of Theorem 4.4,

$$\begin{aligned} 1 - F_n(t) &= P\left(T_1 > t S_1, T_2 > t S_2, \dots, T_{n-1} > t S_{n-1}\right) \\ &= P\left(V_{(n\alpha-1,\alpha)} \geq t U_{(n\alpha-1,\alpha)}, V_{(n\alpha-1,2\alpha)}\right. \\ &\geq t U_{(n\alpha-1,2\alpha)}, \dots, V_{(n\alpha-1,(n-1)\alpha)} \\ &\qquad \left.\geq t U_{(n\alpha-1,(n-1)\alpha)}\right), \end{aligned}$$

where $U_1, \dots, U_{n\alpha-1}, V_1, \dots, V_{n\alpha-1}$ are independent uniform random variables in the interval [0, 1]. We can easily upper bound the latter probability

$$\begin{aligned} & P\left(V_{(n\alpha-1,j)} \geq t U_{(n\alpha-1,j)}, \text{for all } j \in \{\alpha, 2\alpha, \dots, (n-1)\alpha\}\right) \\ &\quad \geq P\left(V_{(n\alpha-1,j)} \geq t U_{(n\alpha-1,j)}, \text{for all } j \in \{1, 2, \dots, n\alpha - 1\}\right) \\ &\quad = 1 - \frac{n\alpha - 1}{n\alpha} t, \end{aligned} \tag{5.35}$$

where the last equality follows from Theorem 4.4 applied to pairs of uniform random points in $\Delta_{\alpha n-1}$. □

Appendix A: Proof of Bolshev's recursion

We denote by

$$G_m(t) := \frac{1}{m}\sum_{k=1}^{m} \mathbf{1}_{(-\infty,t]}(U_k), \tag{A.1}$$

the empirical distribution function of m independent random variables $U_1, \ldots, U_m$ uniformly distributed in [0, 1].

The probability of the event $\{U_{(m,1)} > a_1, \ldots, U_{(m,m)} > a_m\}$ can be found by partitioning the sample space into the following disjoint union of events:

$$\{U_{(m,1)} > a_1, \ldots, U_{(m,m)} > a_m\} \cup \bigcup_{k=1}^{n} \{U_{(m,k)} \leq a_k,\ U_{(m,j)} > a_j \text{ for all } j > k\}. \tag{A.2}$$

Informally, we first isolate the event $\{U_{(m,1)} \leq a_1\}$, and then, we partition the complementary event according to the position of $U_{(m,2)}$ into $\{U_{(m,1)} > a_1, U_{(m,2)} \leq a_2\}$ and $\{U_{(m,1)} > a_1, U_{(m,2)} > a_2\}$, and we iterate the procedure up to m. From (A.2), we obtain

$$\begin{aligned} P_m(a_1, \ldots, a_m) &= 1 - \sum_{k=1}^{m} P\left(\{U_{(m,k)} \leq a_k\} \cap \{U_{(m,j)} > a_j \text{ for all } j > k\}\right) \\ &= 1 - \sum_{k=1}^{m} P(mG_m(a_k) = k) P(mG_m(a_j) \leq j-1 \text{ for all } j > k | mG_m(a_k) = k), \end{aligned}$$

where in the last equality we used the empirical spectral distribution (A.1) and the fact that $\{U_{(m,k)} \leq a_k\} \cap \{U_{(m,k+1)} > a_{k+1}\}$ with $a_{k+1} > a_k$ selects the event where exactly k among the m random variables $(U_1, \ldots, U_m)$ fall within $[0, a_k]$. The corresponding probability is given by the binomial distribution $\mathrm{Bin}(m, a_k)$, so that we have:

$$\begin{aligned} P_m(a_1, \ldots, a_m) = 1 - \sum_{k=1}^{m} \binom{m}{k} a_k^k (1-a_k)^{m-k} P(mG_m(a_j) \\ \leq j-1 \text{ for all } j > k | mG_m(a_k) = k) \end{aligned} \tag{A.3}$$

Now notice that

$$\begin{aligned} &P(mG_m(a_j) \leq j-1 \text{ for all } j > k | mG_m(a_k) = k) \\ &\quad = P((m-k)G_{m-k}(a_j) \leq j-1-k \text{ for all } j > k | (m-k)G_{m-k}(a_k) = 0), \end{aligned} \tag{A.4}$$

that is, the locations of the largest $(m-k)$ among $(U_1, \ldots, U_m)$ given that k of them fall below a_k are identically distributed to the locations of the order statistics of $(m-k)$ uniform random variables $(V_1, \ldots, V_{m-k})$, given that *none* of them falls below a_k. Note also that $(1-a_k)^{m-k}$ corresponds to the probability that among $m-k$ uniform random variables in [0, 1] none of them falls in $[0, a_k]$. Therefore,

$$\begin{aligned}(1-a_k)^{m-k} &P(mG_m(a_j) \le j-1 \text{ for all } j > k | mG_m(a_k) = k) \\ &= P((m-k)G_{m-k}(a_j) \le j-1-k \text{ for all } j > k) \\ &= P(V_{(m-k,1)} > a_{k+1}, \ldots, V_{(m-k,n-k)} > a_m) \\ &= P_{m-k}(a_{k+1}, \ldots, a_m),\end{aligned} \tag{A.5}$$

Substitution of (A.5) into (A.3) yields the claimed recursive relation.

Appendix B: Dirichlet distribution and stick-breaking representation

The Dirichlet measure with parameters $\alpha_1, \ldots, \alpha_n > 0$, denoted $\mathrm{Dir}(\alpha_1, \ldots, \alpha_n)$, is the probability measure on the unit simplex Δ_{n-1} with density

$$f_{X_1,\ldots,X_n}(x) = \frac{\Gamma(\alpha_1 + \cdots + \alpha_n)}{\Gamma(\alpha_1)\cdots\Gamma(\alpha_n)} \prod_{i=1}^{n} x_i^{\alpha_i - 1} \mathbf{1}_{\Delta_n}(x). \tag{B.1}$$

For $n > 2$, the Dirichlet distribution is a natural multidimensional generalisation of the Beta distribution.

When $\alpha_1, \ldots, \alpha_n$ are positive integers, a generalisation of Proposition 5.2 gives a nice and useful representation of a Dirichlet-distributed random vector in terms of spacings of order statistics of independent uniform random variables in the interval [0, 1].

Proposition B.1 (Stick-breaking representation of the Dirichlet distribution) *Let $\alpha_1, \ldots, \alpha_n$ be positive integers. Denote $m_1 = \alpha_1, m_2 = \alpha_1 + \alpha_2, \ldots, m_{n-1} = \alpha_1 + \cdots + \alpha_{n-1}$, and $m = \alpha_1 + \cdots + \alpha_n$. Let $U_1, \ldots, U_{m-1}$ be independent random variables that are uniformly distributed over the interval* [0, 1] *and $U_{(m-1,1)} \le \cdots \le U_{(m-1,m-1)}$ the corresponding order statistics. Then, the vector $(X_1, X_2, \ldots, X_n)$ of the n spacings*

$$\begin{aligned} X_1 &= U_{(m-1,m_1)} \\ X_2 &= U_{(m-1,m_2)} - U_{(m-1,m_1)} \\ &\vdots \\ X_{n-1} &= U_{(m-1,m_{n-1})} - U_{(m-1,m_{n-2})} \\ X_n &= 1 - U_{(m-1,m_{n-1})} \end{aligned}$$

is distributed according to $\mathrm{Dir}(\alpha_1, \ldots, \alpha_n)$.

Acknowledgements We are grateful to F. vom Ende for bibliographical remarks, to M. Michelen for precious insights leading to Theorem 4.2, and to an anonymous referee for valuable suggestions that significantly improved the paper.

Data availability Data sharing was not applicable to this article as no datasets were generated or analysed during the current study.

Declarations

Conflict of interest The authors have no conflict of interest to declare that are relevant to the content of this article.

References

1. Arnold, B.C., Balakrishnan, N., Nagaraja, H.N.: A first course in order statistics. SIAM (2008)
2. Andersen, E.S.: On the fluctuations of sums of random variables. Math Scand **1**(2), 263–285 (1953)
3. Alhambra, A.M., Oppenheim, J., Perry, C.: Fluctuating states: What is the probability of a thermodynamical transition? Phys. Rev. X **6**, 041016 (2016)
4. Alberti, P.M., Uhlmann, A.: Stochasticity and partial order. Deutscher Verlag der Wissenschaften Berlin (1982)
5. Birkhoff, G.: Tres observaciones sobre el algebra lineal. Univ. Nac. Tucuman. Ser. A **5**, 147–154 (1946)
6. Cunden, F.D., Facchi, P., Florio, G., Gramegna, G.: Volume of the set of locc-convertible quantum states. J. Phys. A: Math. Theor. **53**, 175303 (2020)
7. Cunden, F.D., Facchi, P., Florio, G., Gramegna, G.: Generic aspects of the resource theory of quantum coherence. Phys. Rev. A **103**, 022401 (2021)
8. Chitambar, E., Gour, G.: Comparison of incoherent operations and measures of coherence. Phys. Rev. A **94**, 052336 (2016)
9. Chitambar, E., Gour, G.: Quantum resource theories. Rev. Mod. Phys. **91**, 025001 (2019)
10. Du, S., Bai, Z., Guo, Y.: Conditions for coherence transformations under incoherent operations. Phys. Rev. A **91**, 052120 (2015)
11. Devroye, L.: Non-Uniform Random Variate Generation. Springer, New York (1986)
12. Denuit, M., Lefèvre, C., Picard, P.: Polynomial structures in order statistics distributions. J. Stat. Plan. Inference **113**, 151–178 (2003)
13. David, H.A., Nagaraja, H.N.: Order statistics. John Wiley & Sons (2004)
14. de Oliveira Junior, A., Czartowski, J., Życzkowski, K., Korzekwa, K.: Geometric structure of thermal cones. Phys. Rev. E **106**, 064109 (2022)
15. Hardy, G., Littlewood, J., Pólya, G.: Inequalities. Cambridge University Press (1952)
16. Horodecki, M., Oppenheim, J.: Fundamental limitations for quantum and nanoscale thermodynamics. Nat. Commun **4**, 2059 (2013)
17. Hajós, G., Rényi, A.: Elementary proofs of some basic facts concerning order statistics. Acta Math. Acad. Sci. Hungar **5**, 1–6 (1954)
18. Jain, V., Kwan, M., Michelen, M.: Entangled states are typically incomparable. arXiv:2406.03335, (2024)
19. Joe, H.: Majorization and divergence. J. Math. Anal. Appl. **148**, 287–305 (1990)
20. Kotel'Nikova, V., Chmaladze, E.: On computing the probability of an empirical process not crossing a curvilinear boundary. Theory Probab. & Its Appl. **27**, 640–648 (1983)

21. Malamud, S.: Inverse spectral problem for normal matrices and the Gauss-Lucas theorem. Trans. Am. Math. Soc **357**, 4043–4064 (2005)
22. Mirsky, L.: On a convex set of matrices. Archiv der Mathematik **10**, 88–92 (1959)
23. Marshall, A.W., Olkin, I., Arnold, B.C.: Inequalities: theory of majorization and its applications, volume 143. Springer (1979)
24. Narayana, T.V.: A partial order and its applications to probability theory, Sankhya: Indian J. Stat., 91–98 (1959)
25. Narasimhachar, V., Gour, G.: Low-temperature thermodynamics with quantum coherence. Nat. Commun. **6**, 1–6 (2015)
26. Nielsen, M.A.: Conditions for a class of entanglement transformations. Phys. Rev. Lett. **83**, 436 (1999)
27. Nielsen, M.A., Vidal, G.: Majorization and the interconversion of bipartite states. Quantum Inf. Comput. **1**, 76–93 (2001)
28. Oszmaniec, M., Biswas, T.: Operational relevance of resource theories of quantum measurements. Quantum **3**, 133 (2019)
29. OEIS Foundation Inc. The On-Line Encyclopedia of Integer Sequences. Published electronically at http://oeis.org/A002720
30. Parker, D.S., Ram, P.: Greed and majorization. University of California (Los Angeles). Computer Science Department (1996)
31. Rényi, A.: Proceedings of the fourth Berkeley symposium on mathematical statistics and probability. Contributions to Biology and Problems of Medicine **4**, 547–61 (1961)
32. Streltsov, A., Adesso, G., Plenio, M.B.: Colloquium: quantum coherence as a resource. Rev. Mod. Phys. **89**, 041003 (2017)
33. Shiraishi, N.: Two constructive proofs on d-majorization and thermo-majorization. J. Phys. A: Math. Theor. **53**, 425301 (2020)
34. Stanley, R.P., Pitman, J.: A polytope related to empirical distributions, plane trees, parking functions, and the associahedron. Discrete & Computational Geometry **27**, 602–603 (2002)
35. Shorack, G.R., Wellner, J.A.: Empirical processes with applications to statistics. SIAM (2009)
36. vom Ende, F., Dirr, G.: The d-majorization polytope. Linear Algebr. Its Appl. **649**, 152–185 (2022)
37. vom Ende, F., Malvetti, E.: The thermomajorization polytope and its degeneracies. Entropy **26**, 106 (January 2024)
38. Vershynin, R.: High-dimensional probability: an introduction with applications in data science, volume 47. Cambridge university press (2018)
39. Vidal, G.: Entanglement of pure states for a single copy. Phys. Rev. Lett. **83**, 1046 (1999)
40. Zhu, H., Ma, Z., Cao, Z., Fei, S.-M., Vedral, V.: Operational one-to-one mapping between coherence and entanglement measures. Phys. Rev. A **96**, 032316 (2017)

Authors and Affiliations

Fabio Deelan Cunden[1,2] · Jakub Czartowski[3,4,5] · Giovanni Gramegna[2,6] · A. de Oliveira Junior[7]

Fabio Deelan Cunden
fabio.cunden@uniba.it

Jakub Czartowski
jakub.czartowski@ntu.edu.sg

Giovanni Gramegna
giovanni.gramegna@uniba.it

A. de Oliveira Junior
aleoli@dtu.dk

[1] Dipartimento di Matematica, Universitá degli Studi di Bari, 70125 Bari, Italy

[2] INFN, Sezione di Bari, 70126 Bari, Italy

[3] Doctoral School of Exact and Natural Sciences, Jagiellonian University, 30-348 Kraków, Poland

[4] Faculty of Physics, Astronomy and Applied Computer Science, Jagiellonian University, 30-348 Kraków, Poland

[5] School of Physical and Mathematical Sciences, Nanyang Technological University, 21 Nanyang Link, Singapore 637371, Republic of Singapore

[6] Dipartimento di Fisica, Università degli Studi di Bari, 70126 Bari, Italy

[7] Center for Macroscopic Quantum States bigQ, Department of Physics, Technical University of Denmark, Fysikvej 307, 2800 Kongens Lyngby, Denmark