A CERTIFIED UPPER BOUND FOR THE BEAT THE AVERAGE GAME

TEJASVI SINGH TOMAR

Abstract. Let C be the supremum of $\mathbb{P}[X_1+X_2+X_3<2X_4]$ over all probability measures on $[0,\infty)$, where $X_1,\dots,X_4$ are independent with the given law. Determining C is Problem 6.39, the beat the average game, in the list of Georgiev, Gómez-Serrano, Tao and Wagner. Bellec and Fritz proved $0.4006953 \le C \le 673/1615$ and conjectured that the lower bound, $0.4006953723\ldots$, is the exact value. We prove $C \le 0.40069634$, less than 10^{-6} above the conjectured value. The proof applies Razborov's flag algebra method to samples of nine points. Since the points are independent, the positivity constraints of flag algebras hold for every single measure, so that no limit objects are needed. An exact certificate is checked in rational arithmetic, and the bound is also proved in the Lean 4 proof assistant, with Lean's kernel checking the certificate.

hexagon:2611.00476v1

1. Introduction

For a probability measure μ on $[0,\infty)$ let $X_1, X_2, \dots$ be independent random variables with law μ, and put

$$P(\mu) = \mathbb{P}[X_1+X_2+X_3<2X_4], \qquad C = \sup_{\mu} P(\mu).$$

Fritz asked for the value of C on MathOverflow [4]. Determining C is Problem 6.39 in the list of Georgiev, Gómez-Serrano, Tao and Wagner [5], who call it the beat the average game. The name comes from a game with three dice, in which the casino maximizes $\mathbb{P}[2X_1 < X_2+X_3]$ [2, Section 1]. Bellec and Fritz [2] proved

$$0.4006953\ldots = \max_{0<p<1} \frac{p(2-p)}{1+p+p^2+p^3} \;\le\; C \;\le\; \frac{673}{1615} < 0.41672. \tag{1.1}$$

Their lower bound comes from measures with an atom of mass p at 0 and equal masses at the points $1-2^{-i}$ for $1 \le i \le N$, and a scaling construction repeats this configuration at smaller and smaller scales [2, Propositions 3.3 and 4.1]. The maximum in (1.1) is attained at the root $p^* = 0.4743464\ldots$ of $p^4-4p^3-3p^2-2p+2$ in $(0,1)$ [2, Remark 4.2]. Yu [13] pointed out that this value of p improves on $p=1/2$. Bellec and Fritz conjectured that the lower bound in (1.1) is the exact value of C [2, Conjecture 5.1]. A computer search with the method of [10] found a measure with $P(\mu) \ge 0.389$ [5].

The upper bound in (1.1) rests on a lemma that Bellec and Fritz attribute to a MathOverflow answer of Lacker [2, Lemma 2.3]. For every $m \ge 4$, it bounds C by the largest fraction of the $4\binom{m}{4}$ inequalities $x_a+x_b+x_c < 2x_d$ among m nonnegative reals that can hold simultaneously. Bellec and Fritz computed this fraction for $m=20$ by mixed-integer linear programming. For $m=25$ their computation gives $20898/50600 < 0.41301$, but they do not have a proof of this bound [2, Section 4].

Theorem 1.1. *For every probability measure μ on $[0,\infty)$ we have $P(\mu) \le 20034817/50000000$. Together with* (1.1)*, this gives $0.4006953 \le C \le 0.40069634$.*

2020 *Mathematics Subject Classification.* 60E15; 90C22, 52C35, 68V05, 68V20.

Key words and phrases. beat the average game, flag algebras, semidefinite programming, Gordan certificates, computer-assisted proof, Lean 4.

The upper bound of Theorem 1.1 exceeds the conjectured value $C^* = 0.4006953723\ldots$ by $9.6767\ldots \cdot 10^{-7}$. With six, seven and eight points the same method gives the bounds $C \le 0.402546$, $C \le 0.40076$ and $C \le 0.4006983$ (Table 1).

The method. By Lemma 2.1, C is already the supremum of $P(\mu)$ over the absolutely continuous measures μ, so we may assume that μ is absolutely continuous. We sample k points from μ, sort them as $y_1 < \cdots < y_k$, and record which of the $4\binom{k}{4}$ inequalities $2y_d > y_a + y_b + y_c$ with distinct indices hold. This is the sign pattern of the sample. A pattern is realizable if it is the sign pattern of some points $0 < y_1 < \cdots < y_k$ at which none of the inequalities is an equality. Then $P(\mu)$ is the expected fraction of the inequalities that hold (Lemma 2.3), and Lacker's lemma bounds this expectation by its maximum over the realizable patterns. This bound ignores that the sample is independent. Fix t of the sample points and two disjoint groups of $m - t$ further ones, all chosen by their indices before sorting; this needs $2m - t \le k$. Given the fixed points, the two groups are independent and have the same law. Hence the matrix of the probabilities that the first group forms one sign pattern with the fixed points and the second group another is positive semidefinite. This is the positivity used in Razborov's flag algebras [11], and for independent samples it holds for every single measure (Lemma 3.1). We combine these constraints in a semidefinite program (Theorem 3.3 and Section 5). Its solution consists of positive semidefinite matrices, called blocks, each indexed by flags (sign patterns of m points of which t are marked) whose marked points have the same pattern, and it gives each pattern a value: the fraction of its true inequalities plus correction terms from the blocks. The largest value over the realizable patterns is an upper bound for C.

For the proof of Theorem 1.1 we take $k = 9$. The certificate is arranged so that the exact check runs on patterns of 8 points, apart from two small classes of patterns of 9 points (Lemma 5.2), and Gordan certificates, which show that systems of strict linear inequalities have no solution, prove that the patterns that would violate the bound cannot occur (Sections 4 and 5). The certificate is verified by two independently written programs in exact arithmetic and, in a formal proof in the Lean 4 proof assistant that follows Sections 2 to 5, by Lean's kernel (Section 6).

Sections 2 to 5 give the proof, Section 6 the formal proofs, and Section 7 compares the dual solution with the construction of Bellec and Fritz and asks whether a certificate of this kind with some finite k reaches C^*.

2. Sign patterns of samples

We first express $P(\mu)$ through the distribution of a finite combinatorial object. For $k \ge 1$ write $[k] = \{1, \ldots, k\}$ and

$$T_k = \big\{(d; a, b, c) :\ d \in [k],\ a < b < c,\ \{a, b, c\} \subseteq [k] \setminus \{d\}\big\}, \qquad L_\tau(x) = 2x_d - x_a - x_b - x_c$$

for $\tau = (d; a, b, c) \in T_k$ and $x \in \mathbb{R}^k$. Then $|T_k| = 4\binom{k}{4}$; in particular, T_k is empty for $k \le 3$. A point $x \in \mathbb{R}^k$ is *generic* if its coordinates are positive and pairwise distinct and $L_\tau(x) \ne 0$ for all $\tau \in T_k$. A *k-pattern* is a vector $M \in \{0, 1\}^{T_k}$. The *pattern* $\pi(x)$ of a generic point x has $\pi(x)_\tau = 1$ if $L_\tau(y) > 0$ and $\pi(x)_\tau = 0$ otherwise, where y is the increasing rearrangement of x. A pattern is *realizable* if it is the pattern of a generic point, and $\mathcal{R}_k$ denotes the set of realizable k-patterns. The *density* of a k-pattern with $k \ge 4$ is $d(M) = |T_k|^{-1} \sum_\tau M_\tau$. For a set $U = \{u_1 < \cdots < u_m\} \subseteq [k]$ the *restriction* $M|_U$ is the m-pattern with $(M|_U)_{(d;a,b,c)} = M_{(u_d;u_a,u_b,u_c)}$. If x is generic with increasing rearrangement y, then $\pi(x)|_U$ is the pattern of $(y_{u_1}, \ldots, y_{u_m})$.

Lemma 2.1. *C is the supremum of $P(\mu)$ over the absolutely continuous probability measures μ with $\mu((0, \infty)) = 1$.*

Proof. Let μ be a probability measure on $[0,\infty)$ and $\varepsilon > 0$. Let $U_1, \dots, U_4$ be uniform on $[0,1]$, with $X_1, \dots, X_4, U_1, \dots, U_4$ independent, and put $Y_i = X_i + \varepsilon(1 + U_i)$. The Y_i are independent with a common law μ_ε, which is absolutely continuous and satisfies $\mu_\varepsilon([\varepsilon,\infty)) = 1$. We have

$$Y_1 + Y_2 + Y_3 - 2Y_4 = X_1 + X_2 + X_3 - 2X_4 + \varepsilon(1 + U_1 + U_2 + U_3 - 2U_4),$$

and the last term is at most 4ε. Hence $P(\mu_\varepsilon) \geq \mathbb{P}[X_1 + X_2 + X_3 - 2X_4 < -4\varepsilon]$, and the right side tends to $P(\mu)$ as $\varepsilon \to 0$. So the supremum in the lemma is at least $P(\mu)$. The reverse inequality holds because the supremum runs over a subset of the measures in the definition of C. □

Lemma 2.2. *Let μ have no atoms and satisfy $\mu((0,\infty)) = 1$. Then $(X_1, \dots, X_k)$ is generic almost surely.*

Proof. For $i \neq j$ we have $\mathbb{P}[X_i = X_j] = \int \mu(\{x\})\, d\mu(x) = 0$. For $\tau = (d; a, b, c)$ the variable X_d is independent of (X_a, X_b, X_c), so $\mathbb{P}[L_\tau(X) = 0] = \mathbb{E}\,\mu(\{(X_a + X_b + X_c)/2\}) = 0$. The coordinates are positive almost surely because $\mu((0,\infty)) = 1$. □

For such μ we write $p_\mu(M) = \mathbb{P}[\pi(X_1, \dots, X_k) = M]$. By Lemma 2.2, p_μ is a probability vector supported on $\mathcal{R}_k$.

Lemma 2.3. *Let $k \geq 4$, and let μ have no atoms and satisfy $\mu((0,\infty)) = 1$. Then $P(\mu) = \sum_M p_\mu(M)\, d(M)$.*

Proof. Let $x \in \mathbb{R}^k$ be generic and let σ be the permutation of $[k]$ with $x_{\sigma(1)} < \cdots < x_{\sigma(k)}$, so that $y = (x_{\sigma(1)}, \dots, x_{\sigma(k)})$ is the increasing rearrangement of x. For $\tau = (d; a, b, c)$ let $\sigma(\tau) \in T_k$ be the tuple with first entry $\sigma(d)$ and remaining entries $\sigma(a), \sigma(b), \sigma(c)$ in increasing order. Then $\tau \mapsto \sigma(\tau)$ is a bijection of T_k and $L_\tau(y) = L_{\sigma(\tau)}(x)$. Hence $|T_k|\, d(\pi(x))$ is the number of $\tau \in T_k$ with $L_\tau(x) > 0$. We apply this to $X = (X_1, \dots, X_k)$ and take expectations:

$$|T_k| \sum_M p_\mu(M)\, d(M) = \sum_{\tau \in T_k} \mathbb{P}[L_\tau(X) > 0].$$

For $\tau = (d; a, b, c)$ the vector (X_a, X_b, X_c, X_d) has the law of (X_1, X_2, X_3, X_4), so every term on the right equals $P(\mu)$. □

Since p_μ is supported on $\mathcal{R}_k$, Lemmas 2.1 and 2.3 give $C \leq \max_{M \in \mathcal{R}_k} d(M)$. This is the bound of Lacker's lemma with $m = k$, because a point of $[0,\infty)^k$ can be perturbed to a generic point without violating the strict inequalities that it satisfies.

Remark 2.4. Two features of sorted patterns matter below. First, a permutation of the sample points permutes T_k, so the fraction of the inequalities that hold, and the flag counts of Section 3, depend on a sample only through its pattern (proofs of Lemmas 2.3 and 3.1). Second, realizability refers to points with positive coordinates, and this condition cannot be dropped. Indeed, L_τ is not invariant under translations. If $Z_1, \dots, Z_4$ are independent and identically distributed, then

$$\mathbb{P}[(Z_1 - c) + (Z_2 - c) + (Z_3 - c) < 2(Z_4 - c)] = \mathbb{P}[Z_1 + Z_2 + Z_3 - 2Z_4 < c],$$

which tends to 1 as $c \to \infty$. For this reason the inequality $x_1 > 0$ appears in every system of Section 4.

3. Positivity from independence

We now derive the constraints that independence imposes on p_μ. Fix integers $0 \leq t \leq m$. A *flag* of size m with t marked points is a pair $F = (N, J)$ of an m-pattern N and a set $J \subseteq [m]$

with $|J| = t$. Its *type* is the t-pattern $N|_J$. Let $k \geq 2m - t$ and let M be a k-pattern. For disjoint sets $J, A \subseteq [k]$ with $|J| = t$ and $|A| = m - t$ put

$$\mathrm{fl}(M; J, A) = (M|_{J\cup A}, J'),$$

where $J' \subseteq [m]$ is the set of positions of the elements of J in the increasing enumeration of $J \cup A$. For flags F, G of size m with t marked points let

$$c(F, G; M) = \frac{1}{D} \#\big\{(J, A, B) :\ \mathrm{fl}(M; J, A) = F,\ \mathrm{fl}(M; J, B) = G\big\}, \tag{3.1}$$

where (J, A, B) runs over the triples of pairwise disjoint subsets of $[k]$ with $|J| = t$ and $|A| = |B| = m - t$, and $D = \binom{k}{t}\binom{k-t}{m-t}\binom{k-m}{m-t}$ is the number of these triples. Since $\mathrm{fl}(M; J, A)$ and $\mathrm{fl}(M; J, B)$ both have type $M|_J$, we have $c(F, G; M) = 0$ unless F and G have the same type. Exchanging A and B shows that $c(F, G; M) = c(G, F; M)$.

Lemma 3.1. *Let μ have no atoms and satisfy $\mu((0,\infty)) = 1$, and let $k \geq 2m - t$. Let $\mathcal{F}$ be a finite set of flags of size m with t marked points, and let $Q = (Q_{FG})_{F,G\in\mathcal{F}}$ be a positive semidefinite matrix. Then*

$$\sum_M p_\mu(M) \sum_{F,G\in\mathcal{F}} Q_{FG}\, c(F, G; M) \ \geq\ 0.$$

Proof. For a point z with positive coordinates indexed by a finite set I, and for disjoint $J, A \subseteq I$ with $|J| = t$ and $|A| = m - t$, we define $g(z; J, A)$ as follows. If the values z_i with $i \in J \cup A$ are pairwise distinct and, listed in increasing order, form a generic point $w \in \mathbb{R}^m$, then $g(z; J, A) = (\pi(w), J')$, where $J' \subseteq [m]$ is the set of positions of the values z_j with $j \in J$ in this list. Otherwise $g(z; J, A) = *$, a symbol that is not a flag.

Let $x \in \mathbb{R}^k$ be generic and let σ be the permutation of $[k]$ with $x_{\sigma(1)} < \cdots < x_{\sigma(k)}$. Then $\mathrm{fl}(\pi(x); J, A) = g(x; \sigma(J), \sigma(A))$, since both sides are read off from the increasing list of the values $x_{\sigma(i)}$ with $i \in J \cup A$. As σ permutes the triples (J, A, B),

$$c(F, G; \pi(x)) = \frac{1}{D} \#\big\{(J, A, B) :\ g(x; J, A) = F,\ g(x; J, B) = G\big\}. \tag{3.2}$$

Let $X = (X_1, \ldots, X_k)$. By Lemma 2.2 and (3.2),

$$\sum_M p_\mu(M)\, c(F, G; M) = \frac{1}{D} \sum_{(J,A,B)} \mathbb{P}\big[g(X; J, A) = F,\ g(X; J, B) = G\big].$$

Put $J_0 = \{1, \ldots, t\}$, $A_0 = \{t+1, \ldots, m\}$ and $B_0 = \{m+1, \ldots, 2m-t\}$. For each triple (J, A, B) there is a permutation ρ of $[k]$ with $\rho(J_0) = J$, $\rho(A_0) = A$ and $\rho(B_0) = B$. The vector $X \circ \rho = (X_{\rho(1)}, \ldots, X_{\rho(k)})$ has the same law as X, and $g(X; J, A) = g(X \circ \rho; J_0, A_0)$ and $g(X; J, B) = g(X \circ \rho; J_0, B_0)$. Hence each term of the sum equals $\mathbb{P}[g(X; J_0, A_0) = F,\ g(X; J_0, B_0) = G]$.

The vectors $W = (X_1, \ldots, X_t)$, $Y = (X_{t+1}, \ldots, X_m)$ and $Y' = (X_{m+1}, \ldots, X_{2m-t})$ are independent, and Y and Y' have the same law. For $w \in (0,\infty)^t$ let $\varphi_F(w) = \mathbb{P}[g((w, Y); J_0, A_0) = F]$. The flag $g(X; J_0, A_0)$ depends only on (W, Y), and $g(X; J_0, B_0)$ is the same function of (W, Y'). By Fubini's theorem,

$$\mathbb{P}\big[g(X; J_0, A_0) = F,\ g(X; J_0, B_0) = G\big] = \mathbb{E}\big[\varphi_F(W)\, \varphi_G(W)\big].$$

Therefore the sum in the lemma equals $\mathbb{E}\big[\sum_{F,G\in\mathcal{F}} Q_{FG}\, \varphi_F(W)\varphi_G(W)\big]$, which is nonnegative because Q is positive semidefinite. □

Remark 3.2. The flags in Lemma 3.1 need not be realizable or share a type. Up to normalization, $c(F, G; M)$ is the coefficient of M in the averaged product of F and G in Razborov's flag algebra [11], and Lemma 3.1 is the counterpart of the positivity of averaged squares. Razborov

proves this positivity for limits of sequences of finite structures; Lemma 3.1 needs neither limit objects nor de Finetti's theorem. Our structures are linearly ordered and have no nontrivial automorphisms, as in the flag algebras of permutations [1, 12]. This is why a flag can be written as a pattern together with a set of positions.

Theorem 3.3. *Let $k \geq 4$, and let $\mathcal{P}$ be a set of k-patterns with $\mathcal{R}_k \subseteq \mathcal{P}$. For $b = 1, \dots, r$ let $\mathcal{F}_b$ be a finite set of flags of size m_b with t_b marked points, where $2m_b - t_b \leq k$, and let Q_b be a positive semidefinite matrix indexed by $\mathcal{F}_b$. For a k-pattern M put*

$$\operatorname{val}(M) = d(M) + \sum_{b=1}^{r} \sum_{F,G \in \mathcal{F}_b} (Q_b)_{FG}\, c(F, G; M). \tag{3.3}$$

Then $C \leq \max_{M \in \mathcal{P}} \operatorname{val}(M)$.

Proof. Let μ be absolutely continuous with $\mu((0,\infty)) = 1$. By Lemma 2.3 and by Lemma 3.1 for each b,

$$P(\mu) = \sum_M p_\mu(M)\, d(M) \leq \sum_M p_\mu(M) \operatorname{val}(M) \leq \max_{M \in \mathcal{P}} \operatorname{val}(M),$$

where the last step holds because p_μ is a probability vector supported on $\mathcal{R}_k \subseteq \mathcal{P}$. Lemma 2.1 completes the proof. □

Patterns in $\mathcal{P} \setminus \mathcal{R}_k$ can only weaken the bound. Hence it suffices to know a superset of $\mathcal{R}_k$, and we construct one next.

4. Completeness of the list of patterns

We build sets $\mathcal{P}_k \supseteq \mathcal{R}_k$ by adding one point at a time. Restricting a generic point to a subset of its coordinates gives a generic point, so $M|_U \in \mathcal{R}_{|U|}$ for all $M \in \mathcal{R}_k$ and $U \subseteq [k]$. In particular each $M \in \mathcal{R}_k$ extends the pattern $M|_{[k-1]} \in \mathcal{R}_{k-1}$ by its values on the $4\binom{k-1}{3}$ tuples of $T_k \setminus T_{k-1}$, the tuples that contain k.

Let P be a $(k-1)$-pattern, let $n = |T_k \setminus T_{k-1}|$, and fix an enumeration $\tau_1, \dots, \tau_n$ of $T_k \setminus T_{k-1}$, which may depend on P. For $s \in \{0,1\}^i$ with $0 \leq i \leq n$ let $S_P(s)$ be the system of strict inequalities

$$x_1 > 0, \qquad x_{j+1} - x_j > 0 \ \ (1 \leq j < k),$$
$$(2P_\tau - 1)L_\tau(x) > 0 \ \ (\tau \in T_{k-1}), \qquad (2s_j - 1)L_{\tau_j}(x) > 0 \ \ (1 \leq j \leq i)$$

in $x \in \mathbb{R}^k$. Each inequality has the form $r \cdot x > 0$ with $r \in \mathbb{Z}^k$. A *Gordan certificate* for a system of such inequalities consists of nonnegative integers y_r, one for each inequality and not all zero, with $\sum_r y_r\, r = 0$. A system with a Gordan certificate has no solution, since a solution x would give $0 = \sum_r y_r\, (r \cdot x) > 0$. Conversely, by Gordan's theorem [6], every system of strict homogeneous linear inequalities with rational coefficients and without a solution has a certificate with rational entries, and hence one with integer entries.

An *extension tree* for P is a set $\mathcal{T}$ of sign vectors of lengths $0, \dots, n$ that contains the empty vector and all prefixes of its elements, such that for each $s \in \mathcal{T}$ of length $i < n$ and each $e \in \{0,1\}$ either $(s,e) \in \mathcal{T}$ or $S_P(s,e)$ has a Gordan certificate. A vector $s \in \mathcal{T}$ of length n is a *leaf*, and it defines the k-pattern that agrees with P on T_{k-1} and takes the value s_j at τ_j.

Proposition 4.1. *Let $\mathcal{P}_{k-1}$ be a set of $(k-1)$-patterns with $\mathcal{R}_{k-1} \subseteq \mathcal{P}_{k-1}$, and fix an extension tree for each $P \in \mathcal{P}_{k-1}$. Then the set $\mathcal{P}_k$ of the patterns defined by the leaves of these trees satisfies $\mathcal{R}_k \subseteq \mathcal{P}_k$.*

Proof. Let $M \in \mathcal{R}_k$. Then $M = \pi(x)$ for a generic x with $0 < x_1 < \cdots < x_k$, and $P = M|_{[k-1]} \in \mathcal{R}_{k-1} \subseteq \mathcal{P}_{k-1}$. Put $s_j = M_{\tau_j}$ in the enumeration of the tree for P. For each i, the point x solves $S_P(s_1, \ldots, s_i)$, so this system has no Gordan certificate. By induction on i, the vector $(s_1, \ldots, s_i)$ belongs to the tree for all $i \le n$. The leaf $(s_1, \ldots, s_n)$ defines M. □

Starting from the set $\mathcal{P}_3 = \mathcal{R}_3$ of the single empty pattern, Proposition 4.1 produces sets $\mathcal{P}_4, \ldots, \mathcal{P}_7$ that contain all realizable patterns. We computed the trees by linear programming with HiGHS [7], and kept a leaf only together with an integer point that realizes its pattern.

A second enumeration walks through the cells of the arrangement of the hyperplanes $L_\tau = 0$ inside the cone $0 < x_1 < \cdots < x_k$ by changing one sign at a time. Since these hyperplanes are pairwise distinct and distinct from the walls of the cone, a generic segment between two cells crosses them one at a time; hence the walk reaches every cell. Each step is settled in one of three ways: by an integer point, by a Gordan certificate, or, for $k = 7$, by a restriction to 6 points that is not in the list that the walk found for $k = 6$. It finds the same sets for $k \le 7$, with an integer point for each pattern. We check these points against the patterns rebuilt from the trees. Hence $\mathcal{P}_k = \mathcal{R}_k$ for $k \le 7$, although the proof uses only $\mathcal{R}_k \subseteq \mathcal{P}_k$.

The maximal density is $1/2$ on $\mathcal{P}_4$ and on $\mathcal{P}_5$, and Table 1 gives it for $k = 6, 7, 8$. As a check, the value $7/15$ for $k = 6$ agrees with the computation of Bellec and Fritz for $m = 6$ [2, Section 4].

Eight and nine points. For $k \ge 8$ we do not build extension trees; instead, we form the candidates by joins, as described below. For a k-pattern M let $S(M)$ be the system

$$x_1 > 0, \qquad x_{j+1} - x_j > 0 \ \ (1 \le j < k), \qquad (2M_\tau - 1)L_\tau(x) > 0 \ \ (\tau \in T_k).$$

If $M = \pi(x)$ for a generic point x with $0 < x_1 < \cdots < x_k$, then x solves $S(M)$. Hence a k-pattern M for which $S(M)$ has a Gordan certificate is not realizable. Let $\mathcal{J}_8$ be the set of 8-patterns M with $M|_U \in \mathcal{P}_7$ for every 7-element set $U \subseteq [8]$.

Lemma 4.2. *Let $\mathcal{R}_7 \subseteq \mathcal{P}_7$, and let E be a set of 8-patterns M such that $S(M)$ has a Gordan certificate. Then $\mathcal{P}_8 = \mathcal{J}_8 \setminus E$ satisfies $\mathcal{R}_8 \subseteq \mathcal{P}_8$.*

Proof. Let $M \in \mathcal{R}_8$, realized by a generic point x with $0 < x_1 < \cdots < x_8$. Then $M|_U \in \mathcal{R}_7 \subseteq \mathcal{P}_7$ for every 7-element set U, so $M \in \mathcal{J}_8$. The point x solves $S(M)$, so $M \notin E$. □

A k-pattern is determined by its restrictions to the $(k-1)$-element subsets of $[k]$, and these restrictions satisfy compatibility conditions that also suffice.

Lemma 4.3. *Let $k \ge 5$.*

(a) *For a k-pattern M and $1 \le i < j \le k$, the restriction of $M|_{[k]\setminus\{i\}}$ to $[k-1] \setminus \{j-1\}$ and the restriction of $M|_{[k]\setminus\{j\}}$ to $[k-1] \setminus \{i\}$ both equal $M|_{[k]\setminus\{i,j\}}$.*

(b) *Let $N_1, \ldots, N_k$ be $(k-1)$-patterns with $N_i|_{[k-1]\setminus\{j-1\}} = N_j|_{[k-1]\setminus\{i\}}$ for all $1 \le i < j \le k$. Then there is exactly one k-pattern M with $M|_{[k]\setminus\{i\}} = N_i$ for all $i \in [k]$.*

Proof. If $U = \{u_1 < \cdots < u_m\} \subseteq [k]$ and $V \subseteq [m]$, then $(M|_U)|_V = M|_{\{u_v : v \in V\}}$ by the definition of restrictions. For $U = [k] \setminus \{i\}$ the point j is u_{j-1}, and for $U = [k] \setminus \{j\}$ the point i is u_i. This proves (a). For (b), let $\tau \in T_k$. Since τ involves four points and $k \ge 5$, some $u \in [k]$ does not occur in τ. Let M_τ be the entry of N_u at the tuple that τ becomes when the points of $[k] \setminus \{u\}$ are numbered increasingly. If $u < v$ both do not occur in τ, then the entries of N_u and N_v obtained in this way are the entries of $N_u|_{[k-1]\setminus\{v-1\}}$ and $N_v|_{[k-1]\setminus\{u\}}$ at the tuple that τ becomes in $[k] \setminus \{u, v\}$. These two patterns are equal by hypothesis, so M_τ does not depend on the choice of u, and $M|_{[k]\setminus\{u\}} = N_u$ for every u. Every tuple of T_k misses some point, so M is the only k-pattern with these restrictions. □

We obtain $\mathcal{J}_8$ by joining $\mathcal{P}_7$ with itself. By Lemma 4.3, the patterns of $\mathcal{J}_8$ correspond to the families $(N_1, \dots, N_8)$ in $\mathcal{P}_7$ that satisfy the conditions of Lemma 4.3(b), through $N_i = M|_{[8]\setminus\{i\}}$. We let N_8 run over $\mathcal{P}_7$. For $i \in [7]$ the condition with $j = 8$ says that the restriction of N_i to $[6]$, which corresponds to the points $[7] \setminus \{i\}$, equals $N_8|_{[7]\setminus\{i\}}$, and we check the conditions for $1 \le i < j \le 7$ while the N_i are chosen. This gives $|\mathcal{J}_8| = 45{,}319{,}744$. For nine points we form only the parts of the corresponding join of $\mathcal{P}_8$ with itself that are needed for Lemma 5.2 (Section 5.4).

5. The certificate

The repository given under Data availability holds the search programs, the certificates, the checkers with their validation of the certificate format, the second implementation and its differences from the checkers, and the Lean proofs with their formal statements, builds and further checks.

5.1. **The semidefinite program.** For fixed k, $\mathcal{P}$ and flag sets $\mathcal{F}_b$, the best bound that Theorem 3.3 gives is the value of the semidefinite program

$$\text{minimize } \lambda \quad \text{subject to} \quad \lambda \ge \operatorname{val}(M) \ \ (M \in \mathcal{P}), \qquad Q_b \succeq 0 \ \ (b = 1, \dots, r).$$

Since $c(F, G; M)$ vanishes for flags of different types, we lose nothing by giving each type its own block. For $k = 8$ the flags (N, J) have N realizable and (t, m) equal to $(1, 4)$, $(2, 4)$, $(4, 5)$, $(5, 6)$ or $(6, 7)$. These pairs satisfy $2m - t \le 8$. There is one t-pattern for $t \le 3$, and there are 3, 30 and 1,290 realizable patterns of 4, 5 and 6 points; each of these types has its own block.

We solved the program numerically with CSDP [3] and rounded the solution (Section 5.2). The solver need not be trusted, because only the rounded matrices enter the proof. The certificate keeps 596 of the 1,290 blocks of the pair $(6, 7)$ and has 631 blocks with 20,962 flags in all.

The certificates for $k = 6$ and $k = 7$ come from the same program over $\mathcal{P}_6$ and $\mathcal{P}_7$, with the pairs $(1, 3)$, $(2, 4)$, $(3, 4)$, $(4, 5)$ for $k = 6$ and $(1, 4)$, $(2, 4)$, $(4, 5)$, $(5, 6)$ for $k = 7$. For $m \le 3$ all m-patterns are empty. Hence $c(F, G; M)$ does not depend on M, and the pair $(1, 3)$ in the certificate for $k = 6$ only adds a nonnegative constant to every value.

For $k = 9$ we minimized the bound of Lemma 5.2 below in the same way. The blocks of the pairs $(1, 4)$, $(2, 4)$, $(4, 5)$, $(5, 6)$ and $(6, 7)$ define a function f_8 on 8-patterns, and the blocks of the pair $(7, 8)$ have flags on 458 realizable 8-patterns, chosen among those of largest value for $k = 8$. The last program gave $\lambda = 0.4006963373$. In the certificate we take $\lambda = 20034817/50000000$, this value rounded up to eight decimal places, and $\text{cap} = 20039817/50000000 = \lambda + 10^{-4}$. The certificate has 1,418 blocks with 47,632 flags in all.

5.2. **Rounding.** We round each numerical block $\tilde{Q}$ through an eigendecomposition $\tilde{Q} = \sum_i \alpha_i u_i u_i^{\mathsf{T}}$: we discard the smallest eigenvalues, round the scaled vectors Du_i to integer vectors v_i for a power of two D, and round the α_i to positive rationals w_i. The block of the certificate is the rational matrix

$$Q = \sum_i w_i \, (v_i/D)(v_i/D)^{\mathsf{T}},$$

which is positive semidefinite by construction. Rounding changes the values $\operatorname{val}(M)$ slightly. We do not estimate this change but evaluate $\operatorname{val}(M)$ exactly for the rounded matrices. For the certificate with $k = 8$, exactly two patterns of $\mathcal{J}_8$ have $\operatorname{val}(M) > 4006983/10000000$. Both are unrealizable, the certificate contains a Gordan certificate for each of them, and we apply Lemma 4.2 with E consisting of these two patterns. The claimed bound $4006983/10000000$ is the exact maximum $0.4006982904\ldots$ of val over $\mathcal{P}_8 = \mathcal{J}_8 \setminus E$, rounded up to seven decimal places. For the certificate with $k = 9$, exactly 29 patterns of $\mathcal{J}_8$ have $f_8(N) > \text{cap}$, and 19 of

the 9-patterns evaluated in Section 5.4 have $\mathrm{val}(M) > \lambda$. All of them are unrealizable, and the certificate contains a Gordan certificate for each of them.

5.3. **Values through restrictions.** The checkers for $k = 8$ and $k = 9$ evaluate val through patterns of fewer points, by the following identities.

Lemma 5.1. *Let $k \geq 5$ and let M be a k-pattern. Let U run over the $(k-1)$-element subsets of $[k]$, and read $M|_U$ as a $(k-1)$-pattern.*

(a) $d(M) = \frac{1}{k} \sum_U d(M|_U)$.
(b) *If $2m - t \leq k - 1$, then $c(F, G; M) = \frac{1}{k} \sum_U c(F, G; M|_U)$ for flags F, G of size m with t marked points.*
(c) *If $(t, m) = (k-2, k-1)$, then $c(F, G; M)$ is $\frac{1}{k(k-1)}$ times the number of ordered pairs (a, b) of distinct elements of $[k]$ with $F = \mathrm{fl}(M; [k] \setminus \{a, b\}, \{a\})$ and $G = \mathrm{fl}(M; [k] \setminus \{a, b\}, \{b\})$.*

Proof. A triple (J, A, B) covers $s = 2m - t$ points and lies in exactly $k - s$ of the sets U, and $\mathrm{fl}(M; J, A)$ is the same flag whether read in M or in $M|_U$, with the elements of U numbered increasingly. Since $\binom{k}{t}\binom{k-t}{m-t}\binom{k-m}{m-t} = \frac{k}{k-s}\binom{k-1}{t}\binom{k-1-t}{m-t}\binom{k-1-m}{m-t}$, this gives (b). Part (a) follows in the same way, since a tuple of T_k covers four points and $|T_k| = \frac{k}{k-4}|T_{k-1}|$. Part (c) is (3.1) with $J = [k] \setminus \{a, b\}$, $A = \{a\}$, $B = \{b\}$ and $D = k(k-1)$. □

Now let $k = 9$, and suppose that every block has $2m_b - t_b \leq 8$ or $(t_b, m_b) = (7, 8)$. For an 8-pattern N let

$$f_8(N) = d(N) + \sum_{b:\ 2m_b - t_b \leq 8} \ \sum_{F, G \in \mathcal{F}_b} (Q_b)_{FG}\, c(F, G; N), \tag{5.1}$$

that is, the value (3.3) for $k = 8$ of the blocks with $2m_b - t_b \leq 8$. By Lemma 5.1(a) and (b), and by Lemma 5.1(c) for the blocks of the pair $(7, 8)$ together with the symmetry of the matrices Q_b, every 9-pattern M satisfies

$$\mathrm{val}(M) = \frac{1}{9} \sum_{|U|=8} f_8(M|_U) + \frac{1}{36} \sum_{1 \leq i < j \leq 9} q(F_{ij}, F_{ji}), \qquad F_{ij} = \mathrm{fl}(M; [9] \setminus \{i, j\}, \{i\}), \tag{5.2}$$

where $q(F, G)$ is the sum of $(Q_b)_{FG}$ over the blocks b of the pair $(7, 8)$ that contain both F and G. The pattern of the flag F_{ij} is $M|_{[9]\setminus\{j\}}$. Let Z be the set of 8-patterns N such that $(N, J) \in \mathcal{F}_b$ for some block b of the pair $(7, 8)$ and some J. Then $q(F_{ij}, F_{ji}) = 0$ unless $M|_{[9]\setminus\{i\}}$ and $M|_{[9]\setminus\{j\}}$ both lie in Z.

Lemma 5.2. *Let $k = 9$, let $\mathcal{F}_b$ and Q_b be as in Theorem 3.3, with $2m_b - t_b \leq 8$ or $(t_b, m_b) = (7, 8)$ for every block, and let λ and* cap *be real numbers. Put $\theta = 9\lambda - 8\mathrm{cap}$ and assume the following.*

(1) $f_8(N) \leq \mathrm{cap}$ *for every $N \in \mathcal{R}_8$.*
(2) $\mathrm{val}(M) \leq \lambda$ *for every $M \in \mathcal{R}_9$ with $f_8(M|_U) > \theta$ for all 8-element sets $U \subseteq [9]$.*
(3) $\mathrm{val}(M) \leq \lambda$ *for every $M \in \mathcal{R}_9$ with $M|_U \in Z$ for at least two 8-element sets $U \subseteq [9]$.*

Then $C \leq \lambda$.

Proof. Let $M \in \mathcal{R}_9$ be a pattern to which neither (2) nor (3) applies. Then $M|_U \in Z$ for at most one U, so every term of the second sum in (5.2) vanishes. The restrictions $M|_U$ are realizable, so $f_8(M|_U) \leq \mathrm{cap}$ for all U by (1), and $f_8(M|_{U_0}) \leq \theta$ for some U_0 because (2) does not apply. Hence $\mathrm{val}(M) \leq (\theta + 8\mathrm{cap})/9 = \lambda$. Together with (2) and (3), this gives $\mathrm{val}(M) \leq \lambda$ for all $M \in \mathcal{R}_9$, and Theorem 3.3 with $\mathcal{P} = \mathcal{R}_9$ gives $C \leq \lambda$. □

5.4. **The exact check.** For the certificate of Theorem 1.1, our checker verifies the following five facts in exact rational arithmetic. It shares with the search programs only the conventions for indexing T_k and the file formats.

(i) The trees are extension trees in the sense of Section 4: they are given for all patterns of the previous level, each path decides every tuple of $T_k \setminus T_{k-1}$ exactly once, and every absent branch carries a Gordan certificate for $S_P(s,e)$, which is checked against the rows of $S_P(s,e)$ built from their definitions. The sets $\mathcal{P}_4,\dots,\mathcal{P}_7$ are rebuilt from the trees.

(ii) Each block Q_b is the exact rational matrix formed from its terms (w_i,v_i) with $w_i \ge 0$, and every block has $2m_b - t_b \le 7$ or (t_b,m_b) equal to $(6,7)$ or $(7,8)$.

(iii) For each $N \in \mathcal{P}_7$, the value $f_7(N)$, the sum of $d(N)$ and of the contributions of the blocks with $2m_b - t_b \le 7$, is computed from (3.1) with $k=7$ by enumerating all triples (J,A,B). The set $\mathcal{J}_8$ is formed by the join of Section 4, and for each $M \in \mathcal{J}_8$

$$f_8(M) = \frac{1}{8}\sum_{|U|=7} f_7(M|_U) + \frac{1}{28}\sum_{1\le i<j\le 8} q_7(F_{ij},F_{ji}), \qquad F_{ij} = \mathrm{fl}(M;[8]\setminus\{i,j\},\{i\}),$$

where $q_7(F,G)$ is the sum of $(Q_b)_{FG}$ over the blocks b of the pair $(6,7)$ that contain both F and G. This is Lemma 5.1 with $k=8$, together with the symmetry of the matrices Q_b. Every $M \in \mathcal{J}_8$ with $f_8(M) > \mathrm{cap}$ has a Gordan certificate for $S(M)$; let E_8 be the set of these patterns.

(iv) The checker determines which patterns of Z lie in $\mathcal{J}_8$, and the set $\mathcal{S}$ of the patterns $M \in \mathcal{J}_8 \setminus E_8$ with $f_8(M) > \theta$. By joins as in Section 4 it forms

(a) all 9-patterns M with $M|_U \in \mathcal{S}$ for every 8-element set $U \subseteq [9]$, and

(b) for each $N \in Z \cap \mathcal{J}_8$ and each $x \in [9]$, all 9-patterns M with $M|_{[9]\setminus\{x\}} = N$ and $M|_U \in \mathcal{J}_8 \setminus E_8$ for every U, such that $M|_{[9]\setminus\{y\}} \in Z$ for some $y > x$ and for no $y < x$.

For (b), Lemma 4.3(a) gives the restriction of $M|_{[9]\setminus\{y\}}$ to the points $[9]\setminus\{x,y\}$ for each $y \ne x$, which restricts the candidates for $M|_{[9]\setminus\{y\}}$ to the patterns of $\mathcal{J}_8$ with one given restriction to 7 points. The values $\mathrm{val}(M)$ of the patterns of (a) and (b) are computed by (5.2), with f_8 from step (iii).

(v) Every M of step (iv) with $\mathrm{val}(M) > \lambda$ has a Gordan certificate for $S(M)$, and $\mathrm{val}(M) \le \lambda$ for all other M of step (iv).

Proof of Theorem 1.1. The checker accepts the certificate, which states $\lambda = 20034817/50000000$ and $\mathrm{cap} = 20039817/50000000$, so that $\theta = 19994817/50000000$. By step (i) and Proposition 4.1, $\mathcal{R}_7 \subseteq \mathcal{P}_7$. By step (iii) and Lemma 4.2, $\mathcal{R}_8 \subseteq \mathcal{P}_8 = \mathcal{J}_8 \setminus E_8$ and $f_8 \le \mathrm{cap}$ on $\mathcal{P}_8$, which gives (1). By step (ii) the matrices Q_b are positive semidefinite and the blocks are as required in Theorem 3.3 and Lemma 5.2. Let $M \in \mathcal{R}_9$. Its restrictions $M|_U$ are realizable and hence lie in $\mathcal{P}_8$. If $f_8(M|_U) > \theta$ for all U, then all $M|_U$ lie in $\mathcal{S}$, and M is one of the patterns (a) of step (iv) by Lemma 4.3. If $M|_U \in Z$ for at least two sets U, let x be the least y with $M|_{[9]\setminus\{y\}} \in Z$. Then M is one of the patterns (b) for $N = M|_{[9]\setminus\{x\}}$. In both cases $S(M)$ has no Gordan certificate, since M is realizable, so step (v) gives $\mathrm{val}(M) \le \lambda$. This proves (2) and (3), and Lemma 5.2 gives $C \le 20034817/50000000$. The lower bound is (1.1). □

For the certificate with $k=8$, the checker performs steps (i) to (iii), where step (ii) allows only blocks with $2m_b - t_b \le 7$ or $(t_b,m_b) = (6,7)$ and step (iii) compares f_8 with the claimed bound $4006983/10000000$ instead of with cap; for this certificate $f_8 = \mathrm{val}$. By Lemma 4.2 and Theorem 3.3 with $\mathcal{P} = \mathcal{P}_8$, this proves $C \le 4006983/10000000$. For $k=6$ and $k=7$, a simpler checker performs step (i) for the levels 4 to k and the checks of step (ii) with the condition $2m_b - t_b \le k$, and it evaluates $\mathrm{val}(M)$ from the definitions for every $M \in \mathcal{P}_k$.

The independent second implementation covers all four certificates. For $k = 6$ and $k = 7$ it gives every pattern the value $\mathrm{val}(M)$ that the checker computes, and for $k = 8$ and $k = 9$ it reproduces the numbers of patterns, the maximum of val and the patterns above the bounds, whose Gordan certificates it verifies with rows that it builds itself.

TABLE 1. The bounds of Theorem 1.1 ($k = 9$) and of the certificates for $k = 6$, 7 and 8. For $k \le 8$ the largest value is the maximum of val over $\mathcal{P}_k$, where $\mathcal{P}_8$ is $\mathcal{J}_8$ without the two patterns of E; for $k = 9$ it is the maximum of val over the patterns of step (iv) without Gordan certificates, and the largest density is not computed. The excess over the conjectured value $C^* = 0.4006953723\ldots$ is rounded.

Sample points k	6	7	8	9
Bound λ	201273/500000	10019/25000	4006983/10000000	20034817/50000000
Largest value	$0.4025451283\ldots$	$0.4007589336\ldots$	$0.4006982904\ldots$	$0.4006963371\ldots$
$\lvert T_k\rvert$	60	140	280	504
Largest density	7/15	16/35	25/56	–
Excess of λ over C^*	$1.85 \cdot 10^{-3}$	$6.5 \cdot 10^{-5}$	$2.9 \cdot 10^{-6}$	$9.7 \cdot 10^{-7}$

6. Formal verification

The upper bounds of Table 1, and so the upper bound of Theorem 1.1, also have proofs in the Lean 4 proof assistant [9] with the Mathlib library [8]. The formal statements concern probability measures μ on $\mathbb{R}$ with $\mu((-\infty, 0)) = 0$. They define $P(\mu)$ as the measure of the set $\{x \in \mathbb{R}^4 : \ x_1 + x_2 + x_3 < 2x_4\}$ under the product of four copies of μ, and C as the supremum of $P(\mu)$ over these measures. For each certificate they state that $P(\mu) \le \lambda$ for every such μ, the same bound for independent, identically distributed random variables that are nonnegative almost surely, and $C \le \lambda$; a further theorem computes $P(\mu) = 1/4$ for $\mu = (\delta_0 + \delta_1)/2$, so the definitions are not degenerate. The formal proofs follow Sections 2 to 5. For each certificate, a checking program written in Lean performs checks corresponding to those of Section 5.4, and a Lean proof shows that whenever the program accepts its input, there are positive semidefinite blocks with $\mathrm{val}(M) \le \lambda$ for every realizable k-pattern M, as Theorem 3.3 requires. Scripts convert the certificate files into Lean source files; they need not be trusted, since the correctness proofs hold for every input that the programs accept. The lower bound in (1.1) is not formalized.

In one proof for each certificate, the checks are evaluated by `native_decide`, that is, by code that Lean's compiler produces from the programs. Its trusted base includes the Lean compiler and runtime and, for $k = 7$, 8 and 9, a C compiler, and each of these theorems depends, besides the three standard axioms of Lean, on one auxiliary axiom that asserts that the check of its certificate returns true. For $k = 6$, 7 and 9 there are also proofs in which Lean's kernel evaluates the checks itself; the resulting theorems depend only on the three standard axioms. For $k = 9$ this gives a proof of the upper bound of Theorem 1.1 that does not depend on Lean's compiler or on the code it generates: its trusted base is Lean's kernel, including the kernel's built-in arithmetic on natural numbers. The bound for $k = 8$ has only the proof with compiled checks.

7. Concluding remarks

The dual solution. The dual of the semidefinite program of Section 5 asks for a probability vector y on $\mathcal{P}$ that maximizes $\sum_M y(M)\, d(M)$ subject to $\sum_M y(M)\, c_b(M) \succeq 0$ for each b, where $c_b(M) = (c(F, G; M))_{F,G \in \mathcal{F}_b}$. By Lemmas 2.3 and 3.1, the vector p_μ is feasible for every measure μ without atoms that satisfies $\mu((0, \infty)) = 1$, with objective value $P(\mu)$. So the

dual is a relaxation of the problem of maximizing $P(\mu)$. Numerically, the dual solution of the last program for $k = 7$, on 4,146 patterns and with the trace bound ignored (its multiplier is $5 \cdot 10^{-11}$), gives weight above 10^{-6} to 27 patterns. More than 99.97% of its mass lies on the 162 patterns observed in 300,000 samples of 7 points from the construction of Bellec and Fritz [2, Propositions 3.3 and 4.1]. In this construction we take $p = 0.474346$ and $N = 50$ in [2, (4.4)], truncate it after nine scales, put a uniform variable at the finest scale, and choose the ratio between consecutive scales small enough not to affect the patterns. Conversely, more than 92% of the samples fall on the 27 patterns with weight above 10^{-6}. The relaxation and the lower bound construction thus favor the same patterns, which is consistent with Conjecture 5.1 of [2].

An open question. From $k = 6$ to $k = 9$ the excess of the bound over C^* falls by factors of about 29, 22 and 3 (Table 1). We do not know whether a certificate of this kind with some finite k reaches C^*. The construction of Bellec and Fritz suggests why this may be hard. In it, each sample point X_i is a series $\sum_{j\ge1} \eta^{j-1}\xi_j^{(i)}$ with a small ratio $\eta \in (0,1)$ and independent digits $\xi_j^{(i)}$ with a common law. When the first digits satisfy $\xi_1^{(1)} + \xi_1^{(2)} + \xi_1^{(3)} = 2\xi_1^{(4)}$, the event $X_1 + X_2 + X_3 < 2X_4$ is decided by the numbers $\sum_{j\ge2} \eta^{j-2}\xi_j^{(i)}$ formed by the remaining digits, which are independent of the first digits and have the same law as the points themselves [2, proof of Proposition 3.3]. The problem thus reappears at every scale, while a certificate for k points constrains only the patterns of k points.

Remark 7.1. The conjectured value C^* is the algebraic number $\beta(p^*)$, where $\beta(p) = p(2-p)/(1+p+p^2+p^3)$. The numerator of β' is $p^4 - 4p^3 - 3p^2 - 2p + 2$, which is irreducible over $\mathbb{Q}$, and $16z^4 + 8z^3 + 99z^2 - 32z - 4$ is 16 times the characteristic polynomial of multiplication by $\beta(p)$ on $\mathbb{Q}[p]/(p^4 - 4p^3 - 3p^2 - 2p + 2)$. This polynomial is irreducible over $\mathbb{Q}$, so up to the factor 16 it is the minimal polynomial of C^*, and C^* is its only positive root.

Data availability

The certificates, the checkers, the second implementation, the search programs and the Lean proofs of Section 6 are archived at doi:10.5281/zenodo.23716000. The programs, the certificates for $k = 6$ and $k = 7$ and the Lean sources, except for some large data files, are also available at `https://github.com/tejasvi/beat-the-average-bound/tree/v6`. The Lean proofs build with Lean 4 version 4.35.0-rc3 and Mathlib at the tag `v4.35.0-rc3` (commit `c55e6e786f49`).

Use of AI tools

Generative AI tools were used to find and prove the results, to write the programs and the Lean formalization, and to draft the text. The author reviewed the arguments, the programs and the formal proofs, and takes responsibility for the content.

References

[1] J. Balogh, P. Hu, B. Lidický, O. Pikhurko, B. Udvari and J. Volec, *Minimum number of monotone subsequences of length 4 in permutations*, Combin. Probab. Comput. **24** (2015), 658–679.

[2] P. C. Bellec and T. Fritz, *On maximizing the probability of a linear inequality between iid random variables*, Amer. Math. Monthly (2026), doi:10.1080/00029890.2026.2702871; arXiv:2412.15179v5 (listed on arXiv as *Optimizing over iid distributions and the Beat the Average game*), whose numbering we cite.

[3] B. Borchers, *CSDP, a C library for semidefinite programming*, Optim. Methods Softw. **11** (1999), 613–623.

[4] T. Fritz, *How large can* $\mathbf{P}[X_1 + X_2 + X_3 < 2X_4]$ *get?*, MathOverflow question 474916, 2024, `https://mathoverflow.net/q/474916`.

[5] B. Georgiev, J. Gómez-Serrano, T. Tao and A. Z. Wagner, *Mathematical exploration and discovery at scale*, arXiv:2511.02864, 2025.

[6] P. Gordan, *Ueber die Auflösung linearer Gleichungen mit reellen Coefficienten*, Math. Ann. **6** (1873), 23–28.

[7] Q. Huangfu and J. A. J. Hall, *Parallelizing the dual revised simplex method*, Math. Program. Comput. **10** (2018), 119–142.
[8] The mathlib Community, *The Lean mathematical library*, in: Proc. 9th ACM SIGPLAN Int. Conf. Certified Programs and Proofs (CPP 2020), 2020, 367–381.
[9] L. de Moura and S. Ullrich, *The Lean 4 theorem prover and programming language*, in: Automated Deduction – CADE 28, Lecture Notes in Comput. Sci. **12699**, Springer, 2021, 625–635.
[10] A. Novikov et al., *AlphaEvolve: A coding agent for scientific and algorithmic discovery*, arXiv:2506.13131, 2025.
[11] A. A. Razborov, *Flag algebras*, J. Symbolic Logic **72** (2007), 1239–1282.
[12] J. Sliačan and W. Stromquist, *Improving bounds on packing densities of 4-point permutations*, Discrete Math. Theor. Comput. Sci. **19** (2018), no. 2, paper 3.
[13] V. Yu, answer to [4], MathOverflow, 2024, `https://mathoverflow.net/a/484853`.

ORCID: `https://orcid.org/0000-0003-4668-3639`