hexagon:2611.00469v1

CERTIFIED UPPER BOUNDS FOR TWO SUM-DIFFERENCE EXPONENTS

TEJASVI SINGH TOMAR

Abstract. For finite sets A, B of real numbers and $G \subseteq A \times B$, Katz and Tao proved in 1999 that $\{a-b : (a,b) \in G\}$ has at most $N^{11/6}$ elements whenever A, B and $\{a+b : (a,b) \in G\}$ have at most N elements; by an argument of Bourgain, bounds of this kind give lower bounds for the dimension of Kakeya sets. We lower the exponent $11/6$ to 1.8273, and to 1.8295 in every abelian group. We also prove $|X - Y| \leq |X + Y|^{1.4209}$ for all finite sets X, Y of integers, improving the exponent 1.4543 of Kleinwaks; through his transfer theorem, this lowers the upper bound for the sum-difference constant of Gyarmati, Hennecart and Ruzsa from 1.3124 to 1.2962. Both proofs add up entropy inequalities among integer combinations of coordinates of several copies of a random pair, with rational weights found by linear programming; all four bounds are also proved in Lean 4.

1. Introduction

For finite subsets A, B of an abelian group and $G \subseteq A \times B$ we write

$$A \overset{G}{+} B = \{a + b : (a,b) \in G\}, \qquad A \overset{G}{-} B = \{a - b : (a,b) \in G\}.$$

Bourgain [2] showed that bounds $|A \overset{G}{-} B| \leq N^{2-\varepsilon}$, valid whenever $|A|, |B|, |A \overset{G}{+} B| \leq N$, give lower bounds for the dimension of Kakeya sets, and he proved such a bound with $\varepsilon = 1/13$. Katz and Tao [13] proved it with $\varepsilon = 1/6$ in any abelian group. The Kakeya bounds obtained in this way are no longer the best known in any dimension [28, page 3b]. Let C_{3b} be the least exponent C such that

$$|A \overset{G}{-} B| \leq \max(|A|, |B|, |A \overset{G}{+} B|)^C \tag{1.1}$$

for all finite $A, B \subset \mathbb{R}$ and $G \subseteq A \times B$. By [9] (see also [27, §1]), C_{3b} is also the least C such that

$$H(X - Y) \leq C \max(H(X), H(Y), H(X + Y)) \tag{1.2}$$

for all real random variables X, Y with finitely many values, not necessarily independent, where H denotes Shannon entropy. We prove our bounds in both forms (Lemma 3.1). The relation of C_{3b} to the arithmetic Kakeya conjecture is discussed in [14, 9, 27]. The notation C_{3b} is that of Tao's list of optimization constants [28], where it is constant 3b, and C_{3b} is also Problem 6.30 of [8]. The upper bound $11/6 = 1.8333\ldots$ of [13] is the best one recorded in these two sources, and the lower bounds recorded there are slightly above 1.7789 [16, 21]. Farnamfar [5, 6] proves that several families of combinatorial constructions that certify upper bounds for C_{3b} cannot improve $11/6$, and that an improving construction must have dimension at least three. The proof of Theorem 1.1 uses entropy inequalities instead of such constructions.

Theorem 1.1. *There are rational numbers $\lambda_X \approx 0.706543$, $\lambda_Y \approx 0.648446$ and $\lambda_S \approx 0.472307$, given exactly in Appendix B, with sum less than 1.8272963, such that the following hold.*

2020 *Mathematics Subject Classification.* 11B13, 11B30; 94A17, 90C05, 68V05.

Key words and phrases. Sumsets, difference sets, sum-difference exponents, arithmetic Kakeya conjecture, Shannon entropy, copy lemma, non-Shannon inequalities, linear programming, computer-assisted proof.

(a) *If X, Y are random variables with finitely many values in a torsion-free abelian group, then $H(X-Y) \le \lambda_X H(X) + \lambda_Y H(Y) + \lambda_S H(X+Y)$.*

(b) *If A, B are finite subsets of a torsion-free abelian group and $G \subseteq A \times B$, then $|A \overset{G}{-} B| \le |A|^{\lambda_X} |B|^{\lambda_Y} |A \overset{G}{+} B|^{\lambda_S}$.*

In particular $C_{3b} \le 1.8272963$ in both formulations (1.1) *and* (1.2).

Theorem 1.2. *Statements* (a) *and* (b) *of Theorem 1.1 hold in all abelian groups with*

$$(\lambda_X, \lambda_Y, \lambda_S) = (\frac{882}{1\,259}, \frac{1\,609}{2\,518}, \frac{2\,467}{5\,036}),$$

whose sum is $9\,213/5\,036 < 1.8294282$.

For finite subsets X, Y of an abelian group, Ruzsa's triangle inequality and Plünnecke's inequality give $|X-Y| \le |X+Y|^{3/2}$ [10, §1]. Kleinwaks [15, Theorem 2.1] proved the exponent $\lambda_\infty = (9451e - 3286)/(5378e + 787) = 1.454277\ldots$ for finite subsets of any abelian group. Taking $X = Y$ shows that such an exponent is at least the single-set exponent of [8, Problem 6.43], the least C with $|A - A| \le |A + A|^C$ for all finite $A \subset \mathbb{Z}$, which lies between $\log(1+\sqrt{2})/\log 2 = 1.2715\ldots$ [11] and $4/3$ [7].

Kleinwaks also proved that an exponent λ valid for all finite $X, Y \subset \mathbb{Z}$ gives the bound $2 - 1/\lambda$ for the sum-difference constant C_{3a} of Gyarmati, Hennecart and Ruzsa [15, Theorem 2.2]. This constant is the supremum of the exponents t for which there are finite sets $A, B \subset \mathbb{Z}$ with $B \neq \emptyset$, $|A|$ arbitrarily large, $|A+B| \le K|A|$ and $|A-B| \ge c|A+B|^t$ for some constants K and $c > 0$ [28, page 3a]. It is Problem 6.44 of [8]. The sources state its quantifiers differently, and Section 4.4 compares the readings. Gyarmati, Hennecart and Ruzsa proved $C_{3a} \le 4/3$ [10, Corollary 3], and Kleinwaks proved $C_{3a} \le 2 - 1/\lambda_\infty = 1.312373\ldots$ [15, Corollary 2.3]. The lower bounds recorded in [28, page 3a] reach 1.19102809 with an explicit construction and 1.19519192 as a limit of constructions.

Theorem 1.3. *For all finite subsets X, Y of a torsion-free abelian group,*

$$|X - Y| \le |X + Y|^\lambda, \qquad \lambda = \frac{20\,486\,493\,787}{14\,418\,560\,416} < 1.42084184.$$

Corollary 1.4.

$$C_{3a} \le 2 - \frac{1}{\lambda} = \frac{26\,554\,427\,158}{20\,486\,493\,787} < 1.29619190.$$

Remark 1.5. A set A of n integers determines at most $n^{1.8272963}$ distinct positive common differences of three-term arithmetic progressions. This improves the exponent $11/6$ of Katz and Tao [13] for Erdős problem #1097, which asks how many such differences there can be. Chan observed that this problem is equivalent to the sums-differences question of Bourgain recalled above [1]. For the proof, choose for each such difference $d > 0$ an a_d with $a_d, a_d + d, a_d + 2d \in A$, and let G consist of the pairs $(a_d + 2d, a_d)$. Their differences $2d$ are distinct, so $|A \overset{G}{-} A|$ is the number of such d, and their sums $2(a_d + d)$ lie in $\{2a : a \in A\}$, so $|A \overset{G}{+} A| \le n$. Part (b) of Theorem 1.1 with $B = A$ gives $|A \overset{G}{-} A| \le n^{\lambda_X + \lambda_Y + \lambda_S} \le n^{1.8272963}$.

The method. Both proofs are entropy arguments in the style of [15]; entropy methods in additive combinatorics are developed in [25, 26]. For Theorems 1.1 and 1.2 we take a random pair (X, Y) such that $X - Y$ is injective on its support. Three applications of the copy lemma put six copies of the pair on one probability space. This configuration is a variant, in entropy form, of the one behind the bound $11/6$ of [13]: in it

$$6H(X,Y) = 4H(X) + 4H(Y) + 3H(X+Y) - \Delta_{\mathrm{KT}},$$

where $\Delta_{\mathrm{KT}} \geq 0$ is a sum of five mutual informations between coordinates and sums of the copies (Proposition 3.2). An improvement of $11/6$ therefore amounts to a lower bound for Δ_{KT} in terms of $H(X)$, $H(Y)$ and $H(X+Y)$. Integer combinations of the coordinates of the copies are random variables, and Shannon inequalities hold for their joint entropies. The forms x, y, $x+y$ and $x-y$ of a single copy (x, y) do not suffice (Remark 3.5). We call the other combinations *cross forms*. They are new random variables built from several copies at once, and Shannon inequalities that involve them constrain the joint law of the copies in ways that the standard forms do not detect. For example, if two copies (x_0, y_0) and (x_2, y_0) share their second coordinate, then $x_0 + y_0 + x_2$ is a cross form. Shannon inequalities among five cross forms already bound Δ_{KT} from below and give the exponent $841/459 < 11/6$ in every abelian group (Appendix A).

For a fixed finite set of cross forms, a linear program finds a nonnegative combination that bounds $H(X, Y)$ by $\lambda_X H(X) + \lambda_Y H(Y) + \lambda_S H(X + Y)$; it combines Shannon inequalities, the identities given by the copy lemma and non-Shannon inequalities, each of the latter proved from the copy lemma. We call such a combination a *certificate*. Linear programs over linear forms of glued copies of a random pair, with copy-lemma variables, were used before in [20] for the Katz–Tao problem with many slopes. With 41 cross forms this gives Theorem 1.1, and without non-Shannon inequalities 20 cross forms give Theorem 1.2.

For Theorem 1.3 we start from Kleinwaks' coupling, a random pair (X_1, Y_1) uniform on a set $\Gamma \subseteq X \times Y$ with one pair for each element of $X - Y$. Kleinwaks uses two independent copies of this pair and the entropies $h(i, j)$ of sums of independent copies of its marginals. We add three further variables with the law of X_1, independent of the two copies and of each other. Besides symmetries of the whole configuration, we use partial symmetries and partial reflections, which act only on the independent blocks that occur in a set of forms (Section 4.2). Shannon inequalities among integer combinations of these coordinates give an inequality between $\log|X - Y|$, $\log|X + Y|$ and 12 values of h (Proposition 4.1). Finitely many of Kleinwaks' inequalities for h eliminate these values (Lemma 4.2). With the two copies alone, only sums of at most two coordinates are identified with values of h; the added variables bring in sums of up to five independent variables (Remark 4.3). Unlike the two-copy inequality of [15, Proposition 5.1], which uses ten instances of the non-Shannon families of Matúš [19] and of Csirmaz and Csirmaz [3], our certificate uses only Shannon inequalities.

We do not know how far larger configurations, more cross forms or further non-Shannon inequalities can lower the two exponents.

Verification. The weights in the certificates are exact rational numbers, obtained from floating-point solutions of the linear programs. Section 5 describes how each certificate is verified in exact rational arithmetic by two independently written checkers and how the results above are proved in Lean 4.

Section 2 collects the entropy tools. Section 3 proves Theorems 1.1 and 1.2, and Section 4 proves Theorem 1.3 and Corollary 1.4, including a proof of Kleinwaks' transfer theorem. Appendix A works through the short certificate.

2. Entropy tools

All random variables below take finitely many values and are defined on a common probability space, which we enlarge when we construct new variables. We write H for Shannon entropy with the natural logarithm, treat a tuple of random variables as one random variable, and put

$$I(U; V \mid W) = H(U, W) + H(V, W) - H(U, V, W) - H(W), \qquad I(U; V) = I(U; V \mid \emptyset).$$

We use Shannon's inequality $I(U;V \mid W) \geq 0$, the inequality $H(f(U)) \leq H(U)$ for any function f, with equality if f is injective on the range of U, and the identity $H(U,V) = H(U) + H(V)$ for independent U and V.

Lemma 2.1 (copy lemma [30, 4])**.** *Let A, B, C be random variables. On an enlarged probability space there is a random variable A' such that (A', B) has the law of (A, B) and A' is conditionally independent of (A, C) given B. In particular*

$$H(A', A, B, C) + H(B) = H(A, B, C) + H(A', B).$$

Proof. Put $P(A = a, B = b, C = c, A' = a') = P(A = a, B = b, C = c)\, P(A = a' \mid B = b)$ for all b with $P(B = b) > 0$. This law extends the law of (A, B, C), the pair (A', B) has the law of (A, B), and the conditional law of A' given (A, B, C) depends on B only. □

Every relation that holds almost surely between A and B therefore holds between A' and B. For random variables a, b, c, d let

$$[abcd] = I(a; b \mid c) + I(a; b \mid d) + I(c; d) - I(a; b)$$

be the Ingleton expression.

Lemma 2.2. *Let $a_1, \ldots, a_r$ be random variables and let $\psi = \sum_S c_S H(a_S)$, where S runs over the nonempty subsets of $\{1, \ldots, r\}$ and $a_S = (a_i)_{i \in S}$. Suppose that random variables $b_1, \ldots, b_q$ are obtained from $a_1, \ldots, a_r$ by successive applications of Lemma 2.1, and that ψ is identically, as a linear combination of entropies, the sum of nonnegative multiples of Shannon inequalities and of inequalities that hold for all random variables, applied to joint variables made of the a_i and b_j, and of a linear combination of the identities given by these applications: the identity of Lemma 2.1 and the equalities of entropies that follow from the equality of the laws of (A', B) and (A, B). Then $\psi \geq 0$ for all random variables $a_1, \ldots, a_r$.*

Proof. The applications of Lemma 2.1 do not change the joint law of $a_1, \ldots, a_r$. At the constructed variables the identities vanish and the remaining terms are nonnegative. □

Dougherty, Freiling and Zeger prove the inequalities of their list [4, Section VIII] in this way. Matúš's inequality

$$M_s(a, b, c, d) = I(b; c \mid a) + s\,[abcd] + \tbinom{s+1}{2}\big(I(a; c \mid b) + I(a; b \mid c)\big) \geq 0$$

[19, Corollary 3, formula (4)] follows for each $s \geq 1$ from the case $s - 1$ by Lemma 2.2, with one copy of a over (b, c) [4, (80) and Rule [1]].

2.1. **Linear forms and closures.** Let $\mathbf{G}$ be an abelian group, let $Z = (Z_1, \ldots, Z_m)$ be a random vector in $\mathbf{G}^m$, and let $\mathcal{R} \subseteq \mathbb{Z}^m$ be a set of integer vectors r with $\sum_i r_i Z_i = 0$ almost surely. A *form* is a vector $f \in \mathbb{Z}^m$, which defines the random variable $f(Z) = \sum_i f_i Z_i$. For a finite set F of forms, $H[F]$ denotes the entropy of the tuple $(f(Z))_{f \in F}$, and $\langle F \rangle$ denotes the subgroup of $\mathbb{Z}^m$ generated by F and $\mathcal{R}$.

Lemma 2.3. *Let F be a finite set of forms and let f be a form.*

(a) *If $f \in \langle F \rangle$, then $f(Z)$ is a function of $(g(Z))_{g \in F}$.*
(b) *If $cf \in \langle F \rangle$ for an integer $c \neq 0$ such that $u \mapsto cu$ is injective on $\mathbf{G}$, then $f(Z)$ is a function of $(g(Z))_{g \in F}$.*
(c) *Let x, y be forms such that $(x(Z), y(Z))$ lies almost surely in a set on which $(u, v) \mapsto u - v$ is injective. If $(x - y)(Z)$ is a function of $(g(Z))_{g \in F}$, then so are $x(Z)$ and $y(Z)$.*

Proof. For (a) write $f = \sum_{g \in F} a_g g + r$ with r in the subgroup generated by $\mathcal{R}$, so that $f(Z) = \sum_g a_g g(Z)$ almost surely. For (b), $cf(Z)$ is a function of $(g(Z))_{g \in F}$ by (a), and $u \mapsto cu$ is injective. Part (c) is immediate. □

Fix a finite ground set E of forms and a list of pairs (x, y) as in (c). The *closure* $\operatorname{cl} F$ of $F \subseteq E$ is the smallest subset of E that contains F and is closed under the rules (a)–(c), where (b) is used exactly for the integers $c \neq 0$ for which $u \mapsto cu$ is injective on $\mathbf{G}$, for example for all $c \neq 0$ if $\mathbf{G}$ is torsion-free. By Lemma 2.3, $H[\operatorname{cl} F] = H[F]$.

3. The Kakeya-type exponent

3.1. **Reduction to injective pairs.** A pair (X, Y) of random variables in an abelian group is *injective* if $(x, y) \mapsto x - y$ is injective on the support of (X, Y). Then $H(X, Y) = H(X - Y)$.

Lemma 3.1. *Let* $\mathbf{G}$ *be an abelian group and* $\lambda_X, \lambda_Y, \lambda_S \geq 0$. *Suppose that*

$$H(X, Y) \leq \lambda_X H(X) + \lambda_Y H(Y) + \lambda_S H(X + Y) \tag{3.1}$$

for every injective pair (X, Y) *in* $\mathbf{G}$. *Then*

(a) $H(X - Y) \leq \lambda_X H(X) + \lambda_Y H(Y) + \lambda_S H(X + Y)$ *for all pairs* (X, Y) *in* $\mathbf{G}$, *and*

(b) $|A \overset{G}{-} B| \leq |A|^{\lambda_X} |B|^{\lambda_Y} |A \overset{G}{+} B|^{\lambda_S}$ *for all finite* $A, B \subseteq \mathbf{G}$ *and* $G \subseteq A \times B$.

Proof. (b) We may assume $G \neq \emptyset$. Choose $G' \subseteq G$ with exactly one pair (a, b) satisfying $a - b = z$ for each $z \in A \overset{G}{-} B$, and let (X, Y) be uniform on G'. This pair is injective, $H(X, Y) = \log |G'| = \log |A \overset{G}{-} B|$, and X, Y, $X + Y$ take values in A, B, $A \overset{G}{+} B$. Since a random variable with values in a finite set T has entropy at most $\log |T|$, (b) follows from (3.1).

(a) We use the construction in the proof of [27, Proposition 3.6]. Let $D = X - Y$. For each z in the range of D choose a random pair $\varphi(z)$ with the conditional law of (X, Y) given $D = z$, independently for different z and independently of (X, Y), and put $(X', Y') = \varphi(D)$. Given $D = z$, both (X, Y) and (X', Y') have the conditional law of (X, Y) given $D = z$, so (X', Y') has the law of (X, Y), and $X' - Y' = D$. For a value φ_0 of φ, the conditional law of (X', Y') given $\varphi = \varphi_0$ is the law of $\varphi_0(D)$, which is supported on points with distinct differences, so (3.1) applies to it. Averaging over φ_0 gives

$$\begin{aligned} H(X', Y' \mid \varphi) &\leq \lambda_X H(X' \mid \varphi) + \lambda_Y H(Y' \mid \varphi) + \lambda_S H(X' + Y' \mid \varphi) \\ &\leq \lambda_X H(X) + \lambda_Y H(Y) + \lambda_S H(X + Y), \end{aligned}$$

because conditioning does not increase entropy. Given φ, the pair $\varphi(D)$ and D determine each other, so $H(X', Y' \mid \varphi) = H(D \mid \varphi) = H(D)$, since φ is independent of D. □

With $C = \lambda_X + \lambda_Y + \lambda_S$, Lemma 3.1 gives (1.1) and (1.2) with this C for subsets and random variables of $\mathbf{G}$.

3.2. **The configuration.** Let (X, Y) be an injective pair in an abelian group $\mathbf{G}$. We construct six pairs $g_v = (x_v, y_v)$, $0 \leq v \leq 5$, each with the law of (X, Y), by three applications of Lemma 2.1, starting from $g_0 = (X, Y)$:

(S1) g_1 is a copy of g_0 over x_0;
(S2) (g_2, g_3) is a copy of (g_0, g_1) over (y_0, y_1);
(S3) (g_4, g_5) is a copy of (g_0, g_1) over $(x_0 + y_0, y_1)$.

In each step, Lemma 2.1 is applied with A the tuple U of pairs to be copied, B the tuple V of variables it is copied over, and C the pairs constructed before. The new pairs satisfy the relations that U satisfies with V:

$$x_1 = x_0; \qquad y_2 = y_0,\ y_3 = y_1,\ x_3 = x_2; \qquad x_4 + y_4 = x_0 + y_0,\ y_5 = y_1,\ x_5 = x_4$$

almost surely. Hence all twelve coordinates are integer combinations of the five *free coordinates* x_0, y_0, y_1, x_2, x_4:

$$\begin{aligned} &g_0 = (x_0, y_0), \qquad g_1 = (x_0, y_1), \qquad g_2 = (x_2, y_0), \\ &g_3 = (x_2, y_1), \qquad g_4 = (x_4, x_0 + y_0 - x_4), \qquad g_5 = (x_4, y_1). \end{aligned} \tag{3.2}$$

We apply Section 2.1 to the vector $Z \in \mathbf{G}^{12}$ of all coordinates, with $\mathcal{R}$ the seven relations above, so that forms are integer combinations of the free coordinates. The *standard forms* are $ax_v + by_v$ with $(a, b) \in \{(1,0), (0,1), (1,1), (1,-1)\}$ and $0 \le v \le 5$. All other forms are *cross forms*. We write $s_v = x_v + y_v$ and $d_v = x_v - y_v$. The construction gives three further rules.

(CI) For each step, with O the pairs constructed before it, U' the new pairs and V the variables copied over, $H(V, U') + H(O) = H(O, U') + H(V)$. This is the identity of Lemma 2.1.

(EL) For each step and each set S of standard forms of the pairs in U and variables in V, $H[S] = H[\tau S]$, where τ replaces each pair of U by its copy and fixes V. This holds because (U', V) has the law of (U, V).

(INJ) Rule (c) of Lemma 2.3 applies to (x_v, y_v) for $0 \le v \le 5$, because g_v has the law of the injective pair (X, Y).

Proposition 3.2. *Let $a = H(X)$, $b = H(Y)$ and $c = H(X + Y)$. In the configuration* (3.2)*,*

$$6H(X, Y) = 4a + 4b + 3c - \Delta_{\mathrm{KT}},$$

where

$$\Delta_{\mathrm{KT}} = I(y_0; y_1) + I(s_0; y_1) + I\big(x_0; (s_3, s_5, y_4)\big) + I(s_3; y_4) + I\big(s_5; (s_3, y_4)\big) \ge 0.$$

In particular $H(X, Y) \le \frac{2}{3}H(X) + \frac{2}{3}H(Y) + \frac{1}{2}H(X + Y)$ for all injective pairs in all abelian groups.

Proof. Let $L = H(X, Y)$ and $\Omega = (x_0, y_0, y_1, x_2, x_4)$. In step (S1), (CI) and $H(x_0, g_1) = H(g_1) = L$ give $H(g_0, g_1) = 2L - a$. In step (S2), $H(y_0, y_1, g_2, g_3) = H(g_2, g_3) = H(g_0, g_1)$ by the relations and (EL), so (CI) gives $H(g_0, \dots, g_3) = 2H(g_0, g_1) - H(y_0, y_1)$. In the same way step (S3) gives $H(\Omega) = H(g_0, \dots, g_3) + H(g_0, g_1) - H(s_0, y_1)$. Hence

$$H(\Omega) = 6L - 3a - H(y_0, y_1) - H(s_0, y_1).$$

On the other hand $(s_3 - s_5) + (x_0 - y_4) = x_2 - y_0 = d_2$, so by (INJ) the forms x_0, s_3, s_5, y_4 determine x_2 and y_0, then $x_4 = x_2 - s_3 + s_5$ and $y_1 = s_5 - x_4$. Thus $H(\Omega) = H(x_0, s_3, s_5, y_4)$. The pairs g_1, g_3, g_4, g_5 have the law of (X, Y), so $H(y_1) = H(y_4) = b$ and $H(s_3) = H(s_5) = c$, and expanding the five terms gives

$$\Delta_{\mathrm{KT}} = \big(2b - H(y_0, y_1)\big) + \big(b + c - H(s_0, y_1)\big) + \big(a + b + 2c - H(x_0, s_3, s_5, y_4)\big).$$

Comparing the two expressions for $H(\Omega)$, we obtain the identity. No step uses rule (b), so the argument is valid in every abelian group. □

Combined with Lemma 3.1, Proposition 3.2 yields the exponent 11/6. It adapts the counting argument in [13, §3] to entropy.

3.3. **Certificates.** Fix a finite set E of cross forms and take as ground set the standard forms together with E, with the pairs (x_v, y_v) in rule (c). Consider the linear program with one variable $h(F)$ for each closed set F and $h(\emptyset) = 0$, which maximizes $h(\mathrm{cl}\{x_0, y_0\})$ subject to the following constraints, where K runs over closed sets and i, j over the ground set:

(i) $h(\mathrm{cl}(K \cup \{i\})) + h(\mathrm{cl}(K \cup \{j\})) - h(\mathrm{cl}(K \cup \{i, j\})) - h(K) \ge 0$ and $h(\mathrm{cl}(K \cup \{i\})) - h(K) \ge 0$;

(ii) the three identities (CI);

(iii) $h(\mathrm{cl}\, S) = h(\mathrm{cl}\, \tau S)$ for the identifications (EL);

(iv) instances of inequalities proved by Lemma 2.2 for joint variables made of closed sets, each with its proof, where the joint entropy of several closed sets is h of the closure of their union;
(v) $h(\mathrm{cl}\{x_0\}) \le 1$, $h(\mathrm{cl}\{y_0\}) \le 1$ and $h(\mathrm{cl}\{x_0 + y_0\}) \le 1$.

For every injective pair with $\max(a,b,c) > 0$, the vector $F \mapsto H[F]/\max(a,b,c)$ is feasible. A *certificate* consists of a set E, the identifications (iii) that it uses, a finite list of rows R_k of types (i), (ii) and (iv) with rational multipliers μ_k, nonnegative except on rows of type (ii), and rational numbers $\lambda_X, \lambda_Y, \lambda_S \ge 0$, such that

$$h(\mathrm{cl}\{x_0, y_0\}) + \sum_k \mu_k R_k = \lambda_X\, h(\mathrm{cl}\{x_0\}) + \lambda_Y\, h(\mathrm{cl}\{y_0\}) + \lambda_S\, h(\mathrm{cl}\{x_0 + y_0\}) \tag{3.3}$$

holds identically in the variables h after the identifications (iii). At the entropy vector of an injective pair the rows of types (i) and (iv) are nonnegative and the rows of type (ii) vanish, so (3.1) follows from (3.3).

We solve the linear programs with HiGHS [12] in floating point. Rows of type (iv) are added when the current optimum violates them. For the certificate we keep the rows whose floating-point multipliers exceed a tolerance and solve the dual equations on these rows exactly over $\mathbb{Q}$.

3.4. **Proofs of Theorems 1.1 and 1.2.**

Proof of Theorem 1.1. The certificate for Theorem 1.1 [29] uses the configuration (3.2), 41 cross forms and identifications (EL). Its rows are inequalities of both kinds in (i), the identities (CI) and applications of the inequalities that the certificate proves by Lemma 2.2. Its coefficients are the numbers $\lambda_X, \lambda_Y, \lambda_S$ of Theorem 1.1, and the weighted rows satisfy (3.3). The closures use rule (b) of Lemma 2.3, so (3.1) holds for injective pairs in torsion-free groups, and Lemma 3.1 gives (a) and (b). Since $\mathbb{R}$ is torsion-free, $C_{3b} \le \lambda_X + \lambda_Y + \lambda_S$. The proved inequalities include Matúš's inequality for $s \le 62$ and the inequalities numbered 5, 7, 8, 33, 35, 56, 60, 77, 107, 109 in [4, Section VIII], so the proof uses only Shannon inequalities and Lemma 2.1. □

Remark 3.3. This certificate applies rule (b) with $c = 2^k$ only, so Theorem 1.1 holds in all abelian groups without 2-torsion.

Remark 3.4. Among the Shannon inequalities of the certificate for Theorem 1.1, which all have unit coefficients on their entropies, the largest total multipliers fall on those that contain the cross forms $y_0 - x_4$, $x_0 - x_2 - x_4$, $x_0 + y_1 + x_2$ and $x_0 - x_2$. The first and the third of these are among the five cross forms of Appendix A, all of which occur in the certificate.

Proof of Theorem 1.2. The certificate for Theorem 1.2 [29] combines inequalities of type (i) with the identities (CI), after identifications (EL), and uses no row of type (iv). Its coefficients λ are those of Theorem 1.2. For each set of forms in it, the closure with rule (b) equals the closure without it, and the relations of Section 3.2 have integer coefficients. Hence the rules it uses are valid in any abelian group, and Lemma 3.1 applies. □

Remark 3.5. Cross forms are what make the improvement possible: on the configuration (3.2) with standard forms only, the floating-point optimum of the linear program with constraints (i), (ii), (iii) and (v) is 11/6, the value of Proposition 3.2. Adding to the ground set any three of the 65 forms with coefficients ± 1 on at most three free coordinates does not lower this optimum, while the five cross forms of Appendix A do.

4. The two-set exponent

4.1. **Kleinwaks' setting.** Let $X, Y \subset \mathbb{Z}$ be finite and nonempty. Choose $\Gamma \subseteq X \times Y$ with exactly one pair (x, y) satisfying $x - y = z$ for each $z \in X - Y$ [15, Lemma 4.1], let (X_1, Y_1) be uniform on Γ, and let μ and ν be the laws of X_1 and Y_1. Put

$$L = \log|X - Y| = H(X_1, Y_1), \qquad S = \log|X + Y|, \qquad h(i, j) = H(\mu^{*i} * \nu^{*j}) \quad (i, j \geq 0),$$

where $\alpha * \beta$ is the law of $U + V$ for independent $U \sim \alpha$, $V \sim \beta$, and H of a law is the entropy of a random variable with that law. Kleinwaks proved [15, Lemmas 4.2 and 4.3] that for all $i, j \geq 0$

(G1) $h(i+2, j) + h(i, j) \leq 2h(i+1, j)$ and $h(i, j+2) + h(i, j) \leq 2h(i, j+1)$;
(G2) $h(i, j) \leq h(i+1, j)$, $h(i, j) \leq h(i, j+1)$ and $h(i, j) \geq 0$;
(G3) $L + h(i, j) \leq h(i+1, j) + h(i, j+1)$;
(G4) $h(1, 1) \leq S$.

Here (G1) is Madiman's inequality for sums of independent variables [17], and (G3) holds because $(X_1 + V, Y_1 + V)$ determines (X_1, Y_1, V) for V independent of (X_1, Y_1), as $x - y$ determines (x, y) on Γ.

4.2. **The configuration.** Let $D_0 = (x_0, y_0)$ and $D_1 = (x_1, y_1)$ be independent copies of (X_1, Y_1), and let w, v, u be random variables of type $(1, 0)$, that is, with law μ, such that D_0, D_1 and the variables in $\mathcal{W} = \{w, v, u\}$ are mutually independent. We call D_0, D_1 and the variables in $\mathcal{W}$ the *blocks* of the configuration. Each variable in $\mathcal{W}$ carries a *type* (i, j), given with the configuration, and has law $\mu^{*i} * \nu^{*j}$. We apply Section 2.1 to the vector Z of the coordinates x_0, y_0, x_1, y_1 and the variables in $\mathcal{W}$, with $\mathcal{R} = \emptyset$, so that forms are integer combinations of these coordinates. The following rules hold.

(R1) Closure: rules (a)–(c) of Lemma 2.3, with (c) for the pairs (x_k, y_k), $k = 0, 1$.
(R2) Independence: the joint entropy of D_0, D_1 and the variables in $\mathcal{W}$ is the sum of their entropies.
(R3) Symmetry: $H[F] = H[\tau F]$ for all sets F of forms and all permutations τ of the coordinates that map the pairs D_0, D_1 to D_0, D_1 (with x to x and y to y) and each variable in $\mathcal{W}$ to a variable in $\mathcal{W}$ of the same type.
(R4) $H(x_k, y_k) = L$ for $k = 0, 1$.
(R5) Grid identification: if $f = \sum_{k \in I} x_k + \sum_{k \in J} y_k + \sum_{u \in U} u$ with $I \cap J = \emptyset$ and $U \subseteq \mathcal{W}$, then $H(f) = h(i, j)$, where (i, j) is $(|I|, |J|)$ plus the sum of the types of the variables in U.
(R6) Support: $H(a + b) \leq S$ if a is one of x_0, x_1 or a variable in $\mathcal{W}$ of type $(1, 0)$, and b is one of y_0, y_1 or a variable in $\mathcal{W}$ of type $(0, 1)$.

Rule (R1) holds because $\mathbb{Z}$ is torsion-free and $(x_k, y_k) \in \Gamma$. Rule (R3) holds since D_0 and D_1 are independent copies of one pair, variables in $\mathcal{W}$ of the same type are independent copies of one variable, and all of them are independent. Rule (R5) holds because f is a sum of independent variables, each with law μ or ν or one of the laws of the variables in $\mathcal{W}$. Differences such as $w - v$, however, are not identified with entries of h. Rule (R6) holds because x_k and variables of type $(1, 0)$ take values in X, while y_l and variables of type $(0, 1)$ take values in Y. Without $\mathcal{W}$ and without rule (b), rules (R1)–(R6) correspond to the rules (R1)–(R7) of [15, §4.3].

(R7) Reflection: $\Gamma' = \{(-y, -x) : (x, y) \in \Gamma\}$ is a subset of $(-Y) \times (-X)$ with distinct differences, $|\Gamma'| = |X - Y|$ and $|(-Y) + (-X)| = |X + Y|$. Its marginals are the reflections of ν and μ, so the numbers attached to Γ' are L, S and $h'(i, j) = h(j, i)$ [15, proof of Corollary 5.3]. Hence (R1)–(R6) also hold for the configuration built from Γ', in which a variable of type (i, j) has law $\mu'^{*i} * \nu'^{*j}$ for the marginals μ', ν' of Γ', with $h(i, j)$ replaced by $h(j, i)$ in (R5).

(R8) Partial symmetry: $H[\tau F] = H[F]$ for each set F of forms and each permutation τ of the coordinates that maps the coordinates of each block that occur in F into a single block, different blocks into different blocks, and preserves laws in the following sense. If both x_k and y_k occur in F, then $\tau(x_k) = x_m$ and $\tau(y_k) = y_m$ for some m. A coordinate that occurs alone in its block is mapped to one of the same law: x_k and variables of type $(1,0)$, which have law μ, to some x_m or to a variable of type $(1,0)$; y_k and variables of type $(0,1)$, which have law ν, to some y_m or to a variable of type $(0,1)$; and other variables in $\mathcal{W}$ to variables of the same type. The rule holds because the blocks are independent, so the coordinates that occur in F and their images under τ have the same joint law, namely the product of the laws of the pairs and of the single coordinates. The same holds in the configuration built from Γ', with the laws μ', ν' of its marginals in place of μ, ν.

(R9) Partial reflection: $H'[F] = H[\sigma F]$ for each set F of forms of the configuration built from Γ', where H' is entropy in that configuration and σ substitutes $c \mapsto -\tau(c)$ for the coordinates c that occur in F. Here τ maps the coordinates of each block of the reflected configuration that occur in F into a single block of the configuration built from Γ, different blocks into different blocks, as follows. If both x'_k and y'_k occur, then $\tau(x'_k) = y_m$ and $\tau(y'_k) = x_m$ for some m. A coordinate that occurs alone in its block is mapped to one of the transposed law: x'_k and variables of type $(1,0)$ to some y_m or to a variable of type $(0,1)$; y'_k and variables of type $(0,1)$ to some x_m or to a variable of type $(1,0)$; and a variable of type (i,j) to a variable of type (j,i). The rule holds because (x'_k, y'_k) has the law of $(-y_m, -x_m)$, so that x'_k has the law of $-y_m$ and y'_k that of $-x_m$, because a variable of type (i,j) in the reflected configuration has the law of $-u$ for u of type (j,i), and because in both configurations the blocks are independent.

4.3. **Proof of Theorem 1.3.**

Proposition 4.1. *In the setting of Section 4.1,*

$$\begin{aligned} 8\,959\,609\,957\,200\,L \le\ & 13\,623\,518\,368\,355\,S \\ & + 115\,396\,751\,750\,h(0,1) + 500\,183\,187\,638\,h(0,3) \\ & + 248\,591\,603\,812\,h(0,4) + 900\,281\,603\,795\,h(1,2) \\ & + 967\,541\,424\,815\,h(2,1) + 386\,723\,275\,812\,h(3,0) \\ & + 73\,889\,341\,268\,h(3,1) + 156\,315\,592\,518\,h(4,0) \\ & - 115\,396\,751\,750\,h(0,2) - 115\,396\,751\,750\,h(1,1) \\ & - 11\,215\,940\,128\,h(1,3) - 3\,735\,646\,057\,220\,h(2,2). \end{aligned}$$

Proof. The certificate for Theorem 1.3 [29] lists instances of the rules of Section 4.2 with rational multipliers, each for the configuration of Γ or of Γ'. Besides Shannon inequalities among joint entropies of forms and 2 support inequalities (R6), they are identities (R2)–(R5), (R8) and (R9). The multipliers of the inequalities are nonnegative. In the weighted sum, after each $H[F]$ is replaced by $H[\operatorname{cl} F]$, the coefficients of all joint entropies cancel, and what remains is the stated inequality. □

Lemma 4.2. *Let $L \in \mathbb{R}$ and $h\colon \{0,\dots,12\}^2 \to \mathbb{R}$ satisfy (G1)–(G3) whenever all indices involved are at most 12. Then*

$$\begin{aligned}628\,732\,719\,440\,L &+ 115\,396\,751\,750\,h(0,1)\\ &+ 500\,183\,187\,638\,h(0,3) + 248\,591\,603\,812\,h(0,4)\\ &+ 900\,281\,603\,795\,h(1,2) + 967\,541\,424\,815\,h(2,1)\\ &+ 386\,723\,275\,812\,h(3,0) + 73\,889\,341\,268\,h(3,1)\\ &+ 156\,315\,592\,518\,h(4,0)\\ &\le 115\,396\,751\,750\,h(0,2) + 115\,396\,751\,750\,h(1,1)\\ &+ 11\,215\,940\,128\,h(1,3) + 3\,735\,646\,057\,220\,h(2,2).\end{aligned}$$

Proof. The inequality is a sum of instances of (G1)–(G3) with nonnegative rational multipliers, listed in the certificate. □

Proof of Theorem 1.3. If X or Y is empty, both sides vanish. For nonempty $X, Y \subset \mathbb{Z}$, the sum of Proposition 4.1 and Lemma 4.2 is $9\,588\,342\,676\,640\,L \le 13\,623\,518\,368\,355\,S$, that is, $L \le \lambda S$. Finite subsets X, Y of a torsion-free abelian group lie in a subgroup isomorphic to $\mathbb{Z}^d$ for some d. A homomorphism $\phi\colon \mathbb{Z}^d \to \mathbb{Z}$ that is injective on the finite sets $X+Y$ and $X-Y$ exists, because only finitely many hyperplanes must be avoided. Since $\phi(X) \pm \phi(Y) = \phi(X \pm Y)$, the integer case applied to $\phi(X), \phi(Y)$ gives the general case. □

Remark 4.3. The gain over [15] comes from the variables in $\mathcal{W}$: for the configuration of D_0 and D_1 alone, with all forms whose coefficients lie in $\{-1, 0, 1\}$ and rules (R1)–(R7), the floating-point optimum of the Shannon linear program is 1.4621, larger than Kleinwaks' exponent λ_∞ and equal to the optimum of (G1)–(G4) alone for indices up to 12.

4.4. **The transfer theorem.** In [28, page 3a] and [8, Problem 6.44], C_{3a} is the supremum of the t for which there exist arbitrarily large finite $A, B \subset \mathbb{Z}$ with $|A+B| \ll |A|$ and $|A-B| \gg |A+B|^t$ (in [8], $|A-B| \gg |A|^t$), the implied constants being independent of A and B. Gyarmati, Hennecart and Ruzsa [10, (5)] and Kleinwaks [15, §2.2] ask for such sets for every $K > 1$ in $|A+B| \le K|A|$, which can only give a smaller supremum. If $|A+B| \le K|A|$ and $t \ge 0$, then $|A-B| \ge c|A|^t$ implies $|A-B| \ge cK^{-t}|A+B|^t$, so the next theorem, stated with $|A-B| \ge c|A+B|^t$, bounds C_{3a} in each of these readings. Although Kleinwaks formalized it in Lean for his reading [15], we include the proof, which follows [15, §3]. Its main estimate, [15, Proposition 3.5], holds for every K.

Theorem 4.4 (Kleinwaks)**.** *Let $\lambda \ge 1$ be such that $|X-Y| \le |X+Y|^\lambda$ for all finite $X, Y \subset \mathbb{Z}$. Let $K \ge 1$, $c > 0$ and $t \in \mathbb{R}$, and suppose that there are finite nonempty $A, B \subset \mathbb{Z}$ with $|A|$ arbitrarily large, $|A+B| \le K|A|$ and $|A-B| \ge c|A+B|^t$. Then $t \le 2 - 1/\lambda$.*

Proof. Let $A, B \subset \mathbb{Z}$ be finite and nonempty, and write $M_h = |hB - hB|$, where hB is the h-fold sumset.

Packing. For a finite nonempty D, a set $T \subseteq A$ of maximal size such that the translates $t+D$, $t \in T$, are pairwise disjoint satisfies: for each $a \in A$ there is $t \in T$ with $a - t \in D - D$. Otherwise $a \notin T$ and $a + D$ is disjoint from all the translates $t + D$, which contradicts the maximality of T. This is Ruzsa's covering argument [24].

Localization. We show that

$$|A-B|\,|D| \le |D-D-B|^{1-1/\lambda}\,|A+B|\,|D-D+D-B|. \tag{4.1}$$

Assign each $a \in A$ to one $t \in T$ with $a - t \in D - D$, and let A_t be the set of elements assigned to t. Then $A_t - B \subseteq t + (D - D - B)$, so $x = |A_t - B|$ satisfies $x \le U := |D - D - B|$

and, by hypothesis, $x \le |A_t + B|^\lambda$. Hence $x = x^{1-1/\lambda} x^{1/\lambda} \le U^{1-1/\lambda} |A_t + B|$, and $|A - B| \le \sum_t |A_t - B| \le U^{1-1/\lambda} \sum_t |A_t + B|$. If $u \in A_t + B$, say $u = a + b$ with $a \in A_t$ and $a - t = d_1 - d_2$ with $d_1, d_2 \in D$, then $t + D = u - b - d_1 + d_2 + D \subseteq u + (D - D + D - B)$. The translates $t + D$ with $u \in A_t + B$ are disjoint, so there are at most $|D - D + D - B|/|D|$ of them, and $\sum_t |A_t + B| \le |A + B|\,|D - D + D - B|/|D|$.

One scale. Take $D = hB - hB$ with $h \ge 1$. Then $D - D - B = 2hB - (2h+1)B$ and $D - D + D - B = 3hB - (3h+1)B$. Since $B \ne \emptyset$, $|mB - nB|$ does not decrease in m or n, so both sets have at most M_{4h} elements, and (4.1) gives $|A - B|\, M_h \le M_{4h}^{2-1/\lambda} |A + B|$.

Several scales. Let $d = |A - B|$, $s = |A + B|$ and $m_j = M_{4^j}$. By induction on j, $d^j m_0 \le s^j m_j^{1+j(1-1/\lambda)}$: the case $j = 0$ is an equality, and the step multiplies by $d\, m_j \le m_{j+1}^{2-1/\lambda} s$ and uses $m_j \le m_{j+1}$. Now $m_0 \ge 1$, and if $|A + B| \le K|A|$, the Plünnecke–Ruzsa inequality $|kB - lB| \le K^{k+l}|A|$ [23, Theorem 1.2] gives $m_J \le K^{2\cdot 4^J} |A| \le K^{2\cdot 4^J} s$. Hence, for every $J \ge 1$,

$$d \le s \left(K^{2\cdot 4^J} s\right)^{(1+J(1-1/\lambda))/J} = C_{K,J}\, s^{2-1/\lambda+1/J}.$$

For the sets of the theorem $s \ge |A|$ is arbitrarily large and $d \ge cs^t$, so $t \le 2 - 1/\lambda + 1/J$. Letting $J \to \infty$, we conclude that $t \le 2 - 1/\lambda$. □

Proof of Corollary 1.4. The number $\lambda \ge 1$ of Theorem 1.3 satisfies the hypothesis of Theorem 4.4, so all exponents t in the definition of C_{3a} satisfy $t \le 2 - 1/\lambda$. □

5. Verification

Each certificate is checked by two programs, written independently of each other, neither of which imports the programs that found the certificates. For a certificate of Section 3 or Appendix A, the first checker recomputes the relations of each copy step from Lemma 2.1 and the closures by exact elimination over $\mathbb{Q}$ together with rule (INJ), rebuilds each row from its definition, merges the closed sets identified by (EL), checks the proof of each inequality that the certificate proves by Lemma 2.2, including the absence of cycles among these proofs, and compares the two sides of (3.3) coefficient by coefficient. For the certificate of Section 4, it rebuilds the closures with (R1), checks that each row is an instance of its rule (each permutation or partial map respects pairs, types and laws as required, each identification (R5) has the indices of its form, transposed for Γ', each support row has the form required in (R6), and the grid indices are at most 12), and verifies that the weighted rows sum to $\lambda S - L$. Both checkers reject negative multipliers on inequality rows. The second checkers represent the configurations by explicit parametrizations and closed sets by reduced row echelon bases; for the certificates of Section 3 we confirmed that the two checkers compute the same relations and the same closures. For the inequalities proved inside the certificate of Theorem 1.1, the second checker re-implements the check of the first, following the same steps; those from the list of Dougherty, Freiling and Zeger are also re-derived from their copy steps by an exact linear program, and all of them are compared with the published formulas.

Theorems 1.1–1.3, Corollary 1.4 and the bound of Remark 1.5 are also proved in Lean 4 [22] with Mathlib [18]. The proofs build on Kleinwaks' Lean library [15], which we ported to a newer version of Lean and Mathlib without changing its statements. A program translates each certificate into Lean source: each row becomes a lemma proved from general facts, two sets of forms with the same closure are shown to determine each other by explicit derivations with the rules (a)–(c) of Lemma 2.3, which the kernel checks by evaluation, the tactic `linear_combination` checks the weighted sum of the rows, and each inequality that the certificate of Theorem 1.1 proves by Lemma 2.2 becomes a Lean theorem about arbitrary random variables. The formal

statements, listed in the supplement [29] with the definitions they use, do not mention the certificates, so the translation need not be trusted. They are parts (a) and (b) of Theorem 1.1 over $\mathbb{R}$ and in every abelian group without 2-torsion, the same parts with the numbers of Theorem 1.2 in every abelian group, the bound of Remark 1.5, Theorem 1.3, and Corollary 1.4 in the readings of Section 4.4 with fixed K and c and with every $K > 1$. They depend only on the three standard axioms of Lean.

Appendix A. A short certificate

This appendix works through a certificate that is small enough to be checked by hand. The short certificate in [29] proves

$$H(X,Y) \le \tfrac{314}{459}H(X) + \tfrac{302}{459}H(Y) + \tfrac{225}{459}H(X+Y)$$

for injective pairs in any abelian group. By Lemma 3.1 it gives the exponent $841/459 < 1.8322441$ in all abelian groups.

The forms. Besides the standard forms of the configuration (3.2), the certificate uses the five cross forms

$$y_0 - x_4, \quad x_0 + y_0 + x_2, \quad x_0 + y_1 + x_2, \quad x_0 + y_1 - x_4, \quad y_0 + y_1 + x_2.$$

Each of them is built from two pairs of the configuration that share a coordinate or a sum. The form $x_0 + y_0 + x_2$ is the sum of the three coordinates of $g_0 = (x_0, y_0)$ and $g_2 = (x_2, y_0)$, which share y_0; in the same way $x_0 + y_1 + x_2$ belongs to g_1 and g_3, which share y_1, and $y_0 + y_1 + x_2$ to g_2 and g_3, which share x_2. The form $x_0 + y_1 - x_4 = s_1 - x_4$ combines g_1 and g_5, which share y_1, and $y_0 - x_4 = y_4 - x_0$ combines g_0 and g_4, which have the same sum $s_0 = s_4$.

The identity. Let $\mathrm{CI}_t = H(V,U') + H(O) - H(O,U') - H(V)$ be the expression of (CI) for step (St), and let I_k be the 72 mutual informations of Table 2 with weights w_k. Then

$$459\,H(X,Y) + \sum_k w_k I_k - 220\,\mathrm{CI}_1 - 79\,\mathrm{CI}_2 - 69\,\mathrm{CI}_3 = 314\,H(X) + 302\,H(Y) + 225\,H(X+Y) \tag{A.1}$$

holds after each joint entropy is replaced by the entropy of the closure of its forms and the identifications of Table 1 are applied. Since $I_k \ge 0$ and $\mathrm{CI}_t = 0$, (A.1) gives the stated inequality. In the certificate file each I_k is expanded by the chain rule into inequalities of the first kind in (i), all with the weight $w_k/459$.

Checking the identity. Write each $I_k = I(A; B \mid C)$ as $H(A,C) + H(B,C) - H(A,B,C) - H(C)$, and write CI_t out with the pairs O, U' and the variables V of step (St). The closures need only rules (a) and (c): starting from the integer combinations of the given forms, add the two coordinates of a pair g_v whenever d_v is such a combination, and repeat. For example, $d_5 = x_4 - y_1 = (y_0 + x_2) - (y_1 + x_2) - (y_0 - x_4)$, so the closure of $\{y_1 + x_2,\ y_0 - x_4,\ y_0 + x_2\}$ contains x_4 and y_1, and hence also y_0 and x_2. Next, merge the closed sets that Table 1 identifies; each identification is an instance of (EL) for the step shown. After these replacements every entropy except those of x_0, y_0 and $x_0 + y_0$ and of the sets identified with them cancels, and the remaining coefficients are 314, 302 and 225.

What it says about Δ_{KT}. Since $\mathrm{CI}_t = 0$, subtracting 76.5 times the identity of Proposition 3.2 from (A.1) gives

$$\sum_k w_k I_k = 76.5\,\Delta_{\mathrm{KT}} + 8H(X) - 4H(Y) - 4.5H(X+Y),$$

so that

$$153\,\Delta_{\mathrm{KT}} \ge 8H(Y) + 9H(X+Y) - 16H(X).$$

In particular $\Delta_{\mathrm{KT}} \ge 1/153$ when $H(X) = H(Y) = H(X+Y) = 1$.

TABLE 1. The identifications (EL) of the short certificate, by copy step.

Step	Identification	Step	Identification
1	$H[y_0] = H[y_1]$	2	$H[(x_0 + y_1, x_0 + y_0)] = H[(y_1 + x_2, y_0 + x_2)]$
1	$H[x_0 + y_0] = H[x_0 + y_1]$	2	$H[(y_1, x_0 + y_0)] = H[(y_1, y_0 + x_2)]$
1	$H[x_0 - y_0] = H[x_0 - y_1]$	3	$H[x_0] = H[x_4]$
2	$H[x_0] = H[x_2]$	3	$H[y_0] = H[x_0 + y_0 - x_4]$
2	$H[x_0 + y_0] = H[y_0 + x_2]$	3	$H[x_0 - y_0] = H[x_0 + y_0 - 2x_4]$
2	$H[x_0 + y_1] = H[y_1 + x_2]$	3	$H[x_0 - y_1] = H[y_1 - x_4]$
2	$H[(x_0 - y_1, y_0)] = H[(y_1 - x_2, y_0)]$	3	$H[(x_0 - y_1, y_0)] = H[(y_1 - x_4, x_0 + y_0 - x_4)]$
2	$H[(x_0 + y_1, y_0)] = H[(y_1 + x_2, y_0)]$	3	$H[(x_0 + y_1, y_0)] = H[(y_1 + x_4, x_0 + y_0 - x_4)]$

TABLE 2. The 72 Shannon inequalities $I_k \geq 0$ of the short certificate with their weights w_k. A tuple in parentheses is one random variable, and a list after the bar is the conditioning tuple.

w_k	I_k	w_k	I_k
81	$I(y_0; y_1)$	2	$I(y_0; x_4 \mid x_0 + y_1 + x_2)$
68	$I(x_0 + y_0; y_1)$	2	$I(x_0 + y_1; x_4 \mid y_0 + x_2)$
38	$I(y_0; x_0 + y_1)$	2	$I(y_0 - x_4; x_0 + y_0 + x_2 \mid x_0 + y_1 + x_2)$
22	$I(x_0 + y_0; x_0 + y_1)$	1	$I(y_1; y_0 + x_2 \mid x_4)$
20	$I(x_0; y_0 + x_2)$	10	$I(y_0; y_1 \mid x_0, x_2)$
19	$I(x_0 + y_0; x_2)$	7	$I(x_0; x_4 \mid y_1, x_0 + y_0)$
17	$I(x_0 + y_0 + x_2; (y_1 + x_4, x_0 + y_0 - x_4))$	5	$I(y_1; x_2 \mid x_0, y_1 + x_4)$
14	$I(x_0 + y_1 + x_2; (y_1 + x_4, x_0 + y_0 - x_4))$	5	$I(x_2; x_0 \mid y_1 - x_4, x_0 + y_0 + x_2)$
13	$I(x_0; x_0 + y_0 - x_4)$	5	$I(x_0; x_0 + y_0 - x_4 \mid x_2, y_1 + x_4)$
12	$I(x_0; y_1 + x_2)$	5	$I(y_0 - x_4; y_1 + x_4 \mid x_0, y_0 + x_2)$
12	$I(x_0 + y_1; x_2)$	5	$I(x_0 + y_1; y_0 - x_4 \mid x_2, y_1 + x_4)$
11	$I(y_0; x_4)$	4	$I(x_0; y_0 + x_2 \mid x_0 + y_1, x_0 + y_0 - 2x_4)$
11	$I(y_0 - x_4; (y_0 + x_2, y_1 + x_2))$	4	$I(x_0; y_1 \mid x_2, x_0 + y_0 - 2x_4)$
9	$I(x_0 + y_0 - x_4; (y_0 + x_2, y_1 + x_2))$	3	$I(x_0; x_2 \mid y_0, y_1 - x_4)$
7	$I(x_0; x_0 + y_1 - x_4)$	3	$I(x_0; y_1 + x_2 \mid x_4, y_0 + x_2)$
7	$I(x_0 + y_1; x_4)$	3	$I(y_0; y_1 + x_2 \mid x_4, x_0 + y_0 + x_2)$
5	$I(y_0 + x_2; (x_0, y_1 + x_4))$	2	$I(x_0; x_4 \mid y_1, x_0 + y_0 + x_2)$
5	$I(x_0 + y_0 + x_2; y_0 + y_1 + x_2)$	2	$I(x_0; x_0 + y_0 - x_4 \mid x_0 + y_1, y_0 + x_2)$
3	$I(x_0 + y_0 - x_4; (x_0, y_0 + x_2))$	2	$I(x_0 + y_0; y_1 \mid x_4, x_0 + y_1 + x_2)$
2	$I(x_0; (y_0 + x_2, y_1 + x_2))$	2	$I(x_0; y_0 \mid y_1 - x_4, x_0 + y_1 + x_2)$
2	$I(x_0; (y_0 + x_2, x_0 + y_1 - x_4))$	2	$I(x_0; x_0 + y_0 - x_4 \mid x_0 + y_0 + x_2, x_0 + y_1 + x_2)$
2	$I(y_1 + x_2; (x_0, y_0 - x_4))$	2	$I(x_0; y_0 \mid y_0 - x_4, x_0 + y_1 + x_2)$
2	$I(x_0 + y_1 + x_2; y_0 + y_1 + x_2)$	2	$I(x_0; x_2 \mid x_0 + y_1 + x_2, x_0 + y_0 - 2x_4)$
19	$I(x_0; x_2 \mid x_0 + y_0 + x_2)$	2	$I(y_1; x_0 + y_0 - x_4 \mid y_0, x_0 + y_1 + x_2)$
12	$I(x_0; x_2 \mid x_0 + y_1 + x_2)$	2	$I(x_4; y_1 + x_2 \mid y_0 + x_2, x_0 + y_1 - x_4)$
11	$I(x_0; y_0 \mid y_0 - x_4)$	1	$I(x_0; x_4 \mid y_1, y_0 + x_2)$
10	$I(x_0; y_1 + x_4 \mid x_2)$	5	$I(y_0; x_0 \mid x_2, y_1 + x_4, y_0 - x_4)$
9	$I(x_2; x_4 \mid x_0 - y_0)$	5	$I(y_1; y_0 \mid x_2, y_1 + x_4, x_0 + y_0 - x_4)$
7	$I(y_0; y_1 \mid y_0 + y_1 + x_2)$	5	$I(x_0; y_0 \mid x_2, x_0 + y_1, y_1 + x_4)$
7	$I(x_4; (y_0 + x_2, y_1 + x_2) \mid x_0 + y_1 - x_4)$	3	$I(x_0; y_1 \mid y_0, x_4, x_0 + y_0 + x_2)$
4	$I(x_0 + y_1; x_2 \mid x_0 + y_0 - 2x_4)$	3	$I(y_0; x_0 + y_0 \mid x_4, y_0 + x_2, y_1 + x_2)$
3	$I(x_0; x_4 \mid x_0 + y_0 + x_2)$	3	$I(y_1; x_0 \mid x_4, y_1 + x_2, x_0 + y_0 + x_2)$
3	$I(y_0; x_0 + y_0 + x_2 \mid y_1 - x_4)$	2	$I(y_0; y_1 \mid x_0, y_0 + x_2, y_0 - x_4)$
2	$I(x_2; x_4 \mid x_0 - y_1)$	2	$I(x_0; x_2 \mid y_0, x_0 + y_0 - x_4, x_0 + y_1 + x_2)$
2	$I(x_4; y_1 + x_2 \mid x_0 - y_0)$	2	$I(x_2; y_0 \mid x_0 + y_0 - x_4, x_0 + y_0 + x_2, x_0 + y_1 + x_2)$
2	$I(y_0; x_0 + y_0 + x_2 \mid y_1)$	2	$I(x_4; y_0 \mid x_0 + y_1, y_0 + x_2, x_0 + y_0 - x_4)$

Appendix B. The constants of Theorem 1.1

The rational numbers of Theorem 1.1, which are the coefficients of its certificate, are

$$\lambda_X = \frac{925\,252\,334\,339\,885\,103\,560\,423\,126\,733\,856\,553}{1\,309\,548\,008\,918\,246\,261\,431\,347\,658\,091\,038\,786},$$
$$\lambda_Y = \frac{424\,585\,766\,727\,935\,297\,449\,756\,583\,005\,066\,048}{654\,774\,004\,459\,123\,130\,715\,673\,829\,045\,519\,393},$$
$$\lambda_S = \frac{1\,237\,016\,505\,223\,052\,067\,896\,560\,048\,548\,494\,601}{2\,619\,096\,017\,836\,492\,522\,862\,695\,316\,182\,077\,572},$$

and

$$\lambda_X + \lambda_Y + \lambda_S = \frac{4\,785\,864\,240\,814\,563\,464\,816\,432\,634\,036\,471\,899}{2\,619\,096\,017\,836\,492\,522\,862\,695\,316\,182\,077\,572} < 1.8272963.$$

Data availability

The certificates (`c3b_torsionfree.json` for Theorem 1.1, `c3b_allgroups.json` for Theorem 1.2, `c3b_short.json` for Appendix A and `pair_exponent.json` for Theorem 1.3), the checkers, the Lean proofs and the programs that found the certificates are archived in [29] and available at `https://github.com/tejasvi/sum-difference-exponents/tree/v4`.

Acknowledgements

Generative AI tools were used to write code, run analyses and draft text; the author directed the work and takes full responsibility for its content.

References

[1] T. F. Bloom, *Erdős problem #1097*, `https://www.erdosproblems.com/1097`, accessed October 20, 2026; the problem was posed at the 1989 Great Western Number Theory problem session.
[2] J. Bourgain, *On the dimension of Kakeya sets and related maximal inequalities*, Geom. Funct. Anal. **9** (1999), 256–282.
[3] E. P. Csirmaz and L. Csirmaz, *Information inequalities for five random variables*, Computation **14** (2026), no. 2, article 42; arXiv:2512.23316.
[4] R. Dougherty, C. Freiling and K. Zeger, *Non-Shannon information inequalities in four random variables*, arXiv:1104.3602, 2011.
[5] A. Farnamfar, *Sharp optimality of* 11/6 *for literal ladder constructions*, Zenodo, doi:10.5281/zenodo.23337952, 2026.
[6] A. Farnamfar, *Sharp colour bounds and structural obstructions for three-direction sums–differences constructions*, Zenodo, doi:10.5281/zenodo.23385419, 2026.
[7] G. A. Freiman and V. P. Pigarev, *The relation between the invariants R and T* (Russian), Kalinin. Gos. Univ., Moscow, 1973, 172–174.
[8] B. Georgiev, J. Gómez-Serrano, T. Tao and A. Z. Wagner, *Mathematical exploration and discovery at scale*, arXiv:2511.02864, 2025; published as *Mathematical discovery and exploration can be done at scale*, Proc. Natl. Acad. Sci. USA **123** (2026), no. 40, e2536158123. Problem numbers refer to the arXiv version.
[9] B. Green and I. Z. Ruzsa, *On the arithmetic Kakeya conjecture of Katz and Tao*, Period. Math. Hungar. **78** (2019), 135–151.
[10] K. Gyarmati, F. Hennecart and I. Z. Ruzsa, *Sums and differences of finite sets*, Funct. Approx. Comment. Math. **37** (2007), 175–186.
[11] F. Hennecart, G. Robert and A. Yudin, *On the number of sums and differences*, Astérisque **258** (1999), 173–178.
[12] Q. Huangfu and J. A. J. Hall, *Parallelizing the dual revised simplex method*, Math. Program. Comput. **10** (2018), 119–142.
[13] N. H. Katz and T. Tao, *Bounds on arithmetic projections, and applications to the Kakeya conjecture*, Math. Res. Lett. **6** (1999), 625–630.
[14] N. H. Katz and T. Tao, *New bounds for Kakeya problems*, J. Anal. Math. **87** (2002), 231–263.

[15] L. J. Kleinwaks, *Improved sum-difference inequalities in abelian groups*, proof paper, version 1.1 (October 6, 2026), and Lean formalization, `https://github.com/kleinwaks/sum-difference`, archived as doi:10.5281/zenodo.23200803.
[16] M. Lemm, *New counterexamples for sums-differences*, Proc. Amer. Math. Soc. **143** (2015), 3863–3868.
[17] M. Madiman, *On the entropy of sums*, in: Proc. IEEE Information Theory Workshop, Porto, 2008, 303–307.
[18] The mathlib Community, *The Lean mathematical library*, in: Proc. 9th ACM SIGPLAN Int. Conf. Certified Programs and Proofs (CPP 2020), 2020, 367–381.
[19] F. Matúš, *Infinitely many information inequalities*, in: Proc. IEEE Int. Symp. Information Theory, Nice, 2007, 41–44.
[20] Mosaic Intelligence, *A mechanized certificate ladder and a certified Shannon wall for the Katz–Tao sum–difference problem*, Zenodo, doi:10.5281/zenodo.20646386, 2026.
[21] Mosaic Intelligence, *Certified sum-difference lower-bound certificates for C_{3a}, C_{3b}, and C_{3c}*, certificate archive, doi:10.5281/zenodo.20794135, 2026.
[22] 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.
[23] G. Petridis, *New proofs of Plünnecke-type estimates for product sets in groups*, Combinatorica **32** (2012), 721–733.
[24] I. Z. Ruzsa, *An analog of Freiman's theorem in groups*, Astérisque **258** (1999), 323–326.
[25] I. Z. Ruzsa, *Sumsets and entropy*, Random Structures Algorithms **34** (2009), 1–10.
[26] T. Tao, *Sumset and inverse sumset theory for Shannon entropy*, Combin. Probab. Comput. **19** (2010), 603–639.
[27] T. Tao, *Sum-difference exponents for boundedly many slopes, and rational complexity*, arXiv:2511.15135, 2025.
[28] T. Tao et al., *Optimization constants in mathematics*, `https://teorth.github.io/optimizationproblems/`, pages 3a and 3b, accessed October 19, 2026.
[29] T. S. Tomar, *Certificates, programs and Lean proofs for "Certified upper bounds for two sum-difference exponents"*, Zenodo, doi:10.5281/zenodo.23716681, 2026.
[30] Z. Zhang and R. W. Yeung, *On characterization of entropy function via information inequalities*, IEEE Trans. Inform. Theory **44** (1998), 1440–1452.

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