PROPERTIES OF TWO DIMENSIONAL SETS WITH SMALL SUMSET

DAVID GRYNKIEWICZ AND ORIOL SERRA

Abstract. We give tight lower bounds on the cardinality of the sumset of two finite, nonempty subsets A, $B \subseteq \mathbb{R}^2$ in terms of the minimum number $h_1(A,B)$ of parallel lines covering each of A and B. We show that, if $h_1(A,B) \geq s$ and $|A| \geq |B| \geq 2s^2 - 3s + 2$, then

$$|A+B| \geq |A| + (3 - \frac{2}{s})|B| - 2s + 1.$$

More precise estimations are given under different assumptions on $|A|$ and $|B|$.

This extends the 2-dimensional case of the Freiman 2^d–Theorem to distinct sets A and B, and, in the symmetric case $A = B$, improves the best prior known bound for $|A| = |B|$ (due to Stanchescu, and which was cubic in s) to an exact value.

As part of the proof, we give general lower bounds for two dimensional subsets that improve the 2-dimensional case of estimates of Green and Tao and of Gardner and Gronchi, related to the Brunn-Minkowski Theorem.

1. Introduction

Given a pair of finite subsets A and B of an abelian group G, their Minkowski sum, or simply sumset, is $A + B = \{a + b \mid a \in A, b \in B\}$. Furthermore, if $G = \mathbb{R}^d$ and H is a subspace, then we let $\phi_H : \mathbb{R}^d \to \mathbb{R}^d/H$ denote the natural projection modulo H, and we let $h_{d-1}(A,B)$ be the minimal number s such that there exist $2s$ (not necessarily distinct) parallel hyperplanes, $H_1, \ldots, H_s, H'_1, \ldots, H'_s$, with $A \subseteq \bigcup_{i=1}^{s} H_i$ and $B \subseteq \bigcup_{i=1}^{s} H'_i$. Alternatively, $h_{d-1}(A,B)$ is the minimal s such that there exists a $(d-1)$–dimensional subspace H with $|\phi_H(A)|$, $|\phi_H(B)| \leq s$.

It is the central goal of inverse additive theory to describe the structure of sumsets and their summands. One of the most classical results is the Freiman 2^d–Theorem [5] [1] [11] [15], which says that a subset of $\mathbb{R}^d$ with small sumset must be contained in a small number of parallel hyperplanes.

Theorem A (Freiman 2^d–Theorem)**.** *Let $d \geq 2$ be an integer and let $0 < c < 2^d$. There exist constants $k = k(c,d)$ and $s = s(c,d)$ such that if $A \subseteq \mathbb{R}^d$ is a finite, nonempty subset satisfying $|A| \geq k$ and $|A + A| < c|A|$, then $h_{d-1}(A,A) < s$.*

From the pigeonhole principle, one then easily infers there must exist a hyperplane H such that $|H \cap A| \geq \frac{1}{s-1}|A|$, thus containing a significant fraction of the elements of A. In fact, this corollary is sometimes given as the statement of the Freiman 2^d–Theorem itself, in part because it can be shown to easily imply the version given above, illustrating the close dual relationship between being covered by a small number of hyperplanes and having a large intersection with a hyperplane.

Supported in part by the National Science Foundation, as an MPS-DRF postdoctoral fellow, under grant DMS-0502193.

Supported by the Spanish Research Council under project MTM2005-08990-C02-01 and by the Catalan Research Council under project 2005SGR00256.

The Freiman 2^d–Theorem was one of the main tools used in the original proof (as given by Bilu [1] and Freiman [6] [5]) of Freiman's Theorem (a result which shows that any subset $A \subseteq \mathbb{Z}$ with $|A+A| \leq C|A|$ must be a large subset of a multidimensional progression), which has become one of the foundational centerpieces in inverse additive theory. However, like Freiman's Theorem itself, it suffers from lacking even asymptotically correct constants. Remedying such a drawback would greatly magnify the applicability of these results, and in the case of Freiman's Theorem, much effort has been so invested culminating in the achievement by Chang [3] of values that are now almost asymptotically correct.

With the Freiman 2^d–Theorem, there has been less notable success in improving the constants. When $d = 2$ (so that a hyperplane is just a line), independent proofs of the result were found by Fishburn [4] and by Stanchescu [14], with the latter method yielding an optimal value for $s(c, d)$ (specifically, $s = s(c, 2)$ is the ceiling of the smaller root defined by $c|A| = 4|A| + 1 - 2(s + \frac{|A|}{s})$), though the value for $k(c, d)$ was still not asymptotically accurate (the constant obtained was cubic in s rather than quadratic).

The main result of this paper is the following, which extends the 2-dimensional case of the Freiman 2^d–Theorem to distinct sets while at the same time giving exact values for the constants (when $||A| - |B|| \leq s$).

Theorem 1.1. *Let $s \geq 2$ be an integer, and let A, $B \subseteq \mathbb{R}^2$ be finite subsets.*

(i) *If $||A| - |B|| \leq s$, $|A| + |B| \geq 4s^2 - 6s + 3$, and*

$$|A + B| < (2 - \frac{1}{s})(|A| + |B|) - 2s + 1, \tag{1}$$

then $h_1(A, B) < s$.

(ii) *If $|A| \geq |B| + s$, $|B| \geq 2s^2 - \frac{7}{2}s + \frac{3}{2}$, and*

$$|A + B| < |A| + (3 - \frac{2}{s})|B| - s, \tag{2}$$

then $h_1(A, B) < s$.

A slightly less precise but immediate corollary is the following.

Corollary 1.2. *Let $s \geq 2$ be an integer. If A, $B \subseteq \mathbb{R}^2$ are finite subsets with $|A| \geq |B| \geq 2s^2-3s+2$ and $h_1(A, B) \geq s$, then*

$$|A + B| \geq |A| + (3 - \frac{2}{s})|B| - 2s + 1.$$

The following example shows that, for $s \geq 3$, the constant in Theorem 1.1(i) is best possible: let T be a right isosceles triangle in the integer lattice whose equal length sides each cover $x = 2s - 2$ lattice points; then $|T| = (s - 1)(2s - 1)$ and $|2T| = 2(s - 1)(4s - 5) < 4|T| + 1 - 2s - 2\frac{|T|}{s}$, but T is covered by no fewer than $2s - 2 > s - 1$ parallel lines. The same example shows that, even when $|A| + |B| < 4s^2 - 6s + 3$ and $h_1(A, B) \geq s$, the lower bound on $|A + B|$ implied by Theorem 1.1 (i) is quite accurate. Indeed, when $x \geq s$, we have $|T| = \frac{x(x+1)}{2} \geq \frac{s(s+1)}{2}$, $h_1(T, T) \geq s$ and

$$|2T| = x(2x - 1) = 4|T| + \frac{3}{2} - 3\sqrt{\frac{1}{4} + 2|T|}.$$

On the other hand, for $|A|+|B|<4s^2-6s+3$ and $h_1(A,B)\geq s$, one can always choose $s_0<s$ so that the hypothesis of Theorem 1.1 hold. Let $t_0=\frac{1}{2}\sqrt{\frac{1}{4}+|A|+|B|}-\frac{1}{4}$, and let $s_0=\lceil t_0\rceil=t_0+z$, with $0\leq z<1$. Note that $|A|+|B|=4(t_0+1)^2-6(t_0+1)+2>4s_0^2-6s_0+2$. When $|A|+|B|\geq 14$, by applying Theorem 1.1 with s_0, the resulting bound (as a function of z) is minimized for $z=0$. Consequently, we obtain the estimate

$$|A+B|\geq 2|A|+2|B|+\frac{1}{2}-3\sqrt{\frac{1}{4}+|A|+|B|}$$

when $14\leq|A|+|B|\leq 4s^2+2s$, $h_1(A,B)\geq s$, and either $||A|-|B||\leq s_0$ or else $||A|-|B||\leq\lceil\frac{s}{2}\rceil$ and $s(s+1)\leq|A|+|B|$. This shows that the resulting bound for $|A+B|$ using s_0 is surprisingly accurate for $|A|+|B|\geq s(s+1)$. However, once $|A|+|B|<s(s+1)$, the lower bound for $|A+B|$ assuming $h_1(A,B)\geq s$ should begin to become much larger.

The proof of Theorem 1.1 will be given in Section 5, along with the proof of the dual formulation bounding $|A+B|$ when A and B are assumed to contain no s collinear points. Concerning the case $s=2$, a result of Ruzsa [13], generalizing to distinct sets yet another result of Freiman [5, Eq. 1.14.1] [15], shows that if A, $B\subseteq\mathbb{R}^d$ with $|A|\geq|B|$ and $A+B$ d-dimensional, then $|A+B|\geq|A|+d|B|-\frac{d(d+1)}{2}$.

However, as the Freiman 2^d–Theorem indicates, the cardinality of A and B modulo appropriate subspaces also plays an important role contributing to the cardinality of $A+B$. Section 2 is devoted to proving Theorem 1.3 below, which gives a general lower bound for $|A+B|$ based upon $|\phi_H(A)|$ and $|\phi_H(B)|$, with $H=\mathbb{R}x_1$ an arbitrary one-dimensional subspace. It will be a key ingredient in the proof of Theorem 1.1. We remark that the symmetric case (when $A=B$) was first proved by Freiman [5, Eq. 1.15.4].

Theorem 1.3. *Let A, $B\subseteq\mathbb{R}^2$ be finite, nonempty subsets, let $\ell=\mathbb{R}x_1$ be a line, let m be the number of lines parallel to ℓ which intersect A, and let n be the number of lines parallel to ℓ that intersect B. Then*

$$|A+B|\geq(\frac{|A|}{m}+\frac{|B|}{n}-1)(m+n-1). \tag{3}$$

Furthermore, the following bounds are implied by (3).

(i) *If $m\geq n$ and $|A|\leq|B|+m$, then*

$$|A+B|\geq(2-\frac{1}{m})(|A|+|B|)-2m+1.$$

(ii) *If $|A|\geq|B|+m$, then*

$$|A+B|\geq|A|+(3-\frac{2}{m})|B|-m.$$

(iii) *If $1<m<|A|$, let l be an integer such that $\frac{l(l-1)}{m(m-1)}\leq\frac{|B|}{|A|-m}\leq\frac{l(l+1)}{m(m-1)}$, and if $m=1$, let $l=1$. Then*

$$|A+B|\geq|A|+|B|+\frac{l-1}{m}|A|+\frac{m-1}{l}|B|-(m+l-1).$$

(iv) *In general,*

$$|A+B|\geq|A|+|B|+2\sqrt{(m-1)(\frac{|A|}{m}-1)|B|}-(\frac{|A|}{m}+m)+1.$$

Note $l = \lfloor\sqrt{\frac{1}{4} + \frac{(m-1)|B|}{|A|/m-1}} + \frac{1}{2}\rfloor$ satisfies the hypotheses of Theorem 1.3(iii) for $m < |A|$. We remark that Theorem 1.3(iv), along with the compression techniques of Section 2, easily implies (a diagonal compression along $x_1 - x_2$ should also be used when A is contained in two lines, $y_1 + \mathbb{R}x_1$ and $y_2 + \mathbb{R}x_2$, each containing $\frac{|A|+1}{2}$ points of A) the 2-dimensional case of a discrete analog of the Brunn-Minkowski Theorem given by Gardner and Gronchi [7, Theorem 6.6, roles of A and B reversed].

Also, (3) improves the 2-dimensional case of an estimate of Green and Tao [8, Theorem 2.1], with the two bounds equal only when A is a rectangle. In Section 2.2, we briefly exhibit how the discrete methods can be adapted to the continuous case by giving a simple proof of a generalization of the Brunn-Minkowski Theorem, for 2 dimensions, related to Bonneson's generalization of the Brunn-Minkowski Theorem (see eq. (39) in [7]).

The lower bounds for $|A + B|$ from Theorem 1.1(ii) and Theorem 1.3(ii) are estimates based on $\min\{|A|, |B|\}$, much like nearly all other existing estimates for distinct sumsets; however, if $|A|$ is much larger than $|B|$, such bounds can be weak. The bounds in Theorem 1.3(iii) and Theorem 1.3(iv) are more accurate since they take into account the relative size of $|A|$ and $|B|$. It would be desirable to have a similar refinement to Theorem 1.1, i.e., a lower bound for $|A + B|$ based off the parameter $s \le h_1(A, B)$ *and* the relative size of $|A|$ and $|B|$. One possibility would be if the bound in Theorem 1.3(iii) held with the globally defined parameter $s \le h_1(A, B)$ in place of m, for $|A|$ and $|B|$ suitably large with respect to s. This is achieved by Theorem 1.1(i) for the extremal case when $|A|$ and $|B|$ are very close in size. Theorem 1.4 below accomplishes the same aim for the other extremal case, when $|A|$ is much larger than $|B|$. It is proved in Section 6.

Note that the coefficient of $|B|$ in the bound below is much larger than the value of $3-\frac{2}{s}$ obtained from Theorem 1.1(ii). Moreover, the bound on $|B|$ required to apply Theorem 1.4(b) is much smaller than the corresponding requirement for Theorem 1.1, being linear in s rather than quadratic. In fact, Theorem 1.4(a) shows that, by only increasing slightly the requirement of $|A|$ to be much larger than $|B|$—from $|A| \ge \frac{1}{2}s(s-1)|B| + s$ to $|A| > \frac{1}{8}(2s-1)^2|B| - \frac{1}{4}(2s-1) + \frac{(s-1)^2}{2(|B|-2)}$—one can eliminate all need for $|A|$ and $|B|$ to be sufficiently large with respect to s.

Theorem 1.4. *Let s be a positive integer, and let A, $B \subseteq \mathbb{R}^2$ be finite, nonempty subsets with $h_1(A, B) \ge s$ and $|A| \ge \frac{1}{2}s(s-1)|B| + s$. If either*

(a) $|A| > \frac{1}{8}(2s-1)^2|B| - \frac{1}{4}(2s-1) + \frac{(s-1)^2}{2(|B|-2)}$, or

(b) $|B| \ge \frac{2s+4}{3}$, then

$$|A + B| \ge |A| + s(|B| - 1). \tag{4}$$

We remark that the bound $|A| \ge \frac{1}{2}s(s-1)|B| + s$ is not in general sufficient to guarantee $|A + B| \ge |A| + s(|B| - 1)$, and thus the slight increase in the requirement for $|A|$ given by (a) is necessary. For instance, let $s = 34$, and let A' and B be geometrically similar right isosceles triangles whose equal length sides each cover 82 and 3 lattice points, respectively. Suppose A' lies in the positive upper plane with one its equal length sides along the horizontal axis. Let A be obtained from A' by deleting the 3 points in A' farthest away from the horizontal axis. Then $|B| = 6$, $|A| = 3400 = \frac{1}{2}s(s-1)|B| + s$, $h_1(A, B) = 80 > 34$, and $|A + B| = 3567 < 3570 = |A| + s(|B| - 1)$. As a second example, let $A = [0, a-1] \times [0, s+1]$ and $B = [0, b-1] \times \{0, 1\}$ be two rectangles in the integer lattice. We have $|A| = a(s + 2)$, $|B| = 2b$ and $|A + B| = (a + b - 1)(s + 3) =$

$|A|+s(|B|-1)+a-b(s-3)-3$. By taking $b=(s+3)/6$ and $a=(s(s-1)b+s+1)/(s+2)=(s^2+3)/6$ (with $s\equiv 3 \pmod 6$), we have $|A|=\frac{1}{2}s(s-1)|B|+s+1$, $|B|=(s+3)/3$ and $|A+B|<|A|+s(|B|-1)$. Furthermore, $h_1(A,B)\geq h_1(A,A)\geq \min\{s+2,\,(s^2+3)/6\}\geq s$ for $s\geq 9$.

We conclude the introduction with two special cases of Freiman's Theorem for which exact constants are known. The first is folklore [11] [15], while the second is a generalization by Lev and Smeliansky [10] of the Freiman $(3k-4)$–Theorem [5, Theorem 1.9] [11] [15].

Theorem B. *If A and B are finite and nonempty subsets of a torsion-free abelian group, then*

$$|A+B|\geq |A|+|B|-1, \tag{5}$$

with equality possible only when A and B are arithmetic progressions with common difference or when $\min\{|A|,\,|B|\}=1$.

Theorem C. *Let $A,\,B\subseteq\mathbb{Z}$ be finite nonempty subsets with $0=\min A=\min B$, $\max A\geq \max B$ and $\gcd(A)=1$. Let $\delta=1$ if $\max A=\max B$, and let $\delta=0$ otherwise. If*

$$|A+B|=|A|+|B|+r\leq |A|+2|B|-3-\delta,$$

then $\max A\leq |A|+r$.

2. Lower Bound Estimates via Compression

2.1. **Discrete Sets.** Let $X=(x_1,x_2,\ldots,x_d)$ be an ordered basis for $\mathbb{R}^d$, and let $X_i=\langle x_1,\ldots,x_i\rangle$ for $i=0,\ldots,d$. Let $A\subseteq\mathbb{R}^d$ be a finite subset. The linear compression of A with respect to $x_i\in X$, denoted $\mathbf{C}_i(A)=\mathbf{C}_{X,i}(A)$, is the set obtained by compressing and shifting A along each line $\mathbb{R}x_i+a$, where $a\in\mathbb{R}^d$, until the resulting set $\mathbf{C}_i(A)\cap(\mathbb{R}x_i+a)$ is an arithmetic progression with difference x_i whose first term is contained in the hyperplane $H=\langle x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_d\rangle$. More concretely, we define the set $\mathbf{C}_i(A)$ piecewise by its intersections with the lines $(\mathbb{R}x_i+a)$, $a\in\mathbb{R}^d$, by letting $\mathbf{C}_i(A)\cap(\mathbb{R}x_i+a)$ be the subset of $\mathbb{R}x_i+a$ satisfying

$$\phi_H(\mathbf{C}_i(A)\cap(\mathbb{R}x_i+a))=\{0,x_i,2x_i,\ldots,(r-1)x_i\},$$

where $r=|A\cap(\mathbb{R}x_i+a)|$ and the right hand side is considered empty if $r=0$. We let

$$\mathbf{C}_X(A)=\mathbf{C}_d(\mathbf{C}_{d-1}\ldots(\mathbf{C}_1(A)))$$

be the fully compressed subset obtained by iteratively compressing A in all d dimensions. Observe that

$$|\phi_{X_i}(\mathbf{C}_X(A))|=|\phi_{X_i}(A)|, \tag{6}$$

for $i=0,\ldots,d$.

Compression techniques in the study of sumsets have been used by various authors, including Freiman [5], Kleitman [9], Bollobás and Leader [2], and Green and Tao [8]. The reason for introducing the notion of compression is that it gives a useful lower bound for the sumset of an arbitrary pair of finite subsets $A,\,B\subseteq\mathbb{R}^d$. Namely, letting H be as above and letting C_t denote $C\cap(\mathbb{R}x_i+t)$

below, we have in view of Theorem B that

$$\begin{aligned} |A+B| &= \sum_{t\in H} |(A+B)_t| \\ &\geq \sum_{t\in H} \max\{|A_s+B_{t-s}|:\ A_s\neq\emptyset, B_{t-s}\neq\emptyset\} \\ &\geq \sum_{t\in H} \max\{|A_s|+|B_{t-s}|-1:\ A_s\neq\emptyset, B_{t-s}\neq\emptyset\} \\ &= |\mathbf{C}_i(A)+\mathbf{C}_i(B)|, \end{aligned} \tag{7}$$

and consequently (by iterative application of (7)),

$$|A+B| \geq |\mathbf{C}_X(A)+\mathbf{C}_X(B)|. \tag{8}$$

We now restrict our attention to the case $d=2$, which is the object of study for this paper. Let $m=|\phi_{X_1}(A)|$, $n=|\phi_{X_1}(B)|$, $A_i=\mathbf{C}_X(A)\cap(\mathbb{R}x_1+(i-1)x_2)$ and $B_i=\mathbf{C}_X(B)\cap(\mathbb{R}x_1+(i-1)x_2)$. Note that $|A_1|\geq|A_2|\geq\ldots\geq|A_m|$ and $|B_1|\geq|B_2|\geq\ldots\geq|B_n|$. If $|A_i|=a_i$ and $|B_j|=b_j$, then

$$|\mathbf{C}_X(A)+\mathbf{C}_X(B)| = \sum_{l=2}^{m+n}\max_i\{a_i+b_{l-i}\mid 1\leq i\leq m,\ 1\leq l-i\leq n\}-(m+n-1). \tag{9}$$

Consequently, the following lemma provides a lower bound for $|A+B|$ based upon the number of parallel lines that cover A and B, which will imply (3) in Theorem 1.3.

Lemma 2.1. *If $a_1,\ldots,a_m,b_1,\ldots,b_n\in\mathbb{R}$, then*

$$\frac{1}{m+n-1}\sum_{i=2}^{m+n}\max_j\{a_j+b_{i-j}:\ 1\leq j\leq m, 1\leq i-j\leq n\}\geq \frac{1}{m}\sum_{i=1}^{m}a_i+\frac{1}{n}\sum_{i=1}^{n}b_i. \tag{10}$$

Proof. The proof is by induction on $m+n$. The result clearly holds if either $m=1$ or $n=1$. Assume that $m,\ n\geq 2$. Let $a=(a_1,\ldots,a_m)$ and $b=(b_1,\ldots,b_n)$. For a vector $x=(x_1,x_2,\ldots,x_k)$, we denote by $\overline{x}=\frac{1}{k}\sum_{i=1}^{k}x_i$. Also, if $y=(y_1,\ldots,y_l)$, we denote by

$$u(x,y)=\sum_{i=2}^{k+l}\max_j\{x_j+y_{i-j}:\ 1\leq j\leq k, 1\leq i-j\leq l\}.$$

Thus we want to prove

$$u(a,b)\geq (m+n-1)(\bar{a}+\bar{b}).$$

Let $a'=(a_2,\ldots,a_m)$ and $b'=(b_2,\ldots,b_n)$. We may assume that $\bar{a}-\bar{a'}\leq\bar{b}-\bar{b'}$. We clearly have $u(a,b)\geq u(a',b)+a_1+b_1$. Thus by the induction hypothesis,

$$\begin{aligned} u(a,b) &\geq (m+n-2)(\bar{a'}+\bar{b})+a_1+b_1 \\ &= (m+n-2)(\bar{a'}+\bar{b})+m\bar{a}-(m-1)\bar{a'}+n\bar{b}-(n-1)\bar{b'} \\ &= (m+n-1)(\bar{a}+\bar{b})+(n-1)(\bar{a'}-\bar{a})+(n-1)(\bar{b}-\bar{b'}) \\ &\geq (m+n-1)(\bar{a}+\bar{b}), \end{aligned}$$

as claimed. ∎

Note that taking $a_i = \frac{1}{m}\sum_{k=1}^{m} a_k$ and $b_j = \frac{1}{n}\sum_{k=1}^{n} b_k$ for all i and j shows that equality can hold in (10). More generally, equality holds whenever $a_1, \ldots, a_m$ and $b_1, \ldots, b_n$ are arithmetic progressions of common difference. We now prove Theorem 1.3.

Proof. of Theorem 1.3. The bound in (3) follows from Lemma 2.1, (9), (8) and (6). Consider the bound given by (3) as a discrete function in the variable n. If $m = |A|$, then maximizing n will minimize (3). Otherwise, it is a routine discrete calculus minimization question to determine that $l = \lfloor\sqrt{\frac{1}{4} + \frac{(m-1)|B|}{|A|/m-1}} + \frac{1}{2}\rfloor$ is the value of n which minimizes (3), and that $l-1$ also minimizes the bound when $\sqrt{\frac{1}{4} + \frac{(m-1)|B|}{|A|/m-1}} + \frac{1}{2} \in \mathbb{Z}$. Rearranging the expression for l yields (iii). If $m \geq n$ and $|A| \leq |B| + m$, then $l \geq m \geq n$ follows, whence the minimum of (3) occurs instead at the boundary value $n = m$, yielding (i). If $|A| \geq |B| + m$, then (3) implies that

$$|A+B| \geq |A| + |B| + \frac{n-1}{m}(|B|+m) + \frac{m-1}{n}|B| - (m+n-1).$$

Considering the left hand side as a discrete function in n, it is another routine discrete calculus computation to determine $n = m$ minimizes the bound. This yields (ii). Note that when $|B| = |A| + m$ the bounds in (ii) and (i) are equal. Finally, considering the bound given by (3) as a continuous function in n, it follows that $n = \sqrt{\frac{(m-1)|B|}{|A|/m-1}}$ minimizes the bound in (3) when $|A| > m$. This yields (iv) except in the case $|A| = m$, in which case the trivial bound $|A+B| \geq |B|$ implies (iv) instead. □

2.2. **Measurable Sets.** Let μ_d be the Lebesgue measure on the space $\mathbb{R}^d$, $d \geq 1$, and let $\{x_1, \ldots, x_d\}$ be the d standard unit coordinate vectors for $\mathbb{R}^d$. Let $\phi_i : \mathbb{R}^2 \to \mathbb{R}$ denote the canonical projection onto the i-th coordinate, $i = 1, 2$. In this subsection, we briefly show how the results of the previous section are related to sumset volume estimates, such as the Brunn-Minkowski Theorem [15, 7]. In what follows, we make implicit use of the basic analytic theory regarding the Lebesgue measure (see e.g. [12]).

Theorem D (Brunn-Minkowski Theorem). *If A, $B \subseteq \mathbb{R}^d$ and $A + B$ are nonempty, measurable subsets, then*

$$\mu_d(A+B)^{1/d} \geq \mu_d(A)^{1/d} + \mu_d(B)^{1/d}. \tag{11}$$

In 1929, Bonneson gave the following generalization of the Brunn-Minkowksi Theorem (eq. (12) can be shown to imply (11)) [7, eq. (39)].

Theorem E (Bonneson's Generalization). *If A, $B \subseteq \mathbb{R}^d$ are compact and $H \subseteq \mathbb{R}^d$ is a hyperplane, then*

$$\mu_d(A+B) \geq \left(M^{1/(d-1)} + N^{1/(d-1)}\right)^{d-1} \left(\frac{\mu_d(A)}{M} + \frac{\mu_d(B)}{N}\right), \tag{12}$$

where $M = \sup\{\mu_{d-1}((x+H) \cap A) \mid x \in \mathbb{R}^d\}$ and $N = \sup\{\mu_{d-1}((x+H) \cap B) \mid x \in \mathbb{R}^d\}$.

Using either the compression techniques outlined in this section or related Steiner Symmetrization arguments, one can easily derive that the above bound (12) holds when $M = \mu_{d-1}(\varphi(A))$ and $N = \mu_{d-1}(\varphi(B))$, where $\varphi : \mathbb{R}^d \to H$ is any projection onto a hyperplane H. The goal of this section is to give a simple proof of the 2-dimensional case in this variation (though we do not include the details here, the ideas used in Lemma 3.1 can also be adapted to show this variation implies the original Bonneson version), illustrating how the discrete compression methods can be adapted to handle measurable sets.

For simplicity, we state the theorem below only for compact subsets; the compression techniques outlined in the proof of Theorem 2.2 can be made to work when A and B are merely measurable sets, but much extra care must then be taken to deal with issues of measurability, which might obscure the otherwise simple nature of the proof.

Theorem 2.2. *If A, $B \subseteq \mathbb{R}^2$ are compact, then*

$$\mu_2(A+B) \geq \left(\frac{\mu_2(A)}{\mu_1(\phi_1(A))} + \frac{\mu_2(B)}{\mu_1(\phi_1(B))}\right)(\mu_1(\phi_1(A)) + \mu_1(\phi_1(B))). \tag{13}$$

Proof. Note that the case when $\mu_1(\phi_1(A)) = 0$ is somewhat degenerate, being either trivial or meaningless, and so we assume $\mu_1(\phi_1(A)) > 0$ and likewise $\mu_1(\phi_1(B)) > 0$.

For a subset $X \subseteq \mathbb{R}^2$ and $i \in \{1, 2\}$, let $f_{X,i} : \phi_{3-i}(X) \to [0, \infty]$ be defined as $f_{X,i}(\phi_{3-i}(x)) = \mu_1(X \cap (\mathbb{R}x_i + x))$ if $X \cap (\mathbb{R}x_i + x)$ is measurable and otherwise $f_{X,i}(\phi_{3-i}(x)) = 0$. We define the linear compression $\mathbf{C}_i(X)$, for $i = 1, 2$, by its intersections with the lines $(\mathbb{R}x_i + a)$, $a \in \mathbb{R}^2$, by letting $\mathbf{C}_i(X) \cap (\mathbb{R}x_i + a)$ be the subset of $\mathbb{R}x_i + a$ defined by

$$\phi_i(\mathbf{C}_i(X) \cap (\mathbb{R}x_i + a)) = [0, f_{X,i}(\phi_{3-i}(a))],$$

if $X \cap (\mathbb{R}x_i + a)$ is nonempty, and letting $\mathbf{C}_i(X) \cap (\mathbb{R}x_i + a)$ be empty otherwise.

Since A is compact, we have $A = \bigcap_{j=1}^{\infty} S_j$, with $S_1 \supseteq S_2 \supseteq \ldots$ and each S_j a finite union of cubes (a cartesian product of closed intervals). Since A is compact and thus bounded, the lower continuity of μ_1 implies $\mathbf{C}_k(\bigcap_{j=1}^{\infty} S_j) = \bigcap_{j=1}^{\infty} \mathbf{C}_k(S_j)$, for $k = 1, 2$. Note that $\mathbf{C}_k(S_j)$, for $k = 1, 2$, is still a finite union of cubes. Consequently, $\mathbf{C}_k(A)$ is a compact set. We call $\mathbf{C}(A) = \mathbf{C}_1(\mathbf{C}_2(A))$ the *compression* of A. Clearly, we have

$$\mu_1(\phi_1(A)) = \mu_1(\phi_1(\mathbf{C}_2(A))) = \mu_1(\phi_1(\mathbf{C}(A)). \tag{14}$$

Likewise define $\mathbf{C}(B)$ and note that the corresponding equalities in (14) hold for $\mathbf{C}(B)$ as well.

Let $S_z = \mathbf{C}_2(A) \cap (\mathbb{R}x_1 + z)$ be an x_1–section. Observe that if $\phi_2(z) \leq \phi_2(z')$, then $S_{z'} \subseteq S_z$ and thus $\mu_1(S_{z'}) \leq \mu_1(S_z)$. Consequently, $\mathbf{C}(A)$ consists precisely in the area between the graph of the monotonic decreasing L^+–function $f_{\mathbf{C}_2(A),1} : [0, M] \to [0, \mu_1(\phi_1(A))]$ and the x_2-axis, where $M = \sup\{f_{A,2}(x) \mid x \in \phi_1(A)\}$. As both $\mu_1(\phi_1(A))$ and M are finite, $\mathbf{C}(A)$ is Riemann integrable. The same is true for $\mathbf{C}(B)$, from which it is then easily observed that their sumset $\mathbf{C}(A) + \mathbf{C}(B)$ also consists of the area between the graph of a monotonic decreasing L^+–function and the x_2-axis. Now by Fubini's Theorem, we have

$$\begin{aligned} \mu_2(\mathbf{C}(A)) &= \iint \chi_{\mathbf{C}_1(\mathbf{C}_2(A))} dx_1 dx_2 = \iint \chi_{\mathbf{C}_2(A)} dx_1 dx_2 \\ &= \iint \chi_{\mathbf{C}_2(A)} dx_2 dx_1 = \iint \chi_A dx_2 dx_1 = \mu_2(A), \end{aligned} \tag{15}$$

where χ_T denotes the characteristic function of the set T. Likewise,

$$\mu_2(\mathbf{C}(B)) = \mu_2(B). \tag{16}$$

Letting X_z denote in (17) below the x_2-section $(\mathbb{R}x_2 + z) \cap X$ of $X \subseteq \mathbb{R}^2$, we find that

$$\mu_1((A+B)_z) = \mu_1(\bigcup_{x+y=z}(A_x + B_y)) \geq \sup\{\mu_1(A_x + B_y) \mid x+y=z\}$$
$$\geq \sup\{\mu_1(A_x) + \mu_1(B_y) \mid x+y=z\} = \mu_1((\mathbf{C}_2(A) + \mathbf{C}_2(B))_z), \tag{17}$$

where the second inequality follows from the inequality $\mu_1(X+Y) \geq \mu_1(X) + \mu_1(Y)$ (which is the case $d=1$ in the Brunn-Minkowski Theorem). Using Fubini's Theorem and (17) (for the first inequality; the second one follows by an analogous argument), we infer

$$\begin{aligned}\mu_2(A+B) &= \iint \chi_{A+B} dx_2 dx_1 \geq \iint \chi_{\mathbf{C}_2(A)+\mathbf{C}_2(B)} dx_2 dx_1 \\ &= \iint \chi_{\mathbf{C}_2(A)+\mathbf{C}_2(B)} dx_1 dx_2 \geq \iint \chi_{\mathbf{C}_1(\mathbf{C}_2(A))+\mathbf{C}_1(\mathbf{C}_2(B))} dx_1 dx_2 \\ &= \mu_2(\mathbf{C}(A) + \mathbf{C}(B)). \end{aligned}\tag{18}$$

In view of (18), (15), (16) and (14), we see that it suffices to prove the theorem for $A = \mathbf{C}(A)$ and $B = \mathbf{C}(B)$. Since these are Riemann integrable, and thus can be approximated by rectangular strips of fixed height $\log_{2^n}(\mu_1(\phi_2(A)))$ and $\log_{2^n}(\mu_1(\phi_2(B)))$ when $n \to \infty$, it thus suffices to prove the theorem for unions of 2^n rectangular strips of equal height, $n \in \mathbb{Z}^+$. We proceed by induction. If $n = 1$, so that both A and B are themselves rectangles of width $\mu_1(\phi_1(A))$ and $\mu_1(\phi_1(B))$ and height $\frac{\mu_2(A)}{\mu_1(\phi_1(A))}$ and $\frac{\mu_2(B)}{\mu_1(\phi_1(B))}$, respectively, then (13) follows trivially. So we assume $n > 1$. Translate A and B so that the x_2-axis passes through the midpoints of $\phi_1(A)$ and $\phi_1(B)$, and let $A^+ \subseteq A$ and $B^+ \subseteq B$ be those points with nonnegative x_1-coordinate, and let $A^- \subseteq A$ and $B^- \subseteq B$ be those with non-positive x_1-coordinate. Observing that $\mu_2(A+B) \geq \mu_2(A^+ + B^+) + \mu_2(A^- + B^-)$ and applying the induction hypothesis to each of $A^+ + B^+$ and $A^- + B^-$ yields (13), completing the proof. □

3. An Inductive Argument

In this section, we prove the key lemma for an inductive argument analogous to one by Stanchescu [14, Lemma 2.2], which will be used in the proof of Theorem 1.1.

Recall that $h_1(A, B)$ denotes the minimal positive integer s such that there exist $2s$ (not necessarily distinct) parallel lines $\ell_1, \ldots, \ell_s, \ell'_1, \ldots, \ell'_s$ with $A \subseteq \bigcup_{i=1}^s \ell_i$ and $B \subseteq \bigcup_{i=1}^s \ell'_i$.

The inductive argument is collected in Lemma 3.1 below and roughly says that if $h_1(A, B)$ is large enough, we can remove a small number of points from A and B while decreasing substantially the cardinality of their sumset without increasing $||A| - |B||$ unduly.

Lemma 3.1. *Let $s \geq 3$ be an integer, and let A, $B \subseteq \mathbb{R}^2$ be finite subsets, with $|A| \geq |B| \geq s$, such that there are no s collinear points in either A or B. Then either:*

(a) $h_1(A, B) \leq 2s - 3$, or

(b) there exist a, $b \in \mathbb{R}^2$, a line ℓ, a nonempty subset $A_0 \subseteq A$ and a subset $B_0 \subseteq B$, such that $A_0 \subseteq a + \ell$, $B_0 \subseteq b + \ell$,

$$|B_0| \leq |A_0| \leq s - 1,$$

and

$$|A' + B'| \leq |A + B| - 2(|A_0| + |B_0|), \tag{19}$$

where $A' = A \setminus A_0$ *and* $B' = B \setminus B_0$.

Proof. Let Conv(X) denote the boundary of the convex hull of X. Note, since $|A| \geq |B| \geq s$ and since neither A nor B contains s collinear points, that both A and B must be 2-dimensional.

Throughout the proof we assume that (b) is false and proceed to show that (a) holds. The four claims below show that our assumption on (b) not holding leads to geometric structure for A and B (see Figure 1 for an illustration of this, and other soon to be established, information).

Claim 1. *If f and f' are two consecutive edges of Conv(A) incident at the vertex a_0, with a_1, $a_1' \in$ Conv$(A) \cap A$ the closest elements to a_0 in each of the edges f and f', respectively, then the sumset $A+B$ is contained in a translate of the lattice generated by the two vectors $a_1 - a_0$ and $a_1' - a_0$.*

Proof. We use an argument by Ruzsa [13]. Let b_0 be a vertex of Conv(B) such that $A^* = A \setminus \{a_0\}$ and $B^* = (B \setminus \{b_0\}) + (a_0 - b_0)$ are both contained in the same open half plane determined by some line through a_0. We may w.l.o.g. assume that $a_0 = b_0 = (0,0)$ and that both A^* and B^* are contained in the open half plane of points with positive first coordinate. Let $x \in A+B$, $x \neq (0,0)$, and consider all the expressions of x written as a sum of elements taken from $(A+B) \setminus \{(0,0)\}$. Since A and B are finite sets, and since all points in A^* and B^* have positive first coordinate, it follows that the number of summands in any such expression is bounded. Take one expression $x = x_1 + x_2 + \cdots + x_k$ with a maximum number of summands. If $x_i \in A^* + B^*$ for some i, then x_i can be split into two summands, one in A^* and one in B^*, contradicting the maximality of k. Therefore x can be written as a sum of elements in $C = (A+B) \setminus ((A^* + B^*) \cup \{(0,0)\})$.

Observe that if $|C| \geq 3$, then (b) holds with $A_0 = \{(0,0)\} = B_0$. Hence $|C| \leq 2$ and all elements in $A+B$ are contained in the lattice generated by the two elements of C. Let e and e' be the two edges incident with b_0. Note we may assume the convex hull of the two rays parallel to e and e' with base point $b_0 = (0,0)$ is contained in the convex hull of two rays parallel to f and f' with base point $a_0 = (0,0)$, since otherwise by removing a_0 from A we lose all the points in either $a_0 + (B \cap e)$ or $a_0 + (B \cap e')$, yielding (b). However, in this case, it is easily seen that $\{a_1, a_1'\} \subseteq C$, whence $|C| = 2$ implies $C = \{a_1, a_1'\}$, completing the claim. □

Note that Claim 1 implies that A and B are also contained in a translate of the lattice generated by $a_1 - a_0$ and $a_1' - a_0$, though the particular translate may vary from A to B to $A+B$.

Claim 2. *For each side e of Conv(B), there is a side f of Conv(A), parallel to e, such that both $A - f + e$ and B are contained in the same half plane defined by e. Moreover, $|B \cap e| \leq |A \cap f|$.*

Proof. Let ℓ be the line parallel to e that intersects A and for which $A - \ell + e$ and B are both contained in the same half plane defined by e. Let $f = \ell \cap \text{Conv}(A)$ and let $A_f = A \cap \ell$. In view of Theorem B, we see that, by removing the elements of A_f, we lose $|A_f + B_e| \geq |A_f| + |B_e| - 1$ elements from $A+B$, where $B_e = B \cap e$. Since (b) does not hold, it follows that $|A_f| + |B_e| - 1 < 2|A_f|$, whence $2 \leq |B_e| \leq |A_f|$. In particular, f is an edge of the Conv(A). □

Let e and e' be two consecutive edges of Conv(B), and let f and f' be the corresponding parallel edges in Conv(A) as given by Claim 2. Denote the elements in $B_e := B \cap e$ by $b_0, b_1, \ldots, b_t$, ordered as they occur in the edge e, and the ones in $A_f := A \cap f$ by $a_0, a_1, \ldots, a_r$, ordered in the same direction as those of B_e. Likewise define $b_0' = b_0, b_1', \ldots, b_{t'}'$ and $a_0', a_1', \ldots, a_{r'}'$ for the points in

$B_{e'} := B \cap e'$ and $A_{f'} := A \cap f'$. Note $a_0 = a'_0$ need not hold, though as we will soon see (Claim 4), this cannot fail by much.

Claim 3. *With the notation above, $b_0 - b_1 = a_0 - a_1$.*

Proof. Let $f'' \neq f$ be the edge adjacent to a_0 and let $a'' \neq a_1$ be the element of $\mathrm{Conv}(A) \cap A$ adjacent to a_0. If the claim is false, then, by removing a_0 from A_f and b_0 from B_e, we lose from $A + B$ the distinct elements $a_0 + b_0$, $a_0 + b_1$, $a_1 + b_0$ and either $b_0 + a''$ or $a_0 + b'_1$, yielding (b). □

Claim 4. *With the notation above, either: (i) f and f' are also consecutive, or (ii) they are separated by a single edge g of Conv(A) and $A \cap g$ contains exactly two points.*

Proof. Traverse the convex hull of A, beginning at a_0 and in the direction not given by f. Let $a_0, c_1, c_2, \ldots, c_k, a'_0$ be the sequence of points on $\mathrm{Conv}(A)$ encountered until the first point a'_0 of f' is reached. If the claim is false, then $k \geq 1$. Hence, by removing a_0 from A and b_0 from B, we lose from $A + B$ the elements $a_0 + b_0$, $b_0 + a_1$, $b_0 + c_i$ for $i = 1, \ldots, k$, and $b_0 + a'_0$, yielding (b). □

By an appropriate affine transformation, we may assume that $b_0 = (0, 0)$, $b_1 = (1, 0)$ and $b'_1 = (0, 1)$ and that both A and B are contained in the positive first quadrant. We denote by $\pi_1 : \mathbb{R}^2 \to \mathbb{R}$ the projection onto the first coordinate. Let $A_i = A \cap \{y = i\}$ and let $B_i = B \cap \{y = i\}$. We have

$$|A_0 + B_0| \;<\; 2|A_0|, \tag{20}$$

$$|A_0 + B_0| + |(A_0 + B_1) \cup (A_1 + B_0)| \;<\; 2(|A_0| + |B_0|), \tag{21}$$

since otherwise deletion of A_0 or $A_0 \cup B_0$ yields (b) (in view of Claim 2).

It follows from (20) and Theorem B that

$$|B_0| \leq |A_0| \text{ and } \pi_1(b_t) \leq \pi_1(a_r) - \pi_1(a_0). \tag{22}$$

Moreover, both inequalities are strict unless (possibly) A_0 is an arithmetic progression: the former follows in view of Theorem B; equality in the latter would imply $A_0 + B_0 = (A_0 + b_0) \cup (A_0 + b_t) = A_0 \cup (A_0 + B_t)$ while Claims 1 and 3 imply $-a_0 + A_0$ is contained in the integer lattice, and then $A_0 + b_1 = A_0 + (1, 0) \subseteq A_0 \cup (A_0 + b_t)$, from which one inductively finds that $(x, 0) \in A_0$ for $\pi_1(a_0) \leq x \leq \pi_1(a_0) + \pi_1(b_t) = \pi_1(a_r)$, yielding the latter.

We proceed in two cases according to Claim 4 (i) and (ii).

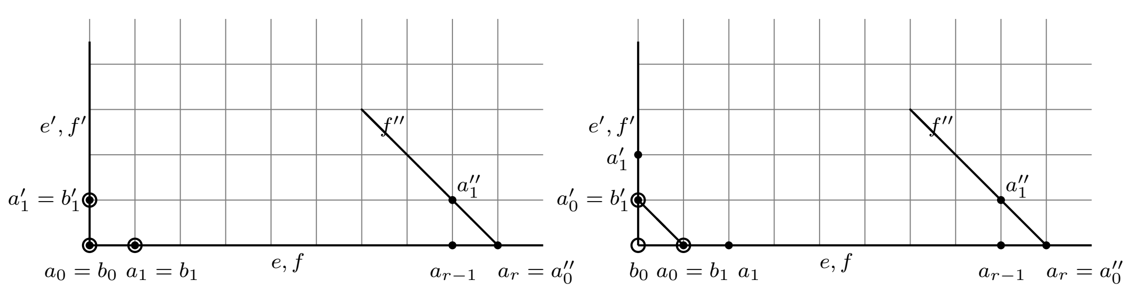

FIGURE 1. A Schematic Illustration of Cases 1 and 2.

Case 1: Claim 4(i) holds for the pair f and f'. In this case, $a_0 = a'_0$ and w.l.o.g. $a_0 = b_0 = (0,0)$. By Claim 3, it follows that

$$b_0 - b_1 = a_0 - a_1 \text{ and } b_0 - b'_1 = a_0 - a'_1.$$

Thus $a_1 = b_1 = (1,0)$ and $a'_1 = b'_1 = (0,1)$. By Claim 1, it follows, in view of $0 \in A \cap B$, that A, B, and $A+B$ are contained in the integer lattice. Moreover, in view of Claim 3 and Claim 1 applied to a_r, it follows that

$$b_t - b_{t-1} = a_r - a_{r-1} = a_1 - a_0 = (1,0) \text{ and } a''_1 \in A_1, \tag{23}$$

where a''_1 is the next element of $A \cap \mathrm{Conv}(A)$ on the edge f'' incident to f with endpoint a_r. Figure 1 shows a picture of the situation.

Let $e'' \neq e$ denote the edge of $\mathrm{Conv}(B)$ incident to b_t. As a consequence of Claim 2, the angle between e and e'' is at most the angle between f and f''. Consequently, it follows in view of (22) that $A \cup B$ is contained in the region defined by the lines $y = 0$, $x = 0$ and the line defined by f''.

We proceed to verify that

$$\pi_1(a''_1) \leq \pi_1(a_r) + 1. \tag{24}$$

Suppose (24) is false. Then $\pi_1(a''_1) \geq \pi_1(a_r) + 1$, along with Theorem B and (21), implies that $|A_0 + B_0| = |A_0| + |B_0| - 1$ and $|(A_0 + B_1) \cup (A_1 + B_0)| = |A_0 + b'_1| + |a''_1 + B_0|$. It follows from the former and Theorem B that A_0 and B_0 are arithmetic progressions with the same difference (which must then be $(1,0)$), and by the second and (24) not holding that the point $a_r + (1,1)$ is not in $A+B$. Combining these two facts, we see that, by removing a_r from A and b_t from B, we lose the elements $a_r + b_t$, $a_r + b_t - (1,0)$, $a''_1 + b_t$ and $a_r + (0,1)$ from the sumset, yielding (b), a contradiction. So (24) is established.

The above argument also shows that, if equality holds in (24), then A_0 is an arithmetic progression. Suppose that this is the case. We may apply the analogous arguments to the set $A \cap \{x = 0\}$. If this set is also an arithmetic progression, then $A \cup B$ is contained in the region defined by the lines $y = 0$, $x = 0$, $y = x - a_r$ and $y = x + a'_{r'}$. Then, since each line, including f and f', contains at most $s-1$ points of A, it follows that $A \cup B$ can be covered by the $2s-3$ lines with slope one passing through the points of A lying on either coordinate axis, yielding (a) as claimed.

Therefore we may assume without loss of generality that A_0 is not an arithmetic progression. Thus the inequalities in (22) and (24) are strict. If $\pi_1(a_r) \leq 2s-4$, then $A \cup B$ is covered (in view of the second paragraph of Case 1 and the strict inequality in (24)) by the $2s-3$ vertical lines $x = i$, for $0 \leq i \leq 2s-4$, and (a) holds. Therefore we may assume that $\pi_1(a_r) \geq 2s-3$, whence $\pi_1(a_r) \geq |A_0| + |B_0| - 1$.

Consequently, by Theorem C applied to A_0 and B_0 with $\delta = 0$ (since the second inequality in (22) is strict), we get

$$|A_0 + B_0| \geq |A_0| + 2|B_0| - 2, \tag{25}$$

which, combined with (21), Theorem B and the fact that A_0 is not an arithmetic progression, yields

$$|B_1| = 1 \text{ and } |(B_0 + A_1) \setminus (A_0 + B_1)| \leq 1. \tag{26}$$

As a result, $B_1 = \{b'_1\}$ and

$$\pi_1(a''_1) \leq \pi_1(a_r) - \pi_1(b_{t-1}), \tag{27}$$

with equality possible only if $a''_1 + b_t$ is a unique expression element in $A+B$.

Let b be the intersection of e'' with the line $y = 1$. Recalling that the angle between e and e'' is at most the angle between f and f'' (see the second paragraph of Case 1), we find that (27) and (23) yield

$$\pi_1(b_t) - \pi_1(b) \geq \pi_1(a_r) - \pi_1(a_1'') \geq \pi_1(b_{t-1}) = \pi_1(b_t) - 1. \tag{28}$$

Consequently, $\pi_1(b) \leq 1$.

If $\pi_1(b) = 0$, then it follows, by the strict inequality in (22), that $|B| = |B_0| + 1 \leq |A_0| \leq s - 1$, a contradiction. Therefore $\pi_1(b) > 0$, which is only possible if equality holds in (27), else the estimate from (28) improves by 1. Thus $a_1'' + b_t$ is a unique expression element, so that if e'' and f'' were parallel, then by removing a_r from A and b_t from B we would lose the elements $a_r + b_t$, $a_r + b_{t-1} = a_{r-1} + b_t$, $a_1'' + b_t$ and $a_r + b_1''$, where b_1'' is the next element from B on the edge e'' after b_t, yielding (b). So we may assume e'' and f'' are not parallel, whence the estimate in (28) becomes strict, yielding $0 < \phi_1(b) < 1$. Thus it follows that $|B_0| = 2$, since otherwise we again get $|B| \leq |B_0| + 1 \leq |A_0| \leq s - 1$, a contradiction. Hence, since $|A_0 + B_0| \geq |A_0| + |B_0| = |A_0| + 2$ (by Theorem B and the fact that A_0 is not an arithmetic progression) and since $|(A_0 \setminus a_r) + (B_0 \setminus b_t)| = |A_0 \setminus a_r|$ (in view of $|B_0 \setminus b_t| = 1$), it follows that removing a_r from A_0 and b_t from B_0 deletes at least three points from $A + B$ contained in $A_0 + B_0$ as well as the unique expression element $a_1'' + b_t$ in $A + B$, yielding (b). This completes the proof of Case 1.

Case 2: Claim 4(ii) holds for the pair f and f'. This case is slightly simpler than Case A, and we use very similar arguments. Recall that g is the line defined by a_0 and a_0', that $b_0 = (0, 0)$, $b_1 = (1, 0)$, $b_1' = (0, 1)$, and that both A and B are contained in the positive first quadrant. We may also assume f is contained in the horizontal axis and f' is contained in the vertical axis; furthermore, by the same arguments used to establish (23), we have $a_0 = (1/d, 0)$, $a_1 = (1/d+1, 0)$, $a_0' = (0, 1/d')$ and $a_1' = (0, 1/d' + 1)$, for some d, $d' \in \mathbb{R}^+$, and $a_1'' \in A_{1/d'}$, where f'' and a_1'' are as there were defined in Case 1. From Claim 1, applied both to f and g and to g and f', we conclude that d, $d' \in \mathbb{Z}^+$ and that A is contained both in the lattice $(1/d, 0) + \langle(1, 0), (1/d, -1/d')\rangle$ and the lattice $(1/d, 0) + \langle(0, 1), (1/d, -1/d')\rangle$. Thus

$$(-1/d, 1/d' + 1) = a_1' - (1/d, 0) \in A - (1/d, 0) \subseteq \langle(1, 0), (1/d, -1/d')\rangle$$

implies $d|d'$, while

$$(1, 0) = a_1 - (1/d, 0) \in A - (1/d, 0) \subseteq \langle(0, 1), (1/d, -1/d')\rangle$$

implies $d'|d$. Hence $d = d'$ and $A + B$ is contained within the lattice $(1/d, 0) + \langle(1, 0), (1/d, -1/d)\rangle$.

Since $A + B$ is contained within the lattice $(1/d, 0) + \langle(1, 0), (1/d, -1/d)\rangle$, by removing b_0 from B and a_0 and a_0' from A, we lose all the elements of $A + B$ contained within the two lines with slope -1 passing through a_0 and a_1, i.e., all the elements from

$$(b_0 + \{a_0, a_0'\}) \cup (b_0 + \{a_1, a_1'\}) \cup (\{a_0, a_0'\} + \{b_1, b_1'\}) =$$
$$\{(0, 1/d), (1/d, 0), (1 + 1/d, 0), (0, 1 + 1/d), (1, 1/d), (1/d, 1)\}.$$

If $d > 1$, then the above 6 elements are distinct, and (b) follows. Therefore we may assume $d = 1$. As a result, $b_0 = (0, 0)$, $a_0 = b_1 = (1, 0)$, $a_1 = (2, 0)$, $a_0' = b_1' = (0, 1)$, $a_1' = (0, 2)$, and A, B and $A + B$ are contained in the integer lattice. The right side of Figure 1 illustrates this case.

As a result, if $\pi_1(a_1'') \geq \pi_1(a_r) + 1$, then (21) and Theorem B would imply

$$|A_0| + |B_0| \geq |(A_0 + B_1) \cup (A_1 + B_0)| \geq |b_1' + A_0| + |a_1'' + B_0| + |\{a_0' + b_0\}| = |A_0| + |B_0| + 1,$$

a contradiction. Therefore

$$\pi_1(a_1'') \leq \pi_1(a_r). \tag{29}$$

As in Case 1, we have from (22) and Claim 2 that $A \cup B$ is contained in the region defined by the lines $x = 0$, $y = 0$ and the line defined by f''. Thus, if $\pi_1(a_r) \leq 2s-4$, then it follows in view of (29) that $A \cup B$ is contained in the $2s-3$ parallel lines $x = i$, $0 \leq i \leq 2s-4$, yielding (a). Therefore we may assume $\pi_1(a_r) \geq 2s-3$. Hence, since $2s-3 \geq s > |A_0|$ for $s \geq 3$, it follows that A_0 is not in arithmetic progression, whence the two inequalities in (22) are strict. By the same arguments used in Case 1, the relation (25) holds—if $\pi_1(b_t) = \pi_1(a_r) - 1$, so that the diameters of A_0 and B_0 are equal, then Theorem C with $\delta = 1$ should instead be applied with the roles of A_0 and B_0 reversed and noting that $|A_0| > |B_0|$ holds in view of the strict inequality in (22).

Now (21) and (25) imply $|(A_0 + B_1) \cup (A_1 + B_0)| \leq |A_0| + 1$. Consequently, since $\{a_1' + b_0\} \cup (b_1' + A_0) \subseteq (A_0 + B_1) \cup (A_1 + B_0)$ with $|\{a_1' + b_0\} \cup (b_1' + A_0)| = |A_0| + 1$, we conclude that

$$(A_0 + B_1) \cup (A_1 + B_0) = \{a_1' + b_0\} \cup (b_1' + A_0),$$

whence

$$\pi_1(a_1'') + \pi_1(b_t) \leq \pi_1(a_r). \tag{30}$$

Let b be the intersection of the edge e'' with the line $y = 1$, where e'' is as defined in Case 1. As we have seen before, Claim 2 implies that the angle between e and e'' is at most the angle between f and f''. Thus (30) implies

$$\pi_1(b_t) - \pi_1(b) \geq \pi_1(a_r) - \pi_1(a_1'') \geq \pi_1(b_t),$$

implying $\pi_1(b) = 0$. Hence $|B| \leq |B_0 \cup \{b_1'\}| \leq |A_0| \leq s-1$ (in view of of the strict inequality $|B_0| < |A_0|$ from (22)), a contradiction. This completes the proof. □

4. A Lemma for Small cases

The following lemma will allow us to improve, in a very particular case, the bound given in Theorem 1.3 (i) by one, which will be a crucial improvement needed in the proof of Theorem 1.1 for the extremal case $|A| + |B| \leq 4s^2 - 5s - 1$.

Lemma 4.1. *Let $X = (x_1, x_2)$ be a basis for $\mathbb{R}^2$, let $s \geq 2$ be an integer, let $A, B \subseteq \mathbb{R}^2$ be finite, nonempty subsets with $||A| - |B|| \leq s$ and $4s^2 - 6s + 3 \leq |A| + |B| \leq 4s^2 - 5s - 1$. Suppose that $|\phi_{X_1}(A)| \leq |\phi_{X_1}(B)| = 2s-2$, where $X_1 = \mathbb{R}x_1$, and that some line parallel to $\mathbb{R}x_1$ intersects A in at least $2s-2$ points. Then*

$$|A + B| \geq 2|A| + 2|B| - 6s + 7. \tag{31}$$

Proof. We may w.l.o.g. assume $\mathbf{C}_X(A) = A$ and $\mathbf{C}_X(B) = B$. Let $m = |\phi_{X_1}(A)|$ and $n = |\phi_{X_1}(B)|$. Let $A_i = A \cap (\mathbb{Z}x_1 + (i-1)x_2)$, $B_j = (\mathbb{Z}x_1 + (j-1)x_2)$, $|A_i| = a_i$ and $|B_i| = b_i$, for $i = 1, \ldots, m$ and $j = 1, \ldots, n$. By hypothesis, we have $a_1 \geq 2s-2$ and $m \leq n = 2s-2$. Assume by contradiction

$$|A + B| \leq 2|A| + 2|B| - 6s + 6. \tag{32}$$

Suppose $m < n = 2s-2$. Then, since $||A| - |B|| \leq s \leq 2s-2$, from the proof of Theorem 1.3 we know that (3) is minimized for the boundary value $m = n-1$. Hence

$$|A + B| \geq |A| + |B| - (n + n - 1 - 1) + \frac{n-1}{n-1}|A| + \frac{n-2}{n}|B| = 2|A| + 2|B| - 4s + 6 - \frac{2}{2s-2}|B|,$$

which together with (32) implies $|B| \geq s(2s-2)$. Consequently, $|A|+|B| \geq 2|B|-s \geq 2s(2s-2)-s = 4s^2-5s$, contradicting our hypotheses. So we may assume $m = n = 2s-2$.

For $j = 1, \ldots, s-1$ consider the following lower estimations of $|A+B|$:

$$\begin{aligned} |A+B| \geq g_j(A,B) &= \sum_{i=1}^{j-1}(a_i+b_1-1) + \sum_{i=1}^{2s-2-j}(a_j+b_i-1) \\ &\quad + \sum_{i=j}^{2s-2}(a_i+b_{2s-j-1}-1) + \sum_{i=2s-j}^{2s-2}(a_{2s-2}+b_i-1) \\ &= |A|+|B| \\ &\quad +(j-1)(a_{2s-2}+b_1)+(2s-2-j)(a_j+b_{2s-j-1})-(4s-5), \end{aligned} \tag{33}$$

$$\begin{aligned} |A+B| \geq g(A,B) &= \sum_{i=1}^{2s-3}(a_i+b_i+a_{i+1}+b_i-2)+(a_{2s-2}+b_{2s-2}-1) \\ &= 2|A|+2|B|-a_1-b_{2s-2}-(4s-5). \end{aligned} \tag{34}$$

By using (32) and $g_1(A,B)$, it follows that $|A|+|B| \geq (2s-3)(a_1+b_{2s-2})+2s-1$. Thus $|A|+|B| \leq 4s^2-5s-1$ implies that $a_1+b_{2s-2} \leq 2s-1$. However, by using $g(A,B)$ instead, it follows that $a_1+b_{2s-2} \geq 2s-1$. Consequently,

$$a_1+b_{2s-2} = 2s-1. \tag{35}$$

Repeating these arguments with $g_1(B,A)$ and $g(B,A)$, we likewise conclude

$$b_1+a_{2s-2} = 2s-1. \tag{36}$$

If $a_j+b_{2s-j-1} \geq 2s$, then, in view of (36), (32) and (33), it follows that

$$|A|+|B| \geq j(2s-1)+(2s-2-j)(2s) = 4s^2-4s-j \geq 4s^2-5s+1,$$

contradicting that $|A|+|B| \leq 4s^2-5s-1$. Therefore we may assume

$$a_j+b_{2s-j-1} \leq 2s-1, \tag{37}$$

for all $j = 1, \ldots, s-1$. Repeating this argument with $g_j(B,A)$ and $g(B,A)$ instead, we likewise conclude

$$b_j+a_{2s-j-1} \leq 2s-1, \tag{38}$$

for all $j = 1, \ldots, s-1$. However, summing (37) and (38) over $j = 1, \ldots, s-1$ yields

$$|A|+|B| \leq 2(s-1)(2s-1) = 4s^2-6s+2,$$

contradicting our hypotheses, and completing the proof. □

5. Proof of Theorem 1.1

The proof of Theorem 1.1 is by induction on s and it uses the following version, which is essentially equivalent to Theorem 1.1.

Theorem 5.1. *Let $s \geq 3$ be an integer, and let A, $B \subseteq \mathbb{R}^2$ be finite subsets such that there are no s collinear points in either A or B.*

(i) *If $||A| - |B|| \leq s$ and $|A| + |B| \geq (s-1)(4s-6)+1$, then*
$$|A+B| \geq 2|A| + 2|B| - 6s + 7.$$

(ii) *If $|A| \geq |B| + s$ and $|B| \geq \frac{1}{2}(s-1)(4s-7)$, then*
$$|A+B| \geq |A| + 3|B| - 5s + 7.$$

We first show that part (ii), in both Theorems 5.1 and 1.1, is a very simple consequence of the corresponding part (i).

Lemma 5.2. *Let $s \geq 2$ be a positive integer. (a) If $s \geq 3$ and Theorem 5.1(i) holds for s, then Theorem 5.1(ii) holds for s. (b) If Theorem 1.1(i) holds for s, then Theorem 1.1(ii) holds for s.*

Proof. We first prove (a). Observe that $|(A \setminus x) + B| < |A+B|$ for any vertex x in the convex hull of A. Thus, by iteratively deleting vertices from the convex hull, we can obtain a subset $A' \subseteq A$ with $|A'| = |B| + s$ and
$$|A' + B| \leq |A+B| - |A \setminus A'|. \tag{39}$$

Since $|B| \geq \frac{1}{2}(s-1)(4s-7)$, it follows that $|A'| + |B| = 2|B| + s \geq (s-1)(4s-6)+1$, whence we can apply Theorem 5.1(i) to $A' + B$. Thus $|A'+B| \geq 2|A'| + 2|B| - 6s + 7 = |A'| + 3|B| - 5s + 7$, whence the theorem follows in view of (39).

Next we prove (b). Suppose by contradiction that $h_1(A,B) \geq s$. As in the previous part, observe that $|(A \setminus x) + B| < |A+B|$ for any vertex x in the convex hull of A. Thus by iteratively deleting vertices from the convex hull we can obtain a sequence of subsets $A_0 = A \supseteq A_1 \supseteq \ldots \supseteq A_{|A|-|B|-s} = A_k$, with $|A_i| = |A| - i$ and
$$|A_i + B| \leq |A+B| - |A \setminus A_i| < |A_i| + 3|B| - s - \frac{2|B|}{s}, \tag{40}$$
where the last inequality follows from (2).

Since $|A_i| = |A_{i-1}| - 1$ and $A_i \subseteq A_{i-1}$, it follows that $h_1(A_i, B) \geq h_1(A_{i-1}, B) - 1$ for all i. Consequently, if $h(A_k, B) < s$, then it would follow in view of $h(A,B) \geq s$ that $h(A_j, B) = s$ for some j, whence Theorem 1.3(i)(ii) would contradict (40) for $i = j$ (note the bound in Theorem 1.3(i) implies that in Theorem 1.3(ii) in view of $|A_j| \geq |A_k| = |B| + s$). Therefore we may assume $h(A_k, B) \geq s$.

Since $|B| \geq 2s^2 - \frac{7}{2}s + \frac{3}{2}$, it follow that $|A_k| + |B| = 2|B| + s \geq 4s^2 - 6s + 3$. Hence we can apply Theorem 1.1(i) to $A_k + B$, whence $h_1(A_k, B) \geq s$ implies
$$|A_k + B| \geq 2|A_k| + 2|B| - 2s + 1 - \frac{|A_k| + |B|}{s} = |A_k| + 3|B| - s - \frac{2|B|}{s},$$
contradicting (40) for $i = k$, and completing the proof. $\square$

We will prove Theorems 5.1 and 1.1 simultaneously using an inductive argument on s: the case $s-1$ of Theorem 1.1 will be used to prove the case s of Theorem 5.1, while the case s of Theorem 5.1 will be used to prove the case s of Theorem 1.1 (except for the case $s = 2$, where a trivial argument will be used instead). Thus both Theorems 5.1 and 1.1 follow immediately from the following two lemmas. This also shows that Theorem 5.1 and Theorem 1.1 are in some sense equivalent statements.

Lemma 5.3. *Let $s \geq 3$ be a positive integer. Suppose that the statement in Theorem 1.1 holds for $s-1$. Then Theorem 5.1 holds for s.*

Proof. In view of Lemma 5.2, it suffices to show part (i) holds, so suppose on the contrary that Theorem 5.1(i) is false for s. Let $A,\, B \subseteq \mathbb{R}^2$ be a counterexample with $|A|+|B|$ minimum. Thus $||A|-|B|| \leq s$, $|A|+|B| \geq (s-1)(4s-6)+1$ and

$$|A+B| < 2|A|+2|B|-6s+7. \tag{41}$$

We may assume $|A| \geq |B|$.

Since neither A nor B contains s collinear points, and since $|A|+|B| \geq (s-1)(4s-6)+1$, it follows from the pigeonhole principle that $h_1(A,B) > 2s-3$. By Lemma 3.1 (in view of (41)), there is a nonempty subset $A_0 \subseteq A$ and $B_0 \subseteq B$ with $|B_0| \leq |A_0| \leq s-1$ and

$$|A'+B'| \leq |A+B| - 2(|A_0|+|B_0|) < 2|A'|+2|B'|-6s+7, \tag{42}$$

where $A' = A \setminus A_0$ and $B' = B \setminus B_0$. Furthermore, $||A'|-|B'|| = ||A|-|B|-(|A_0|-|B_0|)| \leq s$. Therefore, by the minimality of $|A|+|B|$, we have

$$|A'|+|B'| \leq (s-1)(4s-6).$$

As a result,

$$|A|+|B| \leq |A'|+(s-1)+|B'|+(s-1) \leq (s-1)(4s-6)+2(s-1) = 4(s-1)^2. \tag{43}$$

If $|A| < |B|+s$, then, since $h_1(A,B) > 2s-3 \geq s-1$ and since

$$|A|+|B| \geq (s-1)(4s-6)+1 > (s-1)(4s-9)+3 = 4(s-1)^2-5(s-1)+3,$$

it follows, in view of (41) and the case $s-1$ of Theorem 1.1(i), that

$$2|A|+2|B|-2(s-1)+1-\frac{|A|+|B|}{s-1} \leq |A+B| \leq 2|A|+2|B|-6s+6.$$

Hence $|A|+|B| \geq (4s-3)(s-1) > 4(s-1)^2$, contradicting (43). On the other hand, if $|A| = |B|+s$, then, since $h_1(A,B) > 2s-3 \geq s-1$ and since

$$\begin{aligned} 2|B|+s &= |A|+|B| \geq (s-1)(4s-6)+1 = 4s^2-10s+7 \\ &\geq 4s^2-14s+14 = 4(s-1)^2-7(s-1)+3+s, \end{aligned}$$

it follows, in view of (41) and the case $s-1$ of Theorem 1.1(ii), that

$$2|A|+2|B|-2s+1-\frac{|A|+|B|-s}{s-1} = |A|+3|B|-(s-1)-\frac{2|B|}{s-1} \leq |A+B| \leq 2|A|+2|B|-6s+6.$$

Hence $|A|+|B| \geq (4s-5)(s-1)+s = 4s^2-8s+5 > 4(s-1)^2$, contradicting (43), and completing the proof. □

Lemma 5.4. *Let $s \geq 2$ be a positive integer. If $s \geq 3$, suppose that the statement of Theorem 5.1 holds for s. Then Theorem 1.1 holds for s.*

Proof. In view of Lemma 5.2, it suffices to show part (i) holds. Let $A,\, B \subseteq \mathbb{R}^2$ verify the hypothesis of Theorem 1.1(i) for s, and assume by contradiction that $h_1(A,B)\geq s$.

Suppose neither A nor B contain s collinear points. Thus $|A|+|B|\geq 3$ implies that $s\geq 3$. Hence, in view of Theorem 5.1(i) and (1), it follows that

$$2|A|+2|B|-6s+7\leq |A+B|<2|A|+2|B|-2s+1-\frac{|A|+|B|}{s}.$$

Thus $|A|+|B|<4s^2-6s$, contradicting that $|A|+|B|\geq 4s^2-6s+3$. So we may assume w.l.o.g. that A contains at least s collinear points on the line $\mathbb{Z}x_1+a_1$. Let $X=(x_1,x_2)$ be an ordered basis for $\mathbb{R}^2$

Since $h_1(A,B)\geq s$, so that $\max\{|\phi_{X_1}(A)|,\, |\phi_{X_1}(B)|\}\geq s$, it follows in view of (6) that

$$\max\{|\phi_{X_1}(\mathbf{C}_X(A))|,\, |\phi_{X_1}(\mathbf{C}_X(B))|\}\geq s.$$

Hence, since A contains s collinear points on a line parallel to $\mathbb{Z}x_1$, it follows that $h_1(\mathbf{C}_X(A),\mathbf{C}_X(B))\geq s$. Consequently, we conclude from (8) that it suffices to prove the theorem on compressed sets, and w.l.o.g. we assume $A=\mathbf{C}_X(A)$ and $B=\mathbf{C}_X(B)$. Let $|\phi_{X_1}(A)|=m$ and $|\phi_{X_1}(B)|=n$. Let $A_i=A\cap(\mathbb{Z}x_1+(i-1)x_2)$, $1\leq i\leq m$, and $B_i=B\cap(\mathbb{Z}x_1+(i-1)x_2)$, $1\leq i\leq n$. Note, since both A and B are compressed, that $|A_1|\geq|A_2|\geq\ldots\geq|A_m|$ and $|B_1|\geq|B_2|\geq\ldots\geq|B_n|$. Since A contains s collinear points along a line parallel to $\mathbb{Z}x_1$, it follows that $|A_1|\geq s$.

By our assumption to the contrary, we have $\max\{m,\, n\}\geq s$. Thus it follows, from Theorem 1.3(i) (applied with the line $\mathbb{Z}x_1$) and (1), that

$$\max\{m,\, n\}\geq\left\lfloor\frac{|A|+|B|}{2s}\right\rfloor+1. \tag{44}$$

Since $\max\{|A_1|,\, |B_1|\}\geq s$, it follows, from Theorem 1.3(i) (applied with the line $\mathbb{Z}x_2$) and (1), that

$$\max\{|A_1|,\, |B_1|\}\geq\left\lfloor\frac{|A|+|B|}{2s}\right\rfloor+1. \tag{45}$$

Let $k=|A|+|B|$, and let

$$x=\left\lfloor\frac{|A|+|B|}{2s}\right\rfloor+1=\frac{|A|+|B|-\alpha}{2s}+1,$$

so that $k=|A|+|B|\equiv\alpha \mod 2s$, with $0\leq\alpha\leq 2s-1$. With this notation, (1) yields

$$|A+B|\leq 2(k-s-x+1)-\delta, \tag{46}$$

where $\delta=0$ if $\alpha<s$ and otherwise $\delta=1$.

We proceed to show that

$$|A+B|<k-(2x-2)+\frac{x-2}{x}|A|+|B|. \tag{47}$$

Suppose (47) does not hold. In this case, if $\delta=0$, then $\alpha\leq s-1$ whence from (46) we conclude that

$$|A|\geq sx=s(\frac{|A|+|B|-\alpha}{2s}+1)\geq s(\frac{2|A|-s-\alpha}{2s}+1)\geq s(\frac{2|A|-2s+1}{2s}+1)>|A|,$$

a contradiction. On the other hand, if $\delta=1$, then from (46) we instead conclude that

$$2|A|\geq(2s+1)x\geq(2s+1)\frac{|A|+|B|+1}{2s}\geq(2s+1)\frac{2|A|-s+1}{2s},$$

whence

$$|A| \leq s^2 - \frac{s}{2} - \frac{1}{2}. \tag{48}$$

However, since $2|A| + s \geq |A| + |B| \geq 4s^2 - 6s + 3$, it follows that $|A| \geq \lceil 2s^2 - \frac{7}{2}s + \frac{3}{2} \rceil$, which contradicts (48). Thus we conclude that (47) holds.

For each $r \in \{1, \ldots, n\}$, we have the estimate

$$\begin{aligned} |A+B| &\geq |A_1 + \bigcup_{i=1}^{r-1} B_i| + |A + B_r| + |A_m + \bigcup_{i=r+1}^{n} B_i| \\ &= \sum_{i=1}^{r-1} |B_i| + (r-1)(|A_1| - 1) + |A| + m(|B_r| - 1) + \sum_{i=r+1}^{n} |B_i| + (n-r)(|A_m| - 1) \\ &\geq |A| + |B| - 1 + (|A_1| - 1)(r-1) + (m-1)(|B_r| - 1). \end{aligned} \tag{49}$$

Averaging this estimate over all r, we obtain

$$|A+B| \geq |A| + |B| - 1 + (|A_1| - 1)(\frac{n+1}{2} - 1) + (m-1)(\frac{|B|}{n} - 1). \tag{50}$$

In view of (44) and (45), we have $\max\{m,\, n\} \geq x$ and $\max\{|A_1|,\, |B_1|\} \geq x$. We consider two cases according to whether these maxima are achieved in the same set or in different sets.

Case A: Either $\min\{m,\, |B_1|\} \geq x$ or $\min\{n,\, |A_1|\} \geq x$. By symmetry we may assume that the latter holds. We have the estimate

$$\begin{aligned} |A+B| &\geq |A_1 + (B \setminus B_n)| + |A + B_n| \\ &= |B| - |B_n| + (n-1)(|A_1| - 1) + |A| + m(|B_n| - 1) \\ &\geq |A| + |B| - 1 + (n-1)(|A_1| - 1) \\ &\geq |A| + |B| - 1 + (x-1)^2. \end{aligned} \tag{51}$$

In view of (46) and (51), it follows that

$$k \geq x^2 + 2s - 2 + \delta = \frac{k^2 - 2\alpha k + \alpha^2}{4s^2} + \frac{k - \alpha}{s} + 2s - 1 + \delta.$$

Hence,

$$k^2 - 2(2s^2 - 2s + \alpha)k + (8s^3 - 4s^2 + 4\delta s^2 - 4\alpha s + \alpha^2) \leq 0.$$

Thus, since $\alpha - \delta \leq 2s - 2$, it follows that

$$k \leq 2s^2 - 2s + \alpha + 2s\sqrt{s^2 - 4s + 2 + \alpha - \delta} < 4s^2 - 4s + \alpha.$$

Since $|A| + |B| \equiv \alpha \mod 2s$, the above bound implies that

$$|A| + |B| = k \leq 4s^2 - 6s + \alpha \leq 4s^2 - 4s - 1. \tag{52}$$

Hence, since $k \geq 4s^2 - 6s + 3$, it follows that $k = 4s^2 - 6s + \alpha$, with $\alpha \geq 3$ and $x = 2s - 2$.

Suppose $\max\{m, n\} = x$. If $\alpha < s$, then Lemma 4.1 contradicts (46). Therefore $\alpha \geq s$ and $\delta = 1$. Hence Theorem 1.3(i) and (46) imply that

$$2k - 2x - 2s + 1 \geq 2k - 2x + 1 - \left\lfloor \frac{k}{x} \right\rfloor = 2k - 2x + 1 - (2s - 1), \tag{53}$$

a contradiction. So we may assume $\max\{m, n\} > x$.

Suppose $n \geq x+1$. Hence (51) now implies that $|A+B| \geq |A|+|B|-1+x(x-1)$, which, when combined with (46) and $x = 2s-2$, yields $k \geq 4s^2-4s-1+\delta$, contradicting (52). So we can assume $n = x$ and $m > x$. By this same argument, we also conclude that $|A_1| = x$.

If $|B_1| \geq x$, then interchanging the roles of A and B and repeating the above argument completes the proof. Therefore $|B_1| \leq x-1$. Since $|A_1| = x$, we can apply (3) with the line $\mathbb{Z}x_2$ to obtain

$$|A+B| \geq \left(\frac{|A|}{x}+\frac{|B|}{|B_1|}-1\right)(x+|B_1|-1) = k-(x+|B_1|-1)+\frac{|B_1|-1}{x}|A|+\frac{x-1}{|B_1|}|B|.$$

Considering this bound as a function of $|B_1|$, it follows by the same calculation used in the proof of Theorem 1.3, and in view of $|B_1| < x$ and $||A|-|B|| \leq s \leq 2s-2 = x$, that it is minimized when $|B_1| = x-1$, contradicting (47), and completing the case.

Case B: Either $\min\{m, |A_1|\} \geq x$ or $\min\{n, |B_1|\} \geq x$. By symmetry we may assume that the former holds. Note that we can assume $|B_1| < x$ and $n < x$, else the previous case completes the proof.

If $m = x$, then, in view of $n \leq x-1$ and $||A|-|B|| \leq s \leq x = m$, it follows that the bound given by (3), considered as a function of n, is minimized for the boundary value $n = x-1$, contradicting (47). Therefore we may assume $m > x$. Applying the same arguments with the roles of x_1 and x_2 swapped, we also conclude that $|A_1| > x$. Thus (50) implies that

$$|A+B| \geq |A|+|B|-1+\frac{1}{2}x(n+1+\frac{2|B|}{n})-2x \geq k-1+x(\sqrt{2|B|}+\frac{1}{2})-2x.$$

Hence in view of (46), it follows that

$$x(\sqrt{2|B|}+\frac{1}{2}) \leq k-\delta-2s+3, \tag{54}$$

and consequently,

$$(\frac{k-2s+1}{2s}+1)(\sqrt{2|B|}+\frac{1}{2}) \leq k-2s+3.$$

Thus $\sqrt{2|B|}+\frac{1}{2} < 2s$, implying that $|B| \leq 2s^2-s$, whence $|A|+|B| \leq 4s^2-s$. As a result,

$$x = \begin{cases} 2s, & 4s^2-2s \leq k \leq 4s^2-s \\ 2s-1, & 4s^2-4s \leq k \leq 4s^2-2s-1 \\ 2s-2, & 4s^2-6s+3 \leq k \leq 4s^2-4s-1. \end{cases} \tag{55}$$

There are three cases based on the value of x.

If $x = 2s$, then (55) implies that $k-\delta \leq 4s^2-s-1$, whence (54) implies

$$k \leq 2|B|+s \leq (2s-2+\frac{1}{s})^2+s \leq 4s^2-7s+8,$$

contradicting that $k \geq 4s^2-2s$.

If $x = 2s-1$, then (55) implies that $k-\delta \leq 4s^2-2s-2$, whence (54) implies that

$$k \leq 2|B|+s \leq \lfloor (2s-\frac{3}{2})^2+s \rfloor \leq 4s^2-5s+2.$$

Hence $k \geq 4s^2-4s$ implies that $s = 2$, whence the above inequality becomes $k \leq 4s^2-5s+2 = 8$. Thus (54) then implies that $k \leq 2|B|+s \leq (\frac{7}{3}-\frac{1}{2})^2+2 \leq 6$, contradicting that $k \geq 4s^2-4s = 8$.

Finally, if $x = 2s-2$, then (55) implies that $k-\delta \le 4s^2-4s-2$, whence (54) implies

$$k \le 2|B|+s \le \lfloor (2s-\frac{3}{2}-\frac{1}{2s-2})^2+s\rfloor \le 4s^2-5s.$$

However, $k \le 4s^2-5s$ and (54) imply that $k \le 2|B|+s \le (2s-2)^2+s = 4s^2-7s+4$, contradicting that $k \ge 4s^2-6s+3$, and completing the proof. □

6. Proof of Theorem 1.4

Finally, we conclude with the proof of Theorem 1.4.

Proof. of Theorem 1.4. If $s=1$, then the result follows from Theorem B. If $s=2$, then $|A|>|B|$, and the result follows from [15, Corollary 5.16 with $n=|A|$, $t=|A|-|B|\ge 1$, $d=2$]. So we may assume $s\ge 3$. If $|B|=1$, the result is trivial. So $|B|\ge 2$. By hypothesis,

$$|A| \ge \frac{1}{2}s(s-1)|B|+s. \tag{56}$$

Let $X=(x_1,x_2)$ be an arbitrary ordered basis for $\mathbb{R}^2$, where $\mathbb{R}x_1 = Z_1$ and $\mathbb{R}x_2 = Z_2$. Let $m=|\phi_{Z_1}(A)|$ and $n=|\phi_{Z_1}(B)|$. Note $\max\{m,n\}\ge s$ by hypothesis.

Suppose $m<s$. Then $n\ge s>m$ with $|B|<|A|$, whence Theorem 1.3(i) implies that

$$|A+B| \ge 2|A|+2|B|-2n+1-\frac{|A|+|B|}{n}. \tag{57}$$

Note (56) and $s\ge 3$ imply $|A|\ge 3|B|+s$ so that $2\le s\le n\le |B|\le \frac{|A|+|B|}{4}$. As a result, (57) and (56) yield

$$\begin{aligned} |A+B| &\ge 2|A|+2|B|-3-\frac{|A|+|B|}{2} \ge |A| + \frac{\frac{1}{2}s(s-1)|B|+s}{2}+\frac{3}{2}|B|-3 \\ &= |A|+(\frac{1}{4}s^2-\frac{1}{4}s+\frac{3}{2})|B|+\frac{s}{2}-3 \ge |A|+(\frac{1}{4}s^2-\frac{1}{4}s+\frac{3}{2})|B|-s \ge |A|+s|B|-s, \end{aligned}$$

as desired. So we may assume $|\phi_{Z_1}(A)|=m\ge s$. Moreover, if $m=s$, then (4) follows in view of Theorem 1.3(iii) and (56). Therefore $|\phi_{Z_1}(A)|=m>s$. Since X was arbitrary, this means that $|\phi_Z(A)|>s$ for any one-dimensional subspace Z. In particular, by letting Z be a line such that $|\phi_Z(B)|<|B|$ (recall $|B|\ge 2$), we conclude that $|A+B|\ge |A|+|\phi_Z(A)|\ge |A|+s$. Thus we may assume $|B|\ge 3$, else the proof is complete.

If $n=1$, then (4) follows from (3) and $m>s$. Therefore, as X is arbitrary, it follows that $n\ge 2$ and that $|\phi_Z(B)|\ge 2$ for any one-dimensional subspace Z.

Now assume to the contrary that (4) is false. We will throughout the course of the proof find that the following bound holds for varying values of $n'\ge 1$:

$$|A|+|B|-m-n'+1+\frac{n'-1}{m}|A|+\frac{m-1}{n'}|B| \le |A+B| \le |A|+s|B|-s-1. \tag{58}$$

Inequality (3) shows that the lower bound above holds with $n'=n$. Rearranging the terms in (58), we obtain

$$(\frac{|B|}{n'}-1)m^2-(s|B|-|B|+\frac{|B|}{n'}+n'-s-2)m+(n'-1)|A| \le 0. \tag{59}$$

Applying the estimate (56) yields

$$(\frac{|B|}{n'}-1)m^2-(s|B|-|B|+\frac{|B|}{n'}+n'-s-2)m+(n'-1)(\frac{1}{2}s(s-1)|B|+s)\leq 0. \quad (60)$$

When $|B|>n'$, the discriminant of the above quadratic in m must be nonnegative, i.e.,

$$(s|B|-|B|+M-s-2)^2-2(|B|+1-M)(s^2|B|-s|B|+2s)\geq 0, \quad (61)$$

where $M:=\frac{|B|}{n'}+n'$. Collecting terms, we obtain

$$M^2+(2s^2|B|+2s-2|B|-4)M+4+4|B|-4s^2|B|+|B|^2-s^2|B|^2-4s|B|+s^2\geq 0. \quad (62)$$

Noting that $(2s^2|B|+2s-2|B|-4)\geq 0$, we conclude that (62) must hold for the maximum allowed value for M.

Claim 1. *(58) cannot hold with $n'=2$; consequently, $|\phi_Z(B)|\geq 3$ for any one-dimensional subspace Z.*

Proof. We know that (58) holds with $n'=n$. Thus we need only prove the first part of the claim. Suppose to the contrary that (58) holds with $n'=2$. Thus considering (59) as a quadratic in m, we conclude that the discriminant is nonnegative, i.e., that

$$\begin{aligned}|A| \quad &\leq \quad \frac{(s|B|-\frac{|B|}{2}-s)^2}{2|B|-4}=\frac{(2s-1)^2|B|-4(2s-1)s|B|+4s^2}{8|B|-16} && (63)\\ &= \quad \frac{1}{8}(2s-1)^2|B|-\frac{1}{4}(2s-1)+\frac{(s-1)^2}{2|B|-4}, && (64)\end{aligned}$$

which contradicts the hypothesis of (a). Thus we may assume the hypothesis of (b) holds. From (60), we have

$$(|B|-2)m^2-(2s|B|-|B|-2s)m+s(s-1)|B|+2s\leq 0. \quad (65)$$

Considering (65) as a quadratic in m, we see that its minimum occurs for

$$m=\frac{(2s-1)|B|-2s}{2|B|-4}=s-\frac{1}{2}+\frac{s-1}{|B|-2}.$$

However, the hypothesis $|B|\geq\frac{2s+4}{3}$ of (b) implies that $s-\frac{1}{2}+\frac{s-1}{|B|-2}\leq s+1$. Consequently, since $m\geq s+1$, we conclude that (65) is minimized for the boundary value $m=s+1$, whence

$$0\geq(|B|-2)(s+1)^2-(2s|B|-|B|-2s)(s+1)+s(s-1)|B|+2s=2|B|-2,$$

contradicting that $|B|\geq 3$, and completing the claim. □

Claim 2. *If (58) holds with $n'=3$, then $|B|\leq 6$; consequently, if $|B|\geq 7$, then $|\phi_Z(B)|\geq 4$ for any one-dimensional subspace Z.*

Proof. As in the previous claim, we need only prove the first part. Assuming (58) holds with $n'=3$, so that $M=\frac{|B|}{3}+3$, it follows in view of (62) and $s\geq 3$ that

$$\begin{aligned}0 \quad &\leq \quad -s^2|B|^2-10s|B|+6s^2|B|+\frac{4}{3}|B|^2-4|B|+3+18s+3s^2 && (66)\\ &\leq \quad -s^2|B|^2+6s^2|B|+\frac{4}{3}|B|^2+3s^s=-(\frac{23}{27}+\frac{4}{27})s^2|B|^2+6s^2|B|+\frac{4}{3}|B|^2+3s^2\\ &\leq \quad -\frac{23}{27}s^2|B|^2+6s^2|B|+3s^2,\end{aligned}$$

which implies $|B| \leq 7$. However, it can be individually checked that (66) cannot hold for $|B| = 7$, completing the claim. □

Claim 3. *If (58) holds with $n' = 4$, then $|B| \leq 8$; consequently, if $|B| \geq 9$, then $|\phi_Z(B)| \geq 5$ for any one-dimensional subspace Z.*

Proof. Assuming (58) holds with $n' = 4$, so that $M = \frac{|B|}{4} + 4$, it follows in view of (62) and $s \geq 3$ that

$$\begin{aligned} 0 &\leq -s^2|B|^2 - 7s|B| + 8s^2|B| + \frac{9}{8}|B|^2 - 6|B| + 8 + 16s + 2s^2 \\ &< -s^2|B|^2 + 8s^2|B| + \frac{9}{8}|B|^2 + 2s^2 = -(\frac{7}{8} + \frac{1}{8})s^2|B|^2 + 8s^2|B| + \frac{9}{8}|B|^2 + 2s^2 \\ &\leq -\frac{7}{8}s^2|B|^2 + 8s^2|B| + 2s^2 \end{aligned} \tag{67}$$

which implies $|B| \leq 9$. However, it can be individually verified that (67) cannot hold for $|B| = 9$, completing the claim. □

Claim 4. *If $|B| \geq 7$ and Z is any one-dimensional subspace, then*

$$|\phi_Z(A)| > \frac{s|B|}{4}, \quad \textit{when } s \geq 4 \tag{68}$$

$$|\phi_Z(A)| > \frac{s|B|}{5}, \quad \textit{when } s = 3. \tag{69}$$

Proof. Suppose to the contrary that

$$m \leq \frac{s|B|}{4}, \quad \text{when } s \geq 4 \tag{70}$$

$$m \leq \frac{s|B|}{5}, \quad \text{when } s = 3. \tag{71}$$

Note (70) and (71) each implies $m < |A|$. Let $l := \sqrt{\frac{m(m-1)|B|}{|A|-m}}$.

If $s \geq 4$, then (56) and (70) imply

$$l \leq \sqrt{\frac{m^2|B|}{\frac{1}{2}s(s-1)|B| + s - m}} < \sqrt{\frac{s^2|B|^3/16}{\frac{1}{2}s(s-1)|B| - \frac{s|B|}{4}}} = \frac{|B|}{4}\sqrt{\frac{s^2}{\frac{1}{2}s^2 - \frac{3}{4}s}} \leq \frac{\sqrt{5}}{5}|B|. \tag{72}$$

If $s = 3$, then (56) and (71) imply

$$l \leq \sqrt{\frac{m^2|B|}{\frac{1}{2}s(s-1)|B| + s - m}} < \sqrt{\frac{\frac{9}{25}|B|^3}{3|B| - \frac{3}{5}|B|}} \leq \frac{\sqrt{15}}{10}|B|. \tag{73}$$

From the proof of Theorem 1.3, we know that l minimizes (3), and thus that (58) holds with $n' = l$. If $l \leq 3$, then (3) will be minimized for either $n' = 1$, $n' = 2$ or $n' = 3$, whence Claims 1 and 2 imply $|B| \leq 6$. Note that $\frac{1}{3} < \max\{\frac{\sqrt{5}}{5}, \frac{\sqrt{15}}{10}\}$. Hence if $s \geq 4$, then (72) implies that

$$M = \frac{|B|}{l} + l \leq \frac{5}{\sqrt{5}} + \frac{\sqrt{5}}{5}|B| < \frac{9}{20}|B| + \frac{9}{4}, \tag{74}$$

while if $s = 3$, then (73) implies that

$$M = \frac{|B|}{l} + l \leq \frac{10}{\sqrt{15}} + \frac{\sqrt{15}}{10}|B| < \frac{2}{5}|B| + \frac{13}{5}. \tag{75}$$

Combining (74) and (62) and applying the estimate $s \geq 4$, we obtain

$$\begin{aligned} 0 &\leq -\frac{1}{10}s^2|B|^2 - \frac{31}{10}s|B| + \frac{1}{2}s^2|B| + \frac{121}{400}|B|^2 - \frac{11}{40}|B| + \frac{1}{16} + \frac{9}{2}s + s^2 \qquad (76) \\ &\leq -\frac{1}{10}s^2|B|^2 + \frac{1}{2}s^2|B| + \frac{121}{400}|B|^2 + s^2 \leq -(\frac{19}{240} + \frac{1}{48})s^2|B|^2 + \frac{1}{2}s^2|B| + \frac{1}{3}|B|^2 + s^2 \\ &\leq -\frac{19}{240}s^2|B|^2 + \frac{1}{2}s^2|B| + s^2, \end{aligned}$$

which implies $|B| \leq 7$. However, individually checking the case $|B| = 7$ in (76) shows that in fact $|B| \leq 6$. Combining (75) and (62) and assuming $s = 3$, we obtain

$$-36|B|^2 + 12|B| + 624 \geq 0,$$

which implies $|B| \leq 4$, completing the claim. ☐

Claim 5. *There are s collinear points in A.*

Proof. Suppose instead that A contains no s collinear points. Then it follows from the pigeonhole principle and (56) that

$$|\phi_Z(A)| > \frac{1}{2}s|B| + 1, \qquad (77)$$

for any one-dimensional subspace Z. Consequently, if B has at least 3 collinear points contained in a line parallel to (say) Z, then Theorem B implies

$$|A + B| \geq |A| + 2|\phi_Z(A)| > |A| + 2(\frac{1}{2}s|B| + 1) = |A| + s|B| + 2,$$

as desired. Therefore we may assume B contains no 3 collinear points.

Suppose $h_1(B, B) < |B| - 1$. Then, since B contains no 3 collinear points, it follows that there exists a pair of parallel lines each containing 2 points of B. Hence, by an appropriate affine transformation, we may w.l.o.g assume $(0, 0)$, $(1, 0)$, $(0, 1)$, $(x, 1) \in B$, for some $x > 0$. Let $x_1 = (1, 0)$ and $x_2 = (0, 1)$. Let $A_1 \subseteq A$ be the subset obtained by choosing for each element of $\phi_{Z_1}(A)$ the element of A with largest x_1-coordinate. Let $A_2 \subseteq A$ be likewise defined using Z_2 instead of Z_1. Note $A_1 + (1, 0)$ contains $|\phi_{Z_1}(A)|$ points in $A + B$ disjoint from A.

Let $z+\mathbb{R}x_1$ be an arbitrary line parallel to $\mathbb{R}x_1$, and let $a_1, \ldots, a_r$ be the elements of $A_2 \cap (z+\mathbb{R}x_1)$. Moreover, if $A_1 \cap (z + (0, 1) + \mathbb{R}x_1)$ is nonempty, then there is a unique element $y \in A_1 \cap (z + (0, 1) + \mathbb{R}x_1)$, and so let $a_s, \ldots, a_r$ be those elements of $A_2 \cap (z + \mathbb{R}x_1)$ with $\phi_{Z_1}(a_i) \geq \phi_{Z_1}(y) + 1$. If $A_1 \cap (z + (0, 1) + \mathbb{R}x_1)$ is empty, let $s = r + 1$. Note that for each a_i, $i < s$, the element $a_i + (0, 1)$ is an element of $A + B$ contained in neither A nor $A_1 + (1, 0)$, while for each a_i, $i \geq s$, the element $a_i + (x, 1)$ is an element of $A + B$ contained in nether A nor $A_1 + (1, 0)$ (since $x > 0$). Consequently, since z is arbitrary and since $A_1 + (1, 0)$ contains $|\phi_{Z_1}(A)|$ points from $A + B$ disjoint from A, we conclude that

$$|A + B| \geq |A + \{(0, 0),\ (1, 0),\ (0, 1),\ (x, 1)\}| \geq |A| + |\phi_{Z_1}(A)| + |\phi_{Z_2}(A)| \geq |A| + s|B| + 2,$$

where the latter inequality follows by (77) applied both with $Z = \mathbb{R}x_1$ and $Z = \mathbb{R}x_2$. Thus (4) holds, as desired, and so we may assume $h_1(B, B) = |B| - 1$.

Choose x_1 such that $|\phi_{Z_1}(B)| < |B|$, and let $A' = \mathbf{C}_X(A)$, $B' = \mathbf{C}_X(B)$, $A_i = A' \cap (\mathbb{Z}x_1 + (i-1)x_2)$ and $B_j = B' \cap (\mathbb{Z}x_1 + (j - 1)x_2)$, for $i = 1, \ldots, m$ and $j = 1, \ldots, n$. Note, since $h_1(B, B) = |B| - 1$ and $|\phi_{Z_1}(B)| < |B|$, that $n = |B| - 1$, $|B_1| = 2$, and $|B_i| = 1$ for $i > 1$. Since A contains no s

collinear points, we have $|A_i|\leq s-1$ for all i. Observe, for $j=1,\ldots,m$, that we have the following estimate:

$$|A+B|\geq\sum_{i=1}^{j-1}|A_i+B_1|+\sum_{i=1}^{|B|-1}|A_j+B_i|+\sum_{i=j+1}^{m}|A_i+B_n|=$$
$$|A|+(|B|-2)|A_j|+|B|+(j-1)|B_1|+(m-j)|B_n|-(m+|B|-2)=|A|+(|B|-2)|A_j|+j.$$

Thus, assuming (4) is false, we conclude that

$$|A_j|\leq\frac{s(|B|-1)-j-1}{|B|-2}=s+\frac{s-j-1}{|B|-2}, \tag{78}$$

for $j=1,\ldots,m$. Consequently, for j such that $s+(k-1)(|B|-2)\leq j\leq s+k(|B|-2)-1$, where $k=1,2,\ldots$, we infer that

$$|A_j|\leq s-k. \tag{79}$$

Note that

$$|A_j|\leq s-1 \tag{80}$$

for $j=1,\ldots,s-1$, as remarked earlier. Summing (79) and (80) over all possible j, we conclude that

$$|A|\leq(s-1)^2+(|B|-2)\sum_{k=1}^{s-1}(s-k)=(s-1)^2+(|B|-2)\frac{s(s-1)}{2}=\frac{1}{2}s(s-1)|B|-s+1,$$

contradicting (56), and completing the claim. □

In view of Claim 5, choose x_1 so that there are s points on some line parallel to $\mathbb{Z}x_1$. Let $A'=\mathbf{C}_X(A)$ and $B'=\mathbf{C}_X(B)$. Since $|\phi_{Z_1}(A)|\geq s$ and since A contains s collinear points on a line parallel to $\mathbb{Z}x_1$, it follows that $h_1(A',B')\geq h_1(A',A')\geq s$, whence A' and B' also satisfy the hypotheses of the theorem. Furthermore, if $|A'+B'|\geq|A'|+s(|B'|-1)=|A|+s(|B|-1)$, then the proof is complete in view of (8). Thus we can w.l.o.g. assume $A=A'$ and $B=B'$ are compressed subsets.

Let $A_i=A\cap(\mathbb{Z}x_1+(i-1)x_2)$ and $B_j=B\cap(\mathbb{Z}x_1+(j-1)x_2)$ for $i=1,\ldots,m$ and $j=1,\ldots,n$. By the same estimate used for (51), we have

$$\begin{aligned}|A+B| &\geq |A|+|B|+(n-1)(|A_1|-1)+m(|B_n|-1)-|B_n|\\ &\geq |A|+|B|-1+(n-1)(|A_1|-1).\end{aligned} \tag{81}$$

If $|B|\geq 9$, then Claims 1, 2 and 3 imply $n\geq 5$, whence Claim 4 and (81) imply that

$$|A+B|\geq|A|+|B|-1+4(\frac{s|B|+1}{4}-1)=|A|+(s+1)|B|-4,$$

if $s\geq 4$, and that

$$|A+B|\geq|A|+|B|-1+4(\frac{3|B|+1}{5}-1)=|A|+\frac{17}{5}|B|-\frac{21}{5}>|A|+3|B|-1,$$

if $s=3$. In both cases (4) follows, as desired. So we may assume $|B|\leq 8$. In view of Claim 1 applied with $Z=\mathbb{Z}x_1$ and $Z=\mathbb{Z}x_2$, we infer that $|B|\geq 5$.

Using the estimate from (50) (with the roles of A and B reversed), we obtain

$$|A|+|B|-1+(|B_1|-1)\frac{m-1}{2}+(n-1)(\frac{|A|}{m}-1)\leq|A+B|\leq|A|+s|B|-s-1.$$

Multiplying by m, applying (56), and rearranging terms yields

$$\frac{|B_1|-1}{2}\cdot m^2-(s|B|-|B|+\frac{|B_1|-3}{2}+n-s)m+(n-1)(\frac{1}{2}s(s-1)|B|+s)\leq 0.$$

Consequently, the discriminant of the above quadratic in m must be nonnegative, implying

$$(s|B|-|B|+\frac{|B_1|-3}{2}+n-s)^2-(|B_1|-1)(n-1)(s(s-1)|B|+2s)\geq 0 \tag{82}$$

If $|B|=5$, then from Claim 1, applied with $Z=\mathbb{Z}x_1$ and $Z=\mathbb{Z}x_2$, we conclude $n=|B_1|=3$. Thus (82) implies $4s^2+4s-4\leq 0$, contradicting $s\geq 3$. If $|B|=7$, then from Claims 1 and 2, applied with $Z=\mathbb{Z}x_1$ and $Z=\mathbb{Z}x_2$, we conclude $n=|B_1|=4$. Thus (82) implies $27s^2-15s-\frac{25}{4}\leq 0$, contradicting $s\geq 3$. If $|B|=8$, then from Claims 1 and 2, applied with $Z=\mathbb{Z}x_1$ and $Z=\mathbb{Z}x_2$, we conclude $n\geq 4$ and $|B_1|\geq 4$. Thus (82) implies $23s^2-5s-\frac{49}{4}\leq 0$, contradicting $s\geq 3$. Consequently, it remains only to handle the case $|B|=6$.

In view of Claim 1 and by swapping the roles of x_1 and x_2 if necessary, we may assume $n=3$. Hence (3) implies that (58) holds with $n'=3$. Thus considering (59) as a quadratic in m, we conclude that the discriminant is nonnegative, i.e., that

$$|A| \;\leq\; \frac{(5s-3)^2}{8}=\frac{1}{8}(2s-1)^2|B|-\frac{1}{4}(2s-1)+\frac{(s-1)^2}{2(|B|-2)}. \tag{83}$$

This completes the proof in case (a) holds. From (61), we have

$$0\leq(5s-3)^2-24s^2+16s=s^2-14s+9, \tag{84}$$

which implies $s\geq 14$. Thus $|B|\geq\frac{2s+4}{3}\geq\frac{32}{3}>6$, contradicting the hypothesis of (b), and completing the proof. □

Acknowledgements: We thank the referees for their suggestions as well as Y. Stanchescu and G. Freiman for directing us to Bonneson's Generalization of the Brunn-Minkowski Theorem.

References

[1] Y. Bilu, Structure of sets with small sumset, Structure theory of set addition, *Astérisque* (1999) no. **258**, xi, 77–108.

[2] B. Bollobás and I. Leader, Sums in the grid, *Discrete Math.* **162** (1996), no. 1-3, 31-48.

[3] M.C. Chang, A polynomial bound in Freiman's Theorem, *Duke Math J.* **113** (1994), no. 3, 399-419.

[4] P. Fishburn, On a contribution of Freiman to additive number theory, *J. Number Teory* **35** (1990) 325–334.

[5] G. Freiman, *Foundations of Structural Theory of Set Addition*, Transl. Math. Monographs **37**, Amer. Math. Soc. Providence, R.I. 1973.

[6] G. Freiman, What is the structure of K if $K+K$ is small?, *Lecture Notes in Mathematics* **1240**, Springer, New York, 1987, 189–134.

[7] R.J. Gardner and P. Gronchi, A Brunn–Minkowski inequality for the integer lattice, *Trans. Amer. Math. Soc.* **353** (2001), no. 10, 3995–4024.

[8] B. Green and T. Tao, Compressions, Convex Geometry and the Freiman–Bilu Theorem, *Q. J. Math.*, 57 (2006), no. 4, 495–504.

[9] D. Kleitman, Extremal hypergraph problems, *Surveys in Combinatorics*, ed. by B. Bollobás, Cambridge University Press, Cambridge, 1979, 44–65.

[10] S. Lev and P. Y. Smeliansky, On addition of two sets of integers, *Acta Arith.* **70** (1995), 85–91.

[11] M. Nathanson, *Additive number theory: Inverse problems and the geometry of sumsets*, Graduate Texts in Mathematics **165**, Springer-Verlag, New York, 1996.

[12] W. Rudin, *Real and Complex Analysis*, WCB McGraw Hill, San Francisco, 1987.

[13] I. Ruzsa, Sum of sets in several dimensions, *Combinatorica* **14** (1994), 485–490.

[14] Y. Stanchescu, On the structure of sets with small doubling property on the plane (I), *Acta Arith.* **83** (2) (1998) 127–141.

[15] T. Tao and V. Vu, *Additive combinatorics*, Cambridge Studies in Advanced Mathematics **105**, Cambridge University Press, Cambridge, 2006.