COMBINATORIAL THEORY **5 (3)** (2025), #16 combinatorial-theory.org

IMPROVED STABILITY FOR THE SIZE AND STRUCTURE OF ITERATED SUMSETS IN $\mathbb{Z}^d$

Andrew Granville*[1], Jack Smith†[2], and Aled Walker‡[3]

[1]*Université de Montréal, Montréal, Canada*
andrew@dms.umontreal.ca
[2]*University of Cambridge, U.K.*
jackesmith100@gmail.com
[3]*King's College London, U.K.*
aledwalker@gmail.com

Submitted: Jun 17, 2024; Accepted: Apr 24, 2025; Published: Sep 15, 2025

Abstract. Let $A \subset \mathbb{Z}^d$ be a finite set. It is known that the sumset NA has predictable size ($|NA| = P_A(N)$ for some $P_A(X) \in \mathbb{Q}[X]$) and structure (all of the lattice points in some finite cone other than all of the lattice points in a finite collection of exceptional subcones), once N is larger than some threshold. In previous work, the first effective bounds for both of these thresholds were established, for an arbitrary set A. In this article we substantially improve each of these bounds, coming much closer to the corresponding lower bounds known.

Keywords. Sumsets, Set addition, Khovanskii polynomial, Structure Theorem, Explicit Bounds

Mathematics Subject Classifications. 11P21, 05B10, 11B13, 11P70, 05A16

1. Introduction

Let $A \subset \mathbb{Z}^d$ be a finite set, and for each positive integer N consider the sumset

$$NA := \{a_1 + \cdots + a_N : a_i \in A \text{ for all } i\}.$$

When N is sufficiently large, NA becomes rigidly structured. In this article we study two indicators of such structure, establishing that the values of N which are "sufficiently large" are not too large (and indeed are near to what we would guess are the smallest such N).

---

*Partially supported by the Natural Sciences and Engineering Research Council of Canada.

†Supported by St John's College, Cambridge.

‡Supported by the Swedish Research Council under grant no. 2021-06594 while in residence at Institut Mittag-Leffler in Djursholm, Sweden.

The first notion involves the size $|NA|$. Start with the convex hull of A, denoted

$$H(A) := \Big\{ \sum_{a \in A} c_a a : \text{ Each } c_a \in \mathbb{R}_{\geqslant 0}, \sum_{a \in A} c_a = 1 \Big\}.$$

Certainly $NA \subset NH(A) \cap \mathbb{Z}^d$, and therefore $|NA| \leqslant |NH(A) \cap \mathbb{Z}^d|$. Ehrhart showed ([Ehr62] [BR15, Theorem 3.8]) that there is a polynomial $R_A \in \mathbb{Q}[X]$ of degree at most d for which

$$|NH(A) \cap \mathbb{Z}^d| = R_A(N)$$

for all positive integers N. Therefore $|NA| \leqslant R_A(N)$ for all N but one can readily find examples for which $|NA| < R_A(N)$ for all N: for instance, if $d = 1$ and $A = \{0, 3, 5\}$ then $H(A) = [0, 5]$ and $|NA| = 5N - 5 < R_A(N) = 5N + 1$ for all $N \geqslant 3$.

Even though $|NA|$ is not equal to the Ehrhart polynomial in this example, it is still equal to a polynomial in N once N is sufficiently large. This was established when $A \subset \mathbb{Z}$ by Nathanson [Nat72], using an explicit combinatorial argument and, remarkably, this holds in arbitrary dimension:

Theorem 1.1 (Khovanskii [Kho92])**.** *Let $A \subset \mathbb{Z}^d$ be finite. There is a polynomial $P_A \in \mathbb{Q}[X]$ of degree at most d, and a threshold $N_{\mathrm{Kh}}(A)$, such that $|NA| = P_A(N)$ provided $N \geqslant N_{\mathrm{Kh}}(A)$.*

Khovanskii's proof related the sequence $N \mapsto |NA|$ to the Hilbert function of a certain graded module over the polynomial ring $\mathbb{C}[X_1, \ldots, X_\ell]$ (where $\ell = |A|$), and so agrees with the Hilbert polynomial of the graded module once $N \geqslant N_{\mathrm{Kh}}(A)$. But Hilbert's proof [Eis95, Theorem 1.11] does not yield an explicit bound on $N_{\mathrm{Kh}}(A)$. Nathanson and Ruzsa [NR02] later gave a combinatorial proof of Theorem 1.1, but this did not provide an effective bound on $N_{\mathrm{Kh}}(A)$ either, relying on the following well-known principle (proved in [GS20, Lemma 5], say).

Lemma 1.2 (The Mann–Dickson Lemma)**.** *For any $S \subset \mathbb{Z}^d_{\geqslant 0}$ there exists a finite subset $S_{\min} \subset S$ such that for all $s \in S$ there exists $x \in S_{\min}$ with $s - x \in \mathbb{Z}^d_{\geqslant 0}$.*

Prior to 2021, explicit bounds were known when $d = 1$ ([Nat72, WCC11, GS20, GW21], with the strongest results in [Lev22]); when $H(A)$ is a d-simplex ([CG21], with a refinement in [GSW23]); or when $H(A)$ is d-dimensional and $|A| = d + 1$ or $d + 2$ ([CG21], with a refinement in [GSW23]).

In [GSW23], the first effective bounds for arbitrary d and arbitrary $A \subset \mathbb{Z}^d$ were proved. In particular, [GSW23, Theorem 1.1] showed that

$$N_{\mathrm{Kh}}(A) \leqslant (2|A| \cdot \mathrm{width}(A))^{(d+4)|A|}, \tag{1.1}$$

where $\mathrm{width}(A) := \max_{a_1, a_2 \in A} \|a_1 - a_2\|_\infty$ is the 'width' of A. The proof used a complicated explicit linear algebra argument to bound $|S_{\min}|$ in the Nathanson–Ruzsa argument.

By returning to Khovanskii's original approach, and adapting techniques in Gröbner bases from [Stu96b, Chapter 4] as applied to toric ideals, we have been able to greatly improve (1.1). To state the new bound, we define two quantities that will occur frequently throughout.

Definition 1.3. If $A = \{a_1, \dots, a_\ell\} \subset \mathbb{Z}^d$, we define

$$\mathrm{Vol}^{\dagger,\max}(H(A)) := \max_{\{i_0,i_1,\dots,i_d\}\subset\{1,\dots,\ell\}} |\det(a_{i_1} - a_{i_0}, \cdots, a_{i_d} - a_{i_0})|$$
$$\mathrm{Vol}^{\dagger,\min}(H(A)) := \min_{\substack{\{i_0,i_1,\dots,i_d\}\subset\{1,\dots,\ell\} \\ \det(a_{i_1}-a_{i_0},\cdots,a_{i_d}-a_{i_0})\neq 0}} |\det(a_{i_1} - a_{i_0}, \cdots, a_{i_d} - a_{i_0})|.$$

Remark 1.4. Note that $\frac{1}{d!}\mathrm{Vol}^{\dagger,\max}(H(A))$ is equal to the volume of the largest d-simplex subtended by elements of A. In particular $\mathrm{Vol}^{\dagger,\max}(H(A)) \leqslant d!\,\mathrm{Vol}(H(A))$. By Hadamard's inequality we also have $\mathrm{Vol}^{\dagger,\max}(H(A)) \leqslant d^{d/2}\,\mathrm{width}(A)^d$.

Letting $\Lambda_{A-A} \subset \mathbb{Z}^d$ denote the lattice generated by $A - A$, our first result is as follows.

Theorem 1.5 (Improved Khovanskii threshold)**.** *Suppose that $A \subset \mathbb{Z}^d$ is finite and that Λ_{A-A} is d-dimensional. Then*

$$N_{\mathrm{Kh}}(A) \leqslant |A|^2\,\mathrm{Vol}^{\dagger,\max}(H(A)) - |A| + 1. \tag{1.2}$$

Our proof is motivated by notions in algebraic geometry (as in [Stu96b, Chapter 4]), but we present a simpler, more-or-less equivalent, formulation using only linear algebra. This will be useful when considering the second notion of structure for NA, discussed below.

Theorem 1.5 implies the upper bounds

$$N_{\mathrm{Kh}}(A) \leqslant d!|A|^2\,\mathrm{Vol}(H(A)) - |A| + 1$$

and

$$N_{\mathrm{Kh}}(A) \leqslant |A|^2 d^{d/2}\,\mathrm{width}(A)^d - |A| + 1,$$

indicating the scale of improvement over (1.1). These may be compared with lower bounds. For example, when $|A| = d + 2$ and $\Lambda_{A-A} = \mathbb{Z}^d$ it was shown in [CG21, Theorem 1.2] that $N_{\mathrm{Kh}}(A) = d!\,\mathrm{Vol}(H(A)) - d - 1$. This means that Theorem 1.5 is optimal up to the $|A|^2$ term (and the $|A|^2$ term in (1.2) cannot be replaced by 1, as $\mathrm{Vol}^{\dagger,\max}(H(A)) < d!\,\mathrm{Vol}(H(A))$ for some sets with $|A| = d + 2$).

The bound on $N_{\mathrm{Kh}}(A)$ is related to an influential conjecture in algebraic geometry called the Eisenbud–Goto regularity conjecture [MP18]. Though now known to be false in full generality, the conjecture may still be true for projective toric varieties, which is the relevant case for bounding $N_{\mathrm{Kh}}(A)$. A proof of this case of the conjecture would imply $N_{\mathrm{Kh}}(A) \leqslant d!\,\mathrm{Vol}(H(A)) - |A| + O_d(1)$ which, given the above comments on the bounds when $|A| = d + 2$, would be essentially optimal. We direct the interested reader to [Stu97, Conjectures 4.1 and 4.2], also available at [Stu96a], and to [Stu96b, Chapter 4].

Strong, in some ways optimal, bounds on $|NA|$ when $N < N_{\mathrm{Kh}}(A)$ are given in [EM22, EM23].

For the second notion of structure, we consider the inclusion $NA \subset NH(A)\cap\mathbb{Z}^d$ in more detail, an inclusion introduced in [GS20], with antecedents in Khovanskii's original paper [Kho92].

Let $\mathrm{ex}(H(A)) \subset A$ denote the set of extremal points of the polytope $H(A)$, and translate A so that $0 \in \mathrm{ex}(H(A))$ and $\Lambda_A = \mathbb{Z}^d$, without loss of generality. Then $NH(A) \subset C_A$ where

$$C_A := \Big\{ \sum_{a\in A} c_a a : \ c_a \in \mathbb{R}_{\geqslant 0} \text{ for all } a \Big\}$$

is the cone generated by A, and the semigroup generated by A is the nested union (as $0 \in A$)

$$\mathcal{P}(A) := \bigcup_{N=1}^{\infty} NA \subset C_A \cap \mathbb{Z}^d.$$

The set of *exceptional elements* are those lattice points in C_A which do not belong to $\mathcal{P}(A)$,

$$\mathcal{E}(A) := (C_A \cap \mathbb{Z}^d) \setminus \mathcal{P}(A),$$

and so

$$NA \subset (NH(A) \cap \mathbb{Z}^d) \setminus \mathcal{E}(A).$$

Similarly, for all $a \in \mathrm{ex}(H(A))$ we have $0 \in a - \mathrm{ex}(H(A)) = \mathrm{ex}(H(a-A))$. Since $\Lambda_{a-A} = \mathbb{Z}^d$ too, we have

$$N(a-A) \subset (NH(a-A) \cap \mathbb{Z}^d) \setminus \mathcal{E}(a-A).$$

Rearranging and taking the intersection over all $a \in \mathrm{ex}(H(A))$, we get

$$NA \subset (NH(A) \cap \mathbb{Z}^d) \setminus \Big(\bigcup_{a\in \mathrm{ex}(H(A))} (aN - \mathcal{E}(a-A)) \Big). \tag{1.3}$$

It was shown in [GS20] that there is equality in (1.3) once N is sufficiently large. That is, there exists a constant $N_{\mathrm{Str}}(A)$ such that if $N \geqslant N_{\mathrm{Str}}(A)$ then

$$NA = (NH(A) \cap \mathbb{Z}^d) \setminus \Big(\bigcup_{a\in \mathrm{ex}(H(A))} (aN - \mathcal{E}(a-A)) \Big), \tag{1.4}$$

so that NA fills out to its maximal possible size. This was proved by Nathanson [Nat72] when $d = 1$ and for $d \geqslant 2$ in [GS20]; however the proof in [GS20] did not produce an upper bound for the value of $N_{\mathrm{Str}}(A)$ as it relied on the ineffective Lemma 1.2. The article [GSW23, Theorem 1.3] then gave the first effective bound on $N_{\mathrm{Str}}(A)$ for all A:

$$N_{\mathrm{Str}}(A) \leqslant (d|A| \cdot \mathrm{width}(A))^{13d^6}. \tag{1.5}$$

Previous bounds for $N_{\mathrm{Str}}(A)$ were known when $d = 1$ [Nat72, WCC11, GS20, GW21, Lev22] and when $H(A)$ is a d-simplex ([CG21], with refined bounds in [GSW23]).

The proof of (1.5) in [GSW23] was intricate involving an “induction on dimension” strategy. This required repeated use of Siegel’s Lemma from quantitative linear algebra (in the version proved by Bombieri–Vaaler [BV83]) together with delicate geometric considerations, such as the size and shape of the intersection between neighbourhoods of two cones C_A and C_B.

Our second main result gives a strengthening of (1.5), with a much simpler proof. This is based in part on ideas from the proof of Theorem 1.5, developed out of the ideas in [Stu96b, Chapter 4]. Before stating the result, we introduce one final quantity associated to A.

Definition 1.6. Let $A \subset \mathbb{R}^d$ be finite with $\operatorname{span}(A - A) = \mathbb{R}^d$. Given a facet (i.e. $(d-1)$-dimensional face) F of $H(A)$, and a point $a \in A \setminus F$, let $\operatorname{Vol}(F, a)$ denote the volume of the polytope given by the convex hull of F and a. Then set

$$\kappa(A) = \max_F \frac{\max_a \operatorname{Vol}(F,a)}{\min_a \operatorname{Vol}(F,a)}.$$

Remark 1.7. There are several equivalent ways to define $\kappa(A)$. Indeed, for each F we could equivalently replace $\operatorname{Vol}(F, a)$ by $g_F(a)$ for any affine-linear function $g_F : \mathbb{R}^d \to \mathbb{R}$ which vanishes on F and is strictly positive on $H(A) \setminus F$. For example, we could take $g_F(a)$ to be the (signed) orthogonal distance from F to a. Or we could pick linearly independent points $b^{(1)}, \ldots, b^{(d)}$ in $F \cap A$ and let $g_F(a) = \det(b^{(1)} - a, \ldots, b^{(d)} - a)$, where the $b^{(j)}$ are ordered to make this determinant positive for $a \in H(A) \setminus F$. Using the latter choice of g_F we see that

$$\kappa(A) \leqslant \frac{\operatorname{Vol}^{\dagger,\max}(H(A))}{\operatorname{Vol}^{\dagger,\min}(H(A))}. \tag{1.6}$$

Our main result is the first general effective bound for $N_{\mathrm{Str}}(A)$ that captures the geometry of A by involving the quantities $\kappa(A)$, $\operatorname{Vol}(H(A))$, and $\operatorname{Vol}^{\dagger,\max}(H(A))$:

Theorem 1.8 (Improved structural threshold)**.** *Let $A \subset \mathbb{Z}^d$ be a finite set, with $0 \in \operatorname{ex}(H(A))$ and $\Lambda_A = \mathbb{Z}^d$. Then we have the following two upper bounds:*

$$N_{\mathrm{Str}}(A) \leqslant (d+1)\kappa(A)\Big(d! \operatorname{Vol}(H(A)) + \big(|\operatorname{ex}(H(A))| - d - 1\big) \operatorname{Vol}^{\dagger,\max}(H(A))\Big) \tag{1.7}$$

and

$$N_{\mathrm{Str}}(A) \leqslant (d+1)\kappa(A)\big(|A| - d - 1\big) \operatorname{Vol}^{\dagger,\max}(H(A)). \tag{1.8}$$

The bound (1.7) is better when $|A|$ is substantially larger than $\operatorname{ex}(H(A))$; and (1.8) when $d! \operatorname{Vol}(H(A))$ is substantially larger than $\operatorname{Vol}^{\dagger,\max}(H(A))$. Using (1.6) and bounding $\operatorname{Vol}^{\dagger,\min}(H(A)) \geqslant 1$ and $\operatorname{Vol}^{\dagger,\max}(H(A)) \leqslant d! \operatorname{Vol}(H(A))$, (1.7) implies the cleaner but slightly weaker bound

$$N_{\mathrm{Str}}(A) \leqslant (d+1)(d!)^2(|\operatorname{ex}(H(A))| - d) \operatorname{Vol}(H(A))^2, \tag{1.9}$$

but still much stronger than (1.5).

Similar to the Khovanskii threshold, we guess that a bound like $N_{\mathrm{Str}}(A) \leqslant d! \operatorname{Vol}(H(A))$ holds in general. Our (1.9) is roughly the square of this bound, so still far from optimal.

If $H(A)$ is a simplex then $\kappa(A) = 1$ and $|\operatorname{ex}(H(A))| = d+1$, so (1.7) implies that $N_{\mathrm{Str}}(A) \leqslant (d+1)! \operatorname{Vol}(H(A))$, which essentially recovers the best known bound in this case [GSW23, Theorem 1.5].

The paper is structured as follows. In Section 2 we prove Theorem 1.5, while developing general lemmas on equations $\sum_{a \in A} a = \sum_{b \in B} b$ (with $B \subset A$) that will be useful throughout. In Section 3 we use these lemmas, with some convex geometry, to deduce Theorem 1.8.

2. Proof of Theorem 1.5

Let $A = \{a_1, \ldots, a_\ell\}$. To keep track of various quantities throughout the proof, we define the *weight* of a vector $m \in \mathbb{Z}^\ell$ by $\mathrm{wt}(m) := m \cdot 1 = \sum_{i=1}^{\ell} m_i$. We also let $A_{\mathrm{mat}} := (a_1, \ldots, a_\ell)$ be the d-by-ℓ matrix formed with the a_i as column vectors, so that $A_{\mathrm{mat}} m = \sum_i m_i a_i$.

We begin with a result and proof due to Nathanson and Ruzsa.

Proposition 2.1. *There exists a finite set of lattice points $\mathcal{M} \subset \mathbb{Z}^\ell_{\geqslant 0}$, as described in the proof, such that for all positive integers h we have*

$$|hA| = \sum_{T \subset \mathcal{M}} (-1)^{|T|} \binom{h - \mathrm{wt}(m_T) + \ell - 1}{\ell - 1},$$

where m_T is the vector with $(m_T)_i := \max_{m \in T} (m)_i$, and $\binom{N}{\ell-1} = 0$ if $N < \ell - 1$.

Proof. If $x \in hA$ then let

$$\mathrm{rep}_h(x) := \{m \in \mathbb{Z}^\ell_{\geqslant 0} : \mathrm{wt}(m) = h \text{ and } A_{\mathrm{mat}} m = x\},$$

denote the coefficient set of non-negative combinations of h elements of A that represent x. Let $m_h(x)$ be the minimum element in $\mathrm{rep}_h(x)$ with respect to the lexicographic ordering, and let

$$\mathcal{U} := \bigcup_{h \geqslant 0} \bigcup_{x \in hA} \{m \in \mathrm{rep}_h(x) : m \neq m_h(x)\}.$$

Evidently $|hA| = |\{m_h(x) : x \in hA\}|$. We will calculate this size using the relationship

$$\{m_h(x) : x \in hA\} = \{n \in \mathbb{Z}^\ell_{\geqslant 0} :\ \mathrm{wt}(n) = h \text{ and } n \notin \mathcal{U}\}.$$

To this end, note that $\mathcal{U} + \mathbb{Z}^\ell_{\geqslant 0} = \mathcal{U}$. Indeed, if $y \in \mathbb{Z}^\ell_{\geqslant 0}$ then writing $v = A_{\mathrm{mat}} y$ and $k = \mathrm{wt}(y)$ we have $\mathrm{rep}_h(x) + y \subset \mathrm{rep}_{h+k}(x + v)$. So if $m \in \mathrm{rep}_h(x) \cap \mathcal{U}$ then

$$m + y >_{\mathrm{lex}} m_h(x) + y \geqslant_{\mathrm{lex}} m_{h+k}(x + v).$$

So $m + y \in \mathcal{U}$ as needed.

Applying Lemma 1.2 (the Mann–Dickson Lemma) to $\mathcal{U} \subset \mathbb{Z}^\ell_{\geqslant 0}$ we see that the set

$$\mathcal{M} = \mathcal{M}(\mathcal{U}) := \{m \in \mathcal{U} : \text{for all } u \in \mathcal{U},\ u \leqslant_{\mathrm{coord}} m \implies u = m\}$$

of minimal elements is finite, where $x \leqslant_{\mathrm{coord}} y$ if $x_i \leqslant y_i$ for all i. In particular, for every $u \in \mathcal{U}$ there exists some $m \in \mathcal{M}$ with $m \leqslant_{\mathrm{coord}} u$. Therefore we can use inclusion-exclusion to obtain

$$\begin{aligned} &\{n \in \mathbb{Z}^\ell_{\geqslant 0} :\ \mathrm{wt}(n) = h \text{ and } n \notin \mathcal{U}\} \\ &\quad = \sum_{T \subset \mathcal{M}} (-1)^{|T|} \{n \in \mathbb{Z}^\ell_{\geqslant 0} :\ \mathrm{wt}(n) = h \text{ and } m \leqslant_{\mathrm{coord}} n \,\forall m \in T\} \\ &\quad = \sum_{T \subset \mathcal{M}} (-1)^{|T|} \{n \in \mathbb{Z}^\ell_{\geqslant 0} :\ \mathrm{wt}(n) = h \text{ and } m_T \leqslant_{\mathrm{coord}} n\}. \end{aligned}$$

We have written this in terms of sets, and one should think of $+S$ as including the elements of S with multiplicity, and $-S$ as removing one copy of each element of S.

Note that if $n \in \mathbb{Z}_{\geqslant 0}^{\ell}$ with $\mathrm{wt}(n) = h$ and $m_T \leqslant_{\mathrm{coord}} n$, then $\mathrm{wt}(m_T) \leqslant h$. In that case, writing $n = m_T + r$ we obtain

$$\{n \in \mathbb{Z}_{\geqslant 0}^{\ell} : \ \mathrm{wt}(n) = h \text{ and } m_T \leqslant_{\mathrm{coord}} n\} = \{r \in \mathbb{Z}_{\geqslant 0}^{\ell} : \ \mathrm{wt}(r) = h - \mathrm{wt}(m_T)\}$$

and the result follows from the usual 'stars and bars' bound. □

The remainder of the proof of Theorem 1.5 concerns bounding $\mathrm{wt}(m_T)$ above. To describe this argument we introduce some more notation, which will be of use throughout the paper. Given $u \in \mathbb{Z}^{\ell}$ define vectors $u^+, u^- \in \mathbb{Z}_{\geqslant 0}^{\ell}$ where $(u^+)_i = \max\{0, u_i\}$ and $(u^-)_i = \max\{0, -u_i\}$, so that $u = u^+ - u^-$. If $u \in \mathbb{Z}^{\ell}$, we let $\mathrm{supp}(u) = \{i : \ u_i \neq 0\}$. Next let

$$\mathcal{Z} = \mathcal{Z}(A) := \{z \in \mathbb{Z}^{\ell} : \mathrm{wt}(z) = 0 \text{ and } A_{\mathrm{mat}} z = 0\} \tag{2.1}$$

which is a lattice. By taking $m = z^+, m' = z^-$ with $x = A_{\mathrm{mat}} m$ and $h = \mathrm{wt}(m)$ we obtain

$$\mathcal{Z} = \bigcup_{h \geqslant 0} \bigcup_{x \in hA} \{m - m' : m, m' \in \mathrm{rep}_h(x)\}.$$

We continue with a lemma relating $\mathcal{M}$ and $\mathcal{Z}$.

Lemma 2.2. *Given* $m \in \mathcal{M}$ *let* $x = A_{\mathrm{mat}} m$ *and* $h = \mathrm{wt}(m)$. *Then* $\mathrm{supp}(m) \cap \mathrm{supp}(m_h(x)) = \varnothing$. *Moreover if there exists* $v \in \mathcal{Z} \setminus \{0\}$ *with* $v^+ \leqslant_{\mathrm{coord}} m$ *and* $v^- \leqslant_{\mathrm{coord}} m_h(x)$ *then* $v^+ = m$, $v^- = m_h(x)$, *and* $v = m - m_h(x)$.

Proof. Write $n = m_h(x)$ so that $A_{\mathrm{mat}}(m - n) = x - x = 0$. If $\mathrm{supp}(m) \cap \mathrm{supp}(n) \neq \varnothing$, say that $m_i, n_i \geqslant 1$. Then $m - e_i >_{\mathrm{lex}} n - e_i$ with $m - e_i, n - e_i \in \mathbb{Z}_{\geqslant 0}^{\ell}$ so $m - e_i \in \mathcal{U}$, and $m - e_i <_{\mathrm{coord}} m$, contradicting that $m \in \mathcal{M}(\mathcal{U})$. So $\mathrm{supp}(m) \cap \mathrm{supp}(n) = \varnothing$ as claimed.

Now if $v \in \mathcal{Z} \setminus \{0\}$ with $v^+ \leqslant_{\mathrm{coord}} m$ and $v^- \leqslant_{\mathrm{coord}} n$ then $v^+ >_{\mathrm{lex}} v^-$. Indeed, if not then $w := v^+ + n - v^- <_{\mathrm{lex}} n$ and $w \in \mathbb{Z}_{\geqslant 0}^{\ell}$. Moreover $m - w = (m - n) - (v^+ - v^-) \in \mathcal{Z}$, so that $w \in \mathrm{rep}_h(x)$ with $w <_{\mathrm{lex}} n$. However $n = m_h(x) \leqslant_{\mathrm{lex}} w$ by definition, which gives a contradiction. So $v^+ >_{\mathrm{lex}} v^-$.

Finally, let $y = A_{\mathrm{mat}} v^+$ and $k = \mathrm{wt}(v^+)$, so that $v^+, v^- \in \mathrm{rep}_k(y)$ and $v^+ \in \mathcal{U}$ (as $v^+ >_{\mathrm{lex}} v^-$). Then $v^+ \leqslant_{\mathrm{coord}} m$ and $m \in \mathcal{M}$, so $v^+ = m$. Therefore $y = x, k = h$ and so $v^- \in \mathrm{rep}_h(x)$ which implies that $n \leqslant_{\mathrm{lex}} v^-$. Moreover $v^- \leqslant_{\mathrm{coord}} n$ which implies $v^- \leqslant_{\mathrm{lex}} n$, and so $v^- = n$. So $v = v^+ - v^- = m - n$ as claimed. □

We define $\mathcal{Z}^{\dagger} = \mathcal{Z}^{\dagger}(A)$ by

$$\mathcal{Z}^{\dagger} := \{u \in \mathcal{Z} \setminus \{0\} : \text{ If } v \in \mathcal{Z} \setminus \{0\} \text{ with } \mathrm{supp}(v) \subset \mathrm{supp}(u) \text{ then } v = \lambda u \text{ for some } \lambda \in \mathbb{Z}\}.$$

Note that if $v \in \mathcal{Z} \setminus \{0\}$ then there must exist some $u \in \mathcal{Z}^{\dagger}$ with $\mathrm{supp}(u) \subset \mathrm{supp}(v)$. It transpires that elements in $\mathcal{Z}^{\dagger}$ may be strongly controlled, and this in turn will help control $\mathcal{Z}$ and finally $\mathcal{M}$.

Lemma 2.3. *If* $u \in \mathscr{Z}^\dagger$ *then* $\|u\|_\infty := \max_i |u_i| \leqslant \mathrm{Vol}^{\dagger,\max}(H(A))$.

Proof. For each $a_i \in A$ let $b_i = \left(\begin{smallmatrix} a_i \\ 1 \end{smallmatrix}\right) \in \mathbb{Z}^{d+1}$ If $\mathrm{supp}(u) = \{i_1, \ldots, i_r\}$ then the only linear dependence (up to scalars) amongst the the vectors $\{b_{i_1}, \cdots, b_{i_r}\}$ is $\sum_j u_{i_j} b_{i_j} = 0$, so that the $(d+1)$-by-r matrix $M = (b_{i_1}, \cdots, b_{i_r})$ has rank $r-1$. Since $\mathrm{span}(\{b_1, \ldots, b_\ell\}) = \mathbb{R}^{d+1}$ (as Λ_{A-A} is d-dimensional), we can find column vectors $b_{i_{r+1}}, \ldots b_{i_{d+2}}$ such that the $(d+1)$-by-$(d+2)$ matrix $M' = (b_{i_1}, \cdots, b_{i_{d+2}})$ has rank $d+1$ and so has a 1-dimensional null space. Cramer's rule gives a non-zero null vector

$$w := \sum_{j=1}^{d+2} (-1)^j \det(b_{i_1} \cdots, b_{i_{j-1}}, b_{i_{j+1}}, \cdots b_{i_{d+2}}) \cdot e_{i_j} \tag{2.2}$$

(non-zero as the subdeterminants cannot all be zero since M' has rank $d+1$). We already have the null vector u, so w and u must be scalar multiples of one another. In particular $\mathrm{supp}(w) \subset \mathrm{supp}(u)$, and hence $w = \lambda u$ for some $\lambda \in \mathbb{Z}$. This implies that $|u_{i_j}| \leqslant |w_{i_j}| \leqslant |\det(b_{i_1} \cdots, b_{i_{j-1}}, b_{i_{j+1}}, \cdots b_{i_{d+2}})|$ for all j, and hence

$$\|u\|_\infty \leqslant \max_{\{k_0, k_1, \ldots, k_d\} \subset \{1, \ldots, \ell\}} \left| \det \left(\begin{pmatrix} a_{k_0} \\ 1 \end{pmatrix}, \cdots, \begin{pmatrix} a_{k_d} \\ 1 \end{pmatrix} \right) \right| = \mathrm{Vol}^{\dagger,\max}(H(A)).$$

For the final equality, we have relabelled $\{b_{i_1}, \cdots, b_{i_{j-1}},\ b_{i_{j+1}}, \cdots b_{i_{d+2}}\}$ as $\{a_{k_0}, \ldots, a_{k_d}\}$ and then subtracted the first column from the others, expanding the determinant about the bottom row. □

We continue by relating $\mathscr{Z}^\dagger$ and $\mathscr{Z}$. To this end we write $\mathrm{supp}(u^\pm) \subset \mathrm{supp}(v^\pm)$ as shorthand for the two conditions $\mathrm{supp}(u^+) \subset \mathrm{supp}(v^+)$ and $\mathrm{supp}(u^-) \subset \mathrm{supp}(v^-)$

Lemma 2.4. *If* $v \in \mathscr{Z} \setminus \{0\}$ *then there exists* $u \in \mathscr{Z}^\dagger$ *such that* $\mathrm{supp}(u^\pm) \subset \mathrm{supp}(v^\pm)$.

Proof. Let $w = v/\gcd_i v_i$ (where $v = (v_1, \ldots, v_\ell) \in \mathbb{Z}^\ell$). If $w \in \mathscr{Z}^\dagger$ let $u = w$ and we are done, so we may assume that $w \in \mathscr{Z} \setminus \mathscr{Z}^\dagger$. If $|\mathrm{supp}(v)| = 1$ then $w \in \mathscr{Z}^\dagger$ automatically, so we assume that $|\mathrm{supp}(v)| \geqslant 2$ and proceed by induction on $|\mathrm{supp}(v)|$. Since $w \in \mathscr{Z} \setminus \mathscr{Z}^\dagger$ is non-zero there exists $u \in \mathscr{Z} \setminus \{0\}$ with $\mathrm{supp}(u) \subset \mathrm{supp}(w)$ but which is not an integer multiple of w. Select $\lambda := \min_{i: u_i \neq 0} |w_i/u_i|$, which is > 0 as $\mathrm{supp}(u) \subset \mathrm{supp}(w)$. Pick i so that $w_i = \pm \lambda u_i$ with $u_i \neq 0$, and then adjust the sign of u so that $u_i > 0$ (and one can adjust the sign of u since $\mathscr{Z}$ is symmetric by definition). Now let $y := u_i w - w_i u$, so that $y_i = 0$ and for all j either y_j equals 0 or has the same sign as w_j, since $|w_i u_j| = \lambda |u_i u_j| = |u_i| \cdot \lambda |u_j| \leqslant |u_i w_j|$. Therefore $\mathrm{supp}(y) \subset \mathrm{supp}(w) \setminus \{i\}$ with $\mathrm{supp}(y^\pm) \subset \mathrm{supp}(w^\pm)$. Note further that $y \neq 0$, since if $0 = y = u_i w - w_i u$ then $w_i \neq 1$, since u is not an integer multiple of w, but this in turn contradicts the coprimality of the coordinates of w.

Since $y \in \mathscr{Z} \setminus \{0\}$, by the induction hypothesis there exists $u \in \mathscr{Z}^\dagger$ for which

$$\mathrm{supp}(u^\pm) \subset \mathrm{supp}(y^\pm) \subset \mathrm{supp}(w^\pm) = \mathrm{supp}(v^\pm).$$ □

We may iterate this argument to entirely decompose elements $v \in \mathscr{Z}$ in terms of a combination of elements $u \in \mathscr{Z}^\dagger$.

Lemma 2.5 (Decomposing using $\mathcal{Z}^\dagger$)**.** *Any* $v \in \mathcal{Z} \setminus \{0\}$ *can be written as* $\sum_{j=1}^{I} \lambda_j u_j$ *with each* $u_j \in \mathcal{Z}^\dagger, \lambda_j \in \mathbb{Q}_{>0}$ *and* $I \leqslant |\operatorname{supp}(v)|$. *Furthermore each* $\operatorname{supp}(u_j^\pm) \subset \operatorname{supp}(v^\pm)$, *so that* $v^+ = \sum_{j=1}^{I} \lambda_j u_j^+$ *and* $v^- = \sum_{j=1}^{I} \lambda_j u_j^-$.

Proof. We prove Lemma 2.5 by induction on $m := |\operatorname{supp}(v)|$. Select $u \in \mathcal{Z}^\dagger$ by Lemma 2.4, let $\lambda := \min_{i: u_i \neq 0} v_i/u_i$ (noting v_i and u_i have the same sign as $\operatorname{supp}(u^\pm) \subset \operatorname{supp}(v^\pm)$), choose i so that $v_i = \lambda u_i$, and let $y := v - \lambda u$. Now $\operatorname{supp}(y) \subset \operatorname{supp}(w) \setminus \{i\}$ so that $|\operatorname{supp}(y)| \leqslant m-1$. If $y = 0$ (for example if $m = 1$) then $v = \lambda u$. Otherwise the result follows by the induction hypothesis. ☐

The preceding lemmas may be combined to control the size of elements in $\mathcal{M}$.

Lemma 2.6. *If* $m \in \mathcal{M}$ *then* $\|m\|_\infty \leqslant \ell \operatorname{Vol}^{\dagger,\max}(H(A))$.

Proof. Let $x = A_{\mathrm{mat}} m$ and $h = \operatorname{wt}(m)$. By Lemma 2.2, letting $u := m - m_h(x) \in \mathcal{Z}$ we have $u \neq 0$, and if $v \in \mathcal{Z} \setminus \{0\}$ with $v^+ \leqslant_{\mathrm{coord}} u^+$ and $v^- \leqslant_{\mathrm{coord}} u^-$ then $v = u$. Now by Lemma 2.5 we can write $u = \sum_{j=1}^{I} \lambda_j u_j$ with each $u_j \in \mathcal{Z}^\dagger, \lambda_j \in \mathbb{Q}_{>0}$ and $I \leqslant |\operatorname{supp}(u)|$, where each $\operatorname{supp}(u_j^\pm) \subset \operatorname{supp}(u^\pm)$, so that $u^+ = \sum_{j=1}^{I} \lambda_j u_j^+$ and $u^- = \sum_{j=1}^{I} \lambda_j u_j^-$.

We claim each $\lambda_j \leqslant 1$, else $u_j^+ <_{\mathrm{coord}} \lambda_j u_j^+ \leqslant_{\mathrm{coord}} u^+$ and $u_j^- <_{\mathrm{coord}} \lambda_j u_j^- \leqslant_{\mathrm{coord}} u^-$. Then applying Lemma 2.2 as above with $v = u_j$ we conclude $u_j = u$, but this contradicts the strict inequality $u_j^+ <_{\mathrm{coord}} u^+$.

Now m and $m_h(x)$ have disjoint support by Lemma 2.2, and therefore, by the triangle inequality,

$$\|m\|_\infty \leqslant \|u\|_\infty \leqslant \sum_{j=1}^{I} \lambda_j \|u_j\|_\infty \leqslant I \cdot \max_j \|u_j\|_\infty \leqslant \ell \cdot \max_{u \in \mathcal{Z}^\dagger} \|u\|_\infty$$

and the result then follows from Lemma 2.3. ☐

Substituting this control on $\mathcal{M}$ into the Proposition 2.1 will quickly resolve Theorem 1.5.

Proof of Theorem 1.5. Define

$$P_A(x) := \frac{1}{(\ell-1)!} \sum_{T \subset \mathcal{M}} (-1)^{|T|} (x - \operatorname{wt}(m_T) + \ell - 1) \cdots (x - \operatorname{wt}(m_T) + 1).$$

We observe that

$$\frac{(h - \operatorname{wt}(m_T) + \ell - 1) \cdots (x - \operatorname{wt}(m_T) + 1)}{(\ell-1)!} = \binom{h - \operatorname{wt}(m_T) + \ell - 1}{\ell - 1}$$

for all integers $h \geqslant \operatorname{wt}(m_T) - \ell + 1$. Therefore, by Proposition 2.1,

$$|h\mathcal{A}| = P_A(h) \text{ for all } h \geqslant \operatorname{wt}(m_{\mathcal{M}(\mathcal{U})}) - \ell + 1$$

since $\max_{T \subset \mathcal{M}(\mathcal{U})} \operatorname{wt}(m_T) = \operatorname{wt}(m_{\mathcal{M}(\mathcal{U})})$, by definition. Hence

$$N_{\mathrm{Kh}}(A) + \ell - 1 \leqslant \operatorname{wt}(m_{\mathcal{M}(\mathcal{U})}) = \sum_{i=1}^{\ell} \max_{m \in \mathcal{M}(\mathcal{U})} |m|_i \leqslant \ell \max_{m \in \mathcal{M}(\mathcal{U})} \|m\|_\infty \leqslant \ell^2 \operatorname{Vol}^{\dagger,\max}(H(A)).$$

by Lemma 2.6. This is the claimed bound on $N_{\mathrm{Kh}}(A)$. ☐

3. Proof of Theorem 1.8

Let us first describe the general strategy. Let $A \subset \mathbb{Z}^d$ be finite with $0 \in \mathrm{ex}(H(A))$ and $\Lambda_A = \mathbb{Z}^d$. If $v \in \mathcal{P}(A)$, we aim to find u and w such that $v = u + w$, where:

- $u \in MA$, for a bounded M;
- $w \in \mathcal{P}(B \cup \{0\})$, where $B \subset A$ is contained within a single facet of $H(A)$. One may also assume that this facet does not contain the origin.

In some ways, this strategy is similar to [GSW23, Lemma 7.1]. However, in [GSW23, Lemma 7.1] the set B was pre-determined at the outset, with the further assumptions that $v \in \mathcal{P}(A) \cap C_B$ and the further requirement that $u \in C_B$. It turns out to be much easier to prove the weaker version outlined above, where B is found as a consequence of the decomposition $v = u + w$ rather than being fixed in the hypotheses.

To find the decomposition $v = u + w$, one may consider a representation $v = \sum_{i=1}^{\ell} \eta_i a_i$ in which $\eta \in \mathbb{Z}^{\ell}_{\geqslant 0}$ and the weight $\mathrm{wt}(\eta)$ is minimal. Recall that $A = \{a_1, \dots, a_\ell\}$ and $\mathrm{wt}(\eta)$ denotes $\sum_i \eta_i$. In this section it will actually be more convenient to index η directly by A itself, so $v = \sum_{a \in A} \eta_a a$ and $\mathrm{wt}(\eta) = \sum_{a \in A} \eta_a$. The basic idea is then to let

$$u = \sum_{\substack{a \in A \\ \eta_a \text{ is small}}} \eta_a a \quad \text{and} \quad w = \sum_{\substack{a \in A \\ \eta_a \text{ is large}}} \eta_a a.$$

If the set $\{a \in A : \eta_a \text{ is large}\}$ is not contained within a single facet of $H(A)$, one can use properties of the sets $\mathcal{Z}(A)$ and $\mathcal{Z}^{\dagger}(A)$ established previously (Lemmas 2.3 and 2.5) to reduce $\mathrm{wt}(\eta)$, contradicting minimality.

Having proved this decomposition, suppose $v \in NH(A)$ as well, with N at least the right-hand side of (1.8). If the facet on which B lies is defined by $\beta = 1$ for a linear map $\beta : \mathbb{R}^d \to \mathbb{R}$, then one can apply β to both sides of the equation $v = u + w$. Writing $w = \sum_{b \in B} \lambda_b b$, we have $\beta(w) = \mathrm{wt}(\lambda)$, and this enables us to bound $\mathrm{wt}(\lambda)$ above in terms of N and M. Putting everything together, we can place $v \in NA$ as required. (We extend the definition of wt to mean simply the sum of the entries of a vector. We will also implicitly allow ourselves to enlarge the indexing set of a vector, by setting all previously undefined entries to zero.)

This method gives (1.8). In order to prove (1.7), which involves $|\mathrm{ex}(H(A))|$ instead of $|A|$, one first excises the contribution from non-extremal elements (Lemma 3.7 below). This is a simple additive–combinatorial argument, adapted from similar results in [CG21] and [GSW23]. This done, one proceeds as above but with A replaced by $\mathrm{ex}(H(A))$.

To begin the proof proper, we state some standard results on convex polytopes. Let $A \subset \mathbb{Z}^d$ be finite, and assume $0 \in \mathrm{ex}(H(A))$ and $\mathrm{span}(A) = \mathbb{R}^d$. Then from the structure theorem for convex polytopes [Brø83, Theorem 9.2], we know that there are linear maps $\beta_1, \dots, \beta_K, \gamma_1, \dots, \gamma_L, : \mathbb{R}^d \to \mathbb{R}$ for which

$$H(A) = \bigcap_{i=1}^{K} \{x \in \mathbb{R}^d : \beta_i(x) \leqslant 1\} \cap \bigcap_{j=1}^{L} \{x \in \mathbb{R}^d : \gamma_j(x) \geqslant 0\}$$

and the sets $\{x \in H(A) : \beta_i(x) = 1\}$ and $\{x \in H(A) : \gamma_j(x) = 0\}$ form the facets of $H(A)$. For each i and j we call $\{x \in H(A) : \beta_i(x) = 1\}$ an *outer facet* of $H(A)$ and $\{x \in H(A) : \gamma_j(x) = 0\}$ an *inner facet* of $H(A)$.

We continue with a technical lemma which we will use to reduce $\mathrm{wt}(n)$ as discussed above.

Lemma 3.1 (Preparation for reduction step). *Let $A \subset \mathbb{Z}^d$ be finite, and assume $0 \in \mathrm{ex}(H(A))$ and $\mathrm{span}(A) = \mathbb{R}^d$. Let $S \subset A$, and suppose that S does not lie in an outer facet of $H(A)$. Then for any linear map $\alpha : \mathbb{R}^d \to \mathbb{R}$ satisfying $\alpha(s) = 1$ for all $s \in S$ there exists $p \in H(A) \cap \mathrm{span}(S) \cap \mathbb{Q}^d$ for which $\alpha(p) > 1$.*

Proof. Each outer facet of $H(A)$ is defined by $\{x \in H(A) : \beta_i(x) = 1\}$ for some linear map β_i. Now $\beta_i(s) \leqslant 1$ for all $s \in S$ as $S \subset A \subset H(A)$, and we cannot have equality for all $s \in S$ as S is not contained in any outer facet by the hypothesis, and so the *barycentre*

$$q := \frac{1}{|S|} \sum_{s \in S} s$$

of S satisfies $\beta_i(q) < 1$. Letting $\hat{\beta} = \max_i \beta_i(q) \in [0, 1)$, we see that q lies inside $\hat{\beta} H(A)$, and so for any $\varepsilon \in (0, \hat{\beta}^{-1} - 1) \cap \mathbb{Q}$ the point $p = (1 + \varepsilon)q$ lies in $H(A)$. This p is clearly also in $\mathrm{span}(S)$ and $\mathbb{Q}^d$, and satisfies $\alpha(p) = (1 + \varepsilon)\alpha(q) = 1 + \varepsilon > 1$. ☐

We now use this observation to prove the existence of certain relations between sums of elements in A.

Lemma 3.2 (Reduction step). *Let $A \subset \mathbb{Z}^d$ be finite, and assume $0 \in \mathrm{ex}(H(A))$ and $\mathrm{span}(A) = \mathbb{R}^d$. Suppose the elements of $S \subset A$ are linearly independent, and that S is not a subset of any outer facet of $H(A)$. Then there exist non-negative integers $\{\lambda_s\}_{s \in S}$ and $\{\rho_a\}_{a \in A \setminus \{0\}}$ such that*

$$\sum_{s \in S} \lambda_s s = \sum_{a \in A \setminus \{0\}} \rho_a a \quad \textit{and} \quad \mathrm{wt}(\lambda) > \mathrm{wt}(\rho),$$

where $\lambda_s, \rho_a \leqslant \mathrm{Vol}^{\dagger,\max}(H(A))$ for all $s \in S$ and $a \in A \setminus \{0\}$.

Proof. As S is linearly independent, we know that $0 \notin S$ and there exists a linear map $\alpha : \mathbb{R}^d \to \mathbb{R}$ with $\alpha(s) = 1$ for all $s \in S$. From Lemma 3.1, choose $p \in H(A) \cap \mathrm{span}(S) \cap \mathbb{Q}^d$ with $\alpha(p) > 1$. Therefore $p = \sum_{s \in S} \gamma_s s$ for some coefficients $\gamma_s \in \mathbb{Q}$ (and we let $\gamma_a = 0$ for all $a \in A \setminus S$), and $p = \sum_{a \in A} \delta_a a$ where $\mathrm{wt}(\delta) = 1$ and $\delta_a \in [0, 1] \cap \mathbb{Q}$ for all $a \in A$. Then

$$\sum_{s \in S} \gamma_s s = p = \sum_{a \in A} \delta_a a \quad \text{and} \quad \mathrm{wt}(\gamma) = \alpha(p) > 1 = \mathrm{wt}(\delta).$$

Let L be the least common denominator of all the γ_s and δ_a. Define $z_a = L(\delta_a - \gamma_a)$ for $a \in A \setminus \{0\}$, and

$$z_0 = L\big(\mathrm{wt}(\gamma) - \mathrm{wt}(\delta) + \delta_0\big) > 0,$$

so that

$$\mathrm{wt}(z) = z_0 + \sum_{a \in A \setminus \{0\}} L(\delta_a - \gamma_a) = 0.$$

We then have $z \in \mathcal{Z}$ (as defined in (2.1), where we identify $\mathbb{Z}^A$ with $\mathbb{Z}^\ell$) and $\operatorname{supp}(z^-) \subset S$.

By Lemma 2.5, we write $z = \sum_j \eta_j u_j$ with each $u_j \in \mathcal{Z}^\dagger$, $\eta_j \in \mathbb{Q}_{>0}$, and $\operatorname{supp}(u_j^\pm) \subset \operatorname{supp}(z^\pm)$. Now $0 \in \operatorname{supp}(z^+)$ as $z_0 > 0$, so $0 \in \operatorname{supp}(u^+) \subset \operatorname{supp}(z^+)$ for some $u = u_j \in \mathcal{Z}^\dagger$ and $\operatorname{supp}(u^-) \subset \operatorname{supp}(z^-) \subset S$. Define

$$\lambda_s = \begin{cases} (-u_s) & \text{if } s \in \operatorname{supp}(u^-) \\ 0 & \text{if } s \in S \setminus \operatorname{supp}(u^-), \end{cases} \quad \text{and } \rho_a = \begin{cases} u_a & \text{if } a \in \operatorname{supp}(u^+) \setminus \{0\} \\ 0 & \text{otherwise.} \end{cases}$$

Then, since $u \in \mathcal{Z}$, we have

$$\sum_{s \in S} \lambda_s s = \sum_{a \in \operatorname{supp}(u^-)} (-u_a) a = \sum_{a \in \operatorname{supp}(u^+)} u_a a = \sum_{a \in A \setminus \{0\}} \rho_a a$$

$$\text{and} \quad \operatorname{wt}(\lambda) = \sum_{a \in \operatorname{supp}(u^-)} (-u_a) = \sum_{a \in \operatorname{supp}(u^+)} u_a > \operatorname{wt}(\rho).$$

The final inequality uses the fact that $u_0 > 0$. The condition

$$\max_{s \in S, a \in A} \lambda_s, \rho_a = \|u\|_\infty \leqslant \operatorname{Vol}^{\dagger,\max}(H(A))$$

then follows from Lemma 2.3. $\square$

Using the relation from the previous lemma, we can derive the decomposition $v = u + w$ as discussed at the start of the section.

Lemma 3.3 (Regular representation). *Let $A \subset \mathbb{Z}^d$ be finite, and assume $0 \in \operatorname{ex}(H(A))$ and $\operatorname{span}(A) = \mathbb{R}^d$. Let $v \in \mathcal{P}(A)$. Then there is a decomposition $v = u + w$ and an outer facet F of $H(A)$ for which*

- *$w \in \mathcal{P}(B \cup \{0\})$, where $B = A \cap F$;*
- *$u \in MA$ where $M = (|A| - 1 - |B|)(\operatorname{Vol}^{\dagger,\max}(H(A)) - 1)$.*

Proof. By definition we can write $v \in \mathcal{P}(A)$ as a $\mathbb{Z}_{\geqslant 0}$-linear combination of the $a \in A$, so we select the representation

$$v = \sum_{a \in A \setminus \{0\}} \eta_a a \text{ where each } \eta_a \in \mathbb{Z}_{\geqslant 0}$$

for which $\operatorname{wt}(\eta)$ is minimal, and then for which

$$T = T((\eta_a)_a) := \{a \in A \setminus \{0\} : \eta_a \geqslant \operatorname{Vol}^{\dagger,\max}(H(A))\}$$

is also minimal. If T is contained in an outer facet F of $H(A)$ then we obtain the desired decomposition $v = u + w$, where

$$u := \sum_{a \in A \setminus (B \cup \{0\})} \eta_a a \in MA \quad \text{and} \quad w := \sum_{b \in B := A \cap F} \eta_b b,$$

since $\eta_a \leqslant \operatorname{Vol}^{\dagger,\max}(H(A)) - 1$ for all $a \in A \setminus (B \cup \{0\})$, and $|A \setminus (B \cup \{0\})| = |A| - 1 - |B|$.

Henceforth we may assume that no such facet F exists, so that $T \neq \varnothing$. We obtain a contradiction as follows.

Case I: If the elements of T are linearly independent then we apply Lemma 3.2 with $S := T$ to obtain another representation of v,

$$v = \sum_{a \in A \setminus \{0\}} \eta'_a a, \quad \text{where} \quad \eta'_a := \begin{cases} \eta_a - \lambda_a + \rho_a & \text{if } a \in T; \\ \eta_a + \rho_a & \text{otherwise.} \end{cases}$$

The coefficients η'_a are all non-negative since each $\eta_a, \rho_a \geqslant 0$ and

$$\lambda_t \leqslant \mathrm{Vol}^{\dagger,\max}(H(A)) \leqslant \eta_t \text{ for all } t \in T.$$

However

$$\mathrm{wt}(\eta') = \mathrm{wt}(\eta) + \mathrm{wt}(\rho) - \mathrm{wt}(\lambda) < \mathrm{wt}(\eta),$$

contradicting the minimality of $\mathrm{wt}(\eta)$.

Case II: Otherwise the elements of T are linearly dependent and so there exist $z_t \in \mathbb{Z}$, not all zero, for which

$$\sum_{t \in T} z_t t = 0.$$

Define $z_0 := -\sum_{t \in T} z_t$, and multiply through all the z_v-values by -1 if necessary to ensure that $z_0 \geqslant 0$. We also define $z_a = 0$ for all $a \in A \setminus (T \cup \{0\})$, so we can consider z as a non-zero element of $\mathbb{Z}^A$ with $z \in \mathcal{Z}$. By Lemma 2.4 there then exists $\mu \in \mathcal{Z}^\dagger$ with $\mathrm{supp}(\mu^\pm) \subset \mathrm{supp}(z^\pm)$, and by Lemma 2.3 we have $\|\mu\|_\infty \leqslant \mathrm{Vol}^{\dagger,\max}(H(A))$.

Case IIa: If $\mu_0 \neq 0$ then we must have $\mu_0 > 0$ since $\mathrm{supp}(\mu^\pm) \subset \mathrm{supp}(z^\pm)$ and $z_0 > 0$. Now write $v = \sum_{a \in A} \eta'_a a$, where $\eta'_a = \eta_a + \mu_a$ for $a \neq 0$ and $\eta'_0 = 0$. We have $\eta'_a \geqslant 0$ for all a, since η'_a agrees with $\eta_a \geqslant 0$ unless $a \in T$, in which case $\eta_a \geqslant \mathrm{Vol}^{\dagger,\max}(H(A))$ and $\mu_a \geqslant -\|\mu\|_\infty \geqslant -\mathrm{Vol}^{\dagger,\max}(H(A))$. But we also have

$$\mathrm{wt}(\eta') = \mathrm{wt}(\eta) + \mathrm{wt}(\mu) - \mu_0 = \mathrm{wt}(\eta) - \mu_0 < \mathrm{wt}(\eta),$$

contradicting the minimality of $\mathrm{wt}(\eta)$.

Case IIb: Otherwise $\mu_0 = 0$. Since $\mu_a \leqslant \|\mu\|_\infty \leqslant \mathrm{Vol}^{\dagger,\max}(H(A)) \leqslant \eta_a$ for each $a \in T$, the components of $\eta - \mu$ are all non-negative. We select $n \in \mathbb{N}$ to be the largest integer for which the components of $\eta' := \eta - n\mu$ are all non-negative. We obtain $v = \sum_{a \in A} \eta'_a a$ and $\mathrm{wt}(\eta') = \mathrm{wt}(\eta)$. But we must have $\eta'_t < \mathrm{Vol}^{\dagger,\max}(H(A))$ for some $t \in T$, otherwise we can increase n, so $T(\eta')$ must be a proper subset of T, contradicting minimality of T. ☐

In order to leverage the decomposition $v = u + w$ to show that $v \in NA$, we need to control how negative the evaluation $\beta(u)$ can get, when β defines an outer facet of $H(A)$. This is the purpose of the next lemma. Recall from Definition 1.6 and Remark 1.7 that

$$\kappa(A) = \max_F \frac{\max_a g_F(a)}{\min_a g_F(a)},$$

where F ranges over facets of $H(A)$, a ranges over points of $A \setminus F$, and g_F is any affine-linear function $\mathbb{R}^d \to \mathbb{R}$ which vanishes on F and is strictly positive on $H(A) \setminus F$.

Lemma 3.4 (Negative coefficients)**.** *Let* $A\subset\mathbb{R}^d$ *be finite with* $0\in\mathrm{ex}(H(A))$ *and* $\mathrm{span}(A)=\mathbb{R}^d$. *Let* $\beta:\mathbb{R}^d\to\mathbb{R}$ *be a linear map for which* $F=\{x\in H(A):\beta(x)=1\}$ *is an outer facet of* $H(A)$*. Then, for all* $a\in A$*,*

$$\beta(a)\geqslant 1-\kappa(A).$$

Proof. The facet F of the d-dimensional convex polytope $H(A)$ is the convex hull of at least d points of A (see [GSW23, Lemma A.2 (4)] for a discussion). In particular there are d linearly independent $b^{(1)},\dots,b^{(d)}\in A$ for which $\beta(b^{(j)})=1$ (and these uniquely determine β). Let $b_i^{(j)}$ denote the i^{th} coordinate of $b^{(j)}$ with respect to the standard basis, and for $a\in A$ let a_i denote the i^{th} coordinate with respect to the standard basis. Expressing β in coordinates and computing the necessary matrix inverses, we derive

$$\beta(a)=\frac{1}{\det B_{\mathrm{mat}}}\sum_{i,j\leqslant d}(-1)^{i+j}a_iM_{ij},$$

where B_{mat} is the d-by-d matrix with $(B_{\mathrm{mat}})_{ij}=b_i^{(j)}$, and M_{ij} is the minor formed by deleting the i^{th} row and j^{th} column of B_{mat} and taking the determinant. Yet

$$\det B_{\mathrm{mat}}-\sum_{i,j\leqslant d}(-1)^{i+j}a_iM_{ij}=\det\begin{pmatrix}1 & 1 & \cdots & 1\\ a_1 & b_1^{(1)} & \cdots & b_1^{(d)}\\ \vdots & \vdots & & \vdots\\ a_d & b_d^{(1)} & \cdots & b_d^{(d)}\end{pmatrix},$$

as can be seen from expanding the determinant along the top row, and from column operations we have

$$\det\begin{pmatrix}1 & 1 & \cdots & 1\\ a_1 & b_1^{(1)} & \cdots & b_1^{(d)}\\ \vdots & \vdots & & \vdots\\ a_d & b_d^{(1)} & \cdots & b_d^{(d)}\end{pmatrix}=\det(b^{(1)}-a,\cdots,b^{(d)}-a).$$

Letting $g_F(a)=\det(b^{(1)}-a,\cdots,b^{(d)}-a)$, and assuming that the $b^{(j)}$ are ordered so that $\det B_{\mathrm{mat}}=g_F(0)$ is positive, we obtain

$$\beta(a)=1-\frac{\det(b^{(1)}-a,\cdots,b^{(d)}-a)}{\det B_{\mathrm{mat}}}=1-\frac{g_F(a)}{g_F(0)}\geqslant 1-\kappa(A),$$

as required. □

Remark 3.5. Less explicitly, one can argue that $1-\beta$ and $\frac{g_F}{g_F(0)}$ are the unique affine-linear functions $\mathbb{R}^d\to\mathbb{R}$ which vanish on F and map 0 to 1, so they must agree.

We can now deduce part of Theorem 1.8:

Proof of bound (1.8). Let

$$v\in(NH(A)\cap\mathbb{Z}^d)\setminus\Big(\bigcup_{b\in\mathrm{ex}(H(A))}(bN-\mathcal{E}(b-A))\Big)$$

where $N \geqslant (d+1)N_0$ with $N_0 := \kappa(A)(|A|-d-1)\operatorname{Vol}^{\dagger,\max}(H(A))$.

Since $v \in NH(A)$ there is some subset $S = \{s_0, \dots, s_d\} \subset \operatorname{ex}(H(A))$ with $v \in NH(S)$, by Caratheodory's theorem [Brø83, Corollary 2.5]. Writing $v = \sum_{s \in S} c_s s$, with $c_s \geqslant 0$ for all s and $\operatorname{wt}(c) = N$, there must be some $c_s \geqslant N_0$ as $N \geqslant (d+1)N_0$. By re-labelling the vectors in S we may assume that $c_{s_0} \geqslant N_0$, and then

$$v' := s_0 N - v = \sum_{s \in S \setminus \{s_0\}} c_s(s_0 - s) \in (N - c_{s_0})H(s_0 - S) \subset (N - N_0)H(s_0 - S).$$

Letting $A' = s_0 - A$, we have $s_0 - S \subset A'$, so by the preceding equation v' is contained in $(N-N_0)H(A')$. We also have, by assumption, that $v \notin s_0 N - \mathcal{E}(s_0 - A)$, so $v' \notin \mathcal{E}(s_0 - A) = \mathcal{E}(A')$. We conclude that $v' \in \mathcal{P}(A')$. Note also that $|A'| = |A|$, $|\operatorname{ex}(H(A'))| = |\operatorname{ex}(H(A))|$, $\operatorname{Vol}(H(A')) = \operatorname{Vol}(H(A))$, $\operatorname{Vol}^{\dagger,\max}(A') = \operatorname{Vol}^{\dagger,\max}(A)$, and the same for $\min$.

Now apply Lemma 3.3 to v', with A' in place of A. We see that there exists an outer facet F of $H(A')$ such that we can write $v' = u + w$, where

$$w = \sum_{b \in B} \lambda_b b \quad \text{and} \quad u = \sum_{a \in A'} \eta_a a$$

with $B := A' \cap F$, all $\eta_a, \lambda_b \in \mathbb{Z}_{\geqslant 0}$, and

$$\operatorname{wt}(\eta) \leqslant (|A'| - |B| - 1)(\operatorname{Vol}^{\dagger,\max}(H(A')) - 1) \leqslant (|A| - d - 1)\operatorname{Vol}^{\dagger,\max}(H(A)). \tag{3.1}$$

The second inequality here uses that $|B| \geqslant d$, which follows from the fact that every facet of the d-dimensional convex polytope $H(A')$ is the convex hull of at least d points of A'.

We know that $F = \{x \in H(A') : \beta(x) = 1\}$ for some linear map $\beta : \mathbb{R}^d \to \mathbb{R}$. Now $\beta(y) \leqslant 1$ for all $y \in H(A')$ and therefore, for $v' \in (N-N_0)H(A')$, we have

$$N - N_0 \geqslant \beta(v') = \beta(u) + \beta(w) = \sum_{a \in A'} \eta_a \beta(a) + \operatorname{wt}(\lambda)$$

as $\beta(b) = 1$ for each $b \in B$. Moreover, combining (3.1) with Lemma 3.4 applied to A' gives

$$\operatorname{wt}(\eta) - \sum_{a \in A'} \eta_a \beta(a) = \sum_{a \in A'} \eta_a (1 - \beta(a)) \leqslant \kappa(A)\operatorname{wt}(\eta) \leqslant N_0.$$

Summing the last two inequalities we then obtain

$$\operatorname{wt}(\eta) + \operatorname{wt}(\lambda) \leqslant N$$

and so $v' \in NA'$. Therefore $v = s_0 N - v' \in s_0 N - NA' = N(s_0 - A') = NA$ as required. ☐

It remains to prove the bound (1.7), which separates the contribution from $\operatorname{ex}(H(A))$. To effect this separation, we begin with an argument about triangulating polytopes.

Lemma 3.6 (Splitting A into simplices centred at the origin)**.** *Let $A \subset \mathbb{Z}^d$ be finite with $0 \in \mathrm{ex}(H(A))$ and $\mathrm{span}(A) = \mathbb{R}^d$. Then $H(A)$ may be partitioned as a finite union of simplices $\cup_j H(B^{(j)} \cup \{0\})$, where each $B^{(j)} \subset \mathrm{ex}(H(A))$ is a basis of $\mathbb{R}^d$, and for each $i \neq j$ the set $H(B^{(i)} \cup \{0\}) \cap H(B^{(j)} \cup \{0\})$ is contained in a subspace of dimension at most $d-1$. In particular, $H(B^{(i)} \cup \{0\}) \cap H(B^{(j)} \cup \{0\})$ has zero measure.*

When $d = 2$, this is the obvious statement that any polygon with a vertex at the origin may be decomposed into disjoint triangles, all of which have a common vertex at the origin.

Proof. The $d = 1$ case is trivial, so assume that $d \geqslant 2$. We will induct on dimension. Let $F_1, \dots, F_K$ denote the list of outer facets of $H(A)$. Each F_i is a convex polytope of dimension $d-1$, generated by points in $\mathrm{ex}(H(A))$. Therefore, by the induction hypotheses, one may decompose F_i as a union of $(d-1)$-dimensional simplices of the form $H(\{a_1, \dots, a_d\})$, where $\{a_1, \dots, a_d\} \subset \mathrm{ex}(H(A))$ is linearly independent (so as to define a $(d-1)$-dimensional simplex) and the intersection of any two of these simplices is contained in an affine subspace of dimension at most $d-2$. (In fact one may further assume that there is a common vertex a_1 to all these simplices, but that will not be necessary for the induction step.)

Choose $B^{(j)}$ to be the list of such sets $\{a_1, \dots, a_d\}$, taken over all the facets $F_1, \dots, F_K$. We claim that these $B^{(j)}$ satisfy the requirements of the lemma. By construction, each $B^{(j)} \subset \mathrm{ex}(H(A))$ is a basis of $\mathbb{R}^d$. To show that the union of the $H(B^{(j)} \cup \{0\})$ is $H(A)$, fix $x \in H(A) \setminus \{0\}$ and pick (the unique) $\lambda_x \geqslant 1$ such that $\lambda_x x \in \cup_K F_K$. Then $\lambda_x x \in H(B^{(j)})$ for some $B^{(j)}$. Thus there exist coefficients c_b for $b \in B^{(j)}$ such that $c_b \geqslant 0$, $\mathrm{wt}(c) = 1$, and

$$\lambda_x x = \sum_{b \in B^{(j)}} c_b b.$$

We therefore have

$$x = \Big(1 - \frac{1}{\lambda_x}\Big) 0 + \sum_{b \in B^{(j)}} \frac{c_b}{\lambda_x} b \in H(B^{(j)} \cup \{0\}),$$

as wanted.

It remains to show that each intersection $H(B^{(i)} \cup \{0\}) \cap H(B^{(j)} \cup \{0\})$ is contained in a subspace of dimension at most $d-1$. So fix an arbitrary non-zero $x \in H(B^{(1)} \cup \{0\}) \cap H(B^{(2)} \cup \{0\})$. There are coefficients $c_i^{(1)}, c_i^{(2)} \geqslant 0$ with

$$x = \sum_{i \leqslant d} c_i^{(1)} b_i^{(1)} = \sum_{i \leqslant d} c_i^{(2)} b_i^{(2)}$$

and $0 < \mathrm{wt}(c^{(1)}), \mathrm{wt}(c^{(2)}) \leqslant 1$. Letting wt_j denote $\mathrm{wt}(c^{(j)})$, and assuming WLOG that $\mathrm{wt}_2 \geqslant \mathrm{wt}_1$, we can re-scale to obtain

$$y := \frac{x}{\mathrm{wt}_2} = \sum_{i \leqslant d} \frac{c_i^{(1)}}{\mathrm{wt}_2} b_i^{(1)} = \sum_{i \leqslant d} \frac{c_i^{(2)}}{\mathrm{wt}_2} b_i^{(2)},$$

which lies in $H(B^{(1)} \cup \{0\}) \cap H(B^{(2)})$ since $\mathrm{wt}(\frac{c^{(1)}}{\mathrm{wt}_2}) \leqslant 1$ and $\mathrm{wt}(\frac{c^{(2)}}{\mathrm{wt}_2}) = 1$.

We claim that $\mathrm{wt}(\frac{c^{(1)}}{\mathrm{wt}_2}) = 1$. Suppose for contradiction that $\mathrm{wt}(\frac{c^{(1)}}{\mathrm{wt}_2}) < 1$, so $(1+\varepsilon)y \in H(B^{(1)} \cup \{0\}) \subset H(A)$ for all sufficiently small $\varepsilon > 0$. Let $B^{(2)}$ be a subset of the outer facet defined by the linear map $\beta^{(2)} : \mathbb{R}^d \to \mathbb{R}$, so that $B^{(2)} \subset \{u \in \mathbb{R}^d : \beta^{(2)}(u) = 1\}$ and $H(A) \subset \{u \in \mathbb{R}^d : \beta^{(2)}(u) \leqslant 1\}$. Then for all sufficiently small $\varepsilon > 0$ we have

$$1 \geqslant \beta^{(2)}((1+\varepsilon)y) = (1+\varepsilon)\beta^{(2)}(y) = 1+\varepsilon > 1.$$

This gives the desired contradiction, and we deduce that $\mathrm{wt}(\frac{c^{(1)}}{\mathrm{wt}_2}) = 1$. So

$$y \in H(B^{(1)}) \cap H(B^{(2)}),$$

and because $x = \mathrm{wt}_2\, y$ for some $\mathrm{wt}_2 \in [0,1]$ we conclude that

$$H(B^{(1)} \cup \{0\}) \cap H(B^{(2)} \cup \{0\}) = H((H(B^{(1)}) \cap H(B^{(2)})) \cup \{0\}).$$

Hence $H(B^{(1)} \cup \{0\}) \cap H(B^{(2)} \cup \{0\})$ is contained in a subspace of dimension at most $d-1$ by the induction hypothesis (since $H(B^{(1)}) \cap H(B^{(2)})$ is contained in a subspace of dimension at most $d-2$). □

Using this decomposition, we can generalise an additive combinatorial argument from [GSW23] and [CG21] (which was applied when $H(A)$ was a d-simplex).

Lemma 3.7 (Restricting the influence of non-extremal elements)**.** *Let* $A \subset \mathbb{Z}^d$ *be a finite set with* $0 \in \mathrm{ex}(H(A))$ *and* $\mathrm{span}(A) = \mathbb{R}^d$*. Then there exists a finite set* $S = d!\,\mathrm{Vol}(H(A))A$ *for which*

$$\mathcal{P}(A) = S + \mathcal{P}(\mathrm{ex}(H(A))).$$

The proof is similar to (but simpler than) [GSW23, Lemma 3.2] with the set $B := \mathrm{ex}(H(A))$.

Proof. Let $v \in NA$. We will show that $v \in S + \mathcal{P}(\mathrm{ex}(H(A)))$ by induction on N. For $N \leqslant d!\,\mathrm{Vol}(H(A))$ we have $v \in NA \subset S \subset S + \mathcal{P}(\mathrm{ex}(H(A)))$.

Suppose that $N > d!\,\mathrm{Vol}(H(A))$. We can write $v = a_1 + a_2 + \cdots + a_N$ with each $a_i \in A$. By Lemma 3.6, there is a partition $H(A) = \cup_j H(B^{(j)} \cup \{0\})$ where each $B^{(j)} \subset \mathrm{ex}(H(A))$. Therefore we can partition $\{1,\ldots,N\} = \cup_j T_j$ to obtain

$$v = \sum_j \sum_{i \in T_j} a_i,$$

where $i \in T_j$ implies that $a_i \in H(B^{(j)} \cup \{0\})$. (It could be that $a_i \in H(B^{(j)} \cup \{0\}) \cap H(B^{(j')} \cup \{0\})$ in which case we assign i to either T_j or $T_{j'}$, arbitrarily.)

Since $\mathrm{Vol}(H(A)) = \sum_j \mathrm{Vol}(H(B^{(j)} \cup \{0\}))$ by Lemma 3.6 there is some j for which

$$|T_j| > d!\,\mathrm{Vol}(H(B^{(j)} \cup \{0\})) = |\mathbb{Z}^d / \Lambda_{B^{(j)} \cup \{0\}}|,$$

by the pigeonhole principle. Reordering the indices on the a_i we write $T_j = \{1,\ldots,|T_j|\}$. Two of the $|T_j|$ partial sums

$$a_1, a_1 + a_2, \ldots, a_1 + a_2 + \cdots + a_{|T_j|} \bmod \Lambda_{B^{(j)} \cup \{0\}},$$

must be congruent to each other mod $\Lambda_{B^{(j)}\cup\{0\}}$ by the pigeonhole principle. Their difference yields a non-trivial partial sum $\sum_{i\in I} a_i \equiv 0 \bmod \Lambda_{B^{(j)}\cup\{0\}}$ (where $I \subset T_j$ is a non-empty interval) and so this partial sum can be replaced by a sum of elements from $B^{(j)}\cup\{0\}$. Therefore

$$\sum_{i\in I} a_i \in \mathcal{P}(B^{(j)}\cup\{0\}) \subset \mathcal{P}(\mathrm{ex}(H(A))).$$

By the induction hypothesis, we have $v - \sum_{i\in I} a_i \in S + \mathcal{P}(\mathrm{ex}(H(A)))$, and so

$$v \in S + \mathcal{P}(\mathrm{ex}(H(A))) + \mathcal{P}(\mathrm{ex}(H(A))) \subset S + \mathcal{P}(\mathrm{ex}(H(A)))$$

as required. □

We are now ready to finish the argument by modifying the proof of (1.8).

Proof of bound (1.7). Let

$$v \in (NH(A)\cap\mathbb{Z}^d) \setminus \Big(\bigcup_{b\in\mathrm{ex}(H(A))} (bN - \mathcal{E}(b-A))\Big),$$

where $N \geqslant (d+1)N_0$ with

$$N_0 := \kappa(A)\Big(d!\,\mathrm{Vol}(H(A)) + (|\,\mathrm{ex}(H(A))| - d - 1)\,\mathrm{Vol}^{\dagger,\max}(H(A))\Big).$$

As in the proof of (1.8) we use Caratheodory's theorem to determine some $s_0 \in \mathrm{ex}(H(A))$ for which

$$v' := s_0 N - v \in (N-N_0)H(A') \cap \mathcal{P}(A'),$$

where $A' := s_0 - A$. By Lemma 3.7 applied to A', we may write $v' = y + x$ where $y \in d!\,\mathrm{Vol}(H(A'))A'$ and $x \in \mathcal{P}(\mathrm{ex}(H(A')))$. Applying Lemma 3.3 to $x \in \mathcal{P}(\mathrm{ex}(H(A')))$ (in place of $v \in \mathcal{P}(A)$) we write $x = u + w$, where $w \in \mathcal{P}(B\cup\{0\})$ and $u \in M\,\mathrm{ex}(H(A'))$, with $B = \mathrm{ex}(H(A')) \cap F$ for some outer facet F of $H(A')$, and $M = (|\,\mathrm{ex}(H(A'))| - 1 - |B|)(\mathrm{Vol}^{\dagger,\max}(H(A')) - 1)$.

Now let

$$z = y + u = \sum_{a\in A'} \rho_a a \quad \text{and} \quad w = \sum_{b\in B} \lambda_b b,$$

and note that $\rho_a, \lambda_b \in \mathbb{Z}_{\geqslant 0}$ for all a and b. We obtain $v = z + w$ and

$$\mathrm{wt}(\rho) \leqslant d!\,\mathrm{Vol}(H(A)) + (|\,\mathrm{ex}(H(A))| - 1 - d)\,\mathrm{Vol}^{\dagger,\max}(H(A)), \tag{3.2}$$

using $|\,\mathrm{ex}(H(A'))| = |\,\mathrm{ex}(H(A))|$, $\mathrm{Vol}^{\dagger,\max}(H(A')) = \mathrm{Vol}^{\dagger,\max}(H(A))$, and $|B| \geqslant d$, as in the proof of (1.8). The outer facet F is given by $\{x \in H(\mathrm{ex}(H(A'))) : \beta(x) = 1\}$ for some linear $\beta : \mathbb{R}^d \to \mathbb{R}$ so we again obtain

$$N - N_0 \geqslant \beta(v) = \beta(z) + \beta(w) = \sum_{a\in A'} \rho_a \beta(a) + \mathrm{wt}(\lambda).$$

By again applying Lemma 3.4 to A', this time using the bound (3.2) in place of (3.1), we obtain $\mathrm{wt}(\rho) + \mathrm{wt}(\lambda) \leqslant N$ with the modified value for N_0, and so $v' \in NA'$. Therefore $v = s_0 N - v' \in s_0 N - NA' = N(s_0 - A') = NA$ as required. □

Acknowledgements

Our proof of Theorem 1.5 is pretty much that presented in [Stu96b, Chapter 4], albeit written in a different mathematical language and context. It also helped inspire the proof of Theorem 1.8. Many thanks to the anonymous referees for their helpful remarks.

References

[BR15] M. Beck and S. Robins. *Computing the continuous discretely*. Undergraduate Texts in Mathematics. Springer, New York, second edition, 2015. Integer-point enumeration in polyhedra, With illustrations by David Austin. `doi:10.1007/978-1-4939-2969-6`.

[Brø83] A. Brøndsted. *An introduction to convex polytopes*, volume 90 of *Graduate Texts in Mathematics*. Springer-Verlag, New York-Berlin, 1983.

[BV83] E. Bombieri and J. Vaaler. On Siegel's lemma. *Invent. Math.*, 73(1):11–32, 1983. `doi:10.1007/BF01393823`.

[CG21] M. J. Curran and L. Goldmakher. Khovanskii's theorem and effective results on sumset structure. *Discrete Anal.*, pages Paper No. 27, 25, 2021. `doi:10.19086/da.28814`.

[Ehr62] E. Ehrhart. Sur les polyèdres homothétiques bordés à n dimensions. *C. R. Acad. Sci. Paris*, 254:988–990, 1962.

[Eis95] D. Eisenbud. *Commutative algebra*, volume 150 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1995. With a view toward algebraic geometry. `doi:10.1007/978-1-4612-5350-1`.

[EM22] Shalom Eliahou and Eshita Mazumdar. Iterated sumsets and Hilbert functions. *J. Algebra*, 593:274–294, 2022. `doi:10.1016/j.jalgebra.2021.11.019`.

[EM23] Shalom Eliahou and Eshita Mazumdar. Optimal bounds on the growth of iterated sumsets in abelian semigroups. 2023. `arXiv:2310.10241`.

[GS20] A. Granville and G. Shakan. The Frobenius postage stamp problem, and beyond. *Acta Math. Hungar.*, 161(2):700–718, 2020. `doi:10.1007/s10474-020-01073-y`.

[GSW23] A. Granville, G. Shakan, and A. Walker. Effective results on the size and structure of sumsets. *Combinatorica*, 43(6):1139–1178, 2023. `doi:10.1007/s00493-023-00055-2`.

[GW21] A. Granville and A. Walker. A tight structure theorem for sumsets. *Proc. Amer. Math. Soc.*, 149(10):4073–4082, 2021. `doi:10.1090/proc/15608`.

[Kho92] A. G. Khovanskiĭ. Newton polyhedron, Hilbert polynomial, and sums of finite sets. *Funktsional. Anal. i Prilozhen.*, 26(4):57–63, 96, 1992. `doi:10.1007/BF01075048`.

[Lev22] V. F. Lev. The structure of higher sumsets. *Proc. Amer. Math. Soc.*, 150(12):5165–5177, 2022. `doi:10.1090/proc/16128`.

[MP18] J. McCullough and I. Peeva. Counterexamples to the Eisenbud-Goto regularity conjecture. *J. Amer. Math. Soc.*, 31(2):473–496, 2018. `doi:10.1090/jams/891`.

[Nat72] M. B. Nathanson. Sums of finite sets of integers. *Amer. Math. Monthly*, 79(9):1010–1012, 1972. `doi:10.2307/2318072`.

[NR02] M. B. Nathanson and I. Z. Ruzsa. Polynomial growth of sumsets in abelian semigroups. *J. Théor. Nombres Bordeaux*, 14(2):553–560, 2002. URL: `http://jtnb.cedram.org/item?id=JTNB_2002__14_2_553_0`.

[Stu96a] B. Sturmfels. Equations defining toric varieties. 1996. `arXiv:alg-geom/9610018`.

[Stu96b] B. Sturmfels. *Gröbner bases and convex polytopes*, volume 8 of *University Lecture Series*. American Mathematical Society, Providence, RI, 1996. `doi:10.1090/ulect/008`.

[Stu97] B. Sturmfels. Equations defining toric varieties. In *Algebraic geometry—Santa Cruz 1995*, volume 62, Part 2 of *Proc. Sympos. Pure Math.*, pages 437–449. Amer. Math. Soc., Providence, RI, 1997. `doi:10.1090/pspum/062.2/1492542`.

[WCC11] J.-D. Wu, F.-J. Chen, and Y.-G. Chen. On the structure of the sumsets. *Discrete Math.*, 311(6):408–412, 2011. `doi:10.1016/j.disc.2010.11.014`.