ON A NEW ABSOLUTE VERSION OF SIEGEL'S LEMMA

MAXWELL FORST AND LENNY FUKSHANSKY

ABSTRACT. We establish a new version of Siegel's lemma over a number field k, providing a bound on the maximum of heights of basis vectors of a subspace of k^N, $N \geq 2$. In addition to the small-height property, the basis vectors we obtain satisfy certain sparsity condition. Further, we produce a nontrivial bound on the heights of all the possible subspaces generated by subcollections of these basis vectors. Our bounds are absolute in the sense that they do not depend on the field of definition. The main novelty of our method is that it uses only linear algebra and does not rely on the geometry of numbers or the Dirichlet box principle employed in the previous works on this subject.

1. INTRODUCTION AND STATEMENT OF RESULTS

The name Siegel's lemma is usually attributed to results on small-size nontrivial solutions to underdetermined systems of homogeneous linear equations over a field or a ring of arithmetic interest. The original Siegel's lemma asserts that given a nonzero integer $M \times N$ matrix A, $M < N$, there exists a nonzero $\boldsymbol{\xi} \in \mathbb{Z}^N$ such that $A\boldsymbol{\xi} = \mathbf{0}$ and

$$(1.1) \qquad |\boldsymbol{\xi}| \leq 1 + (N|A|)^{\frac{M}{N-M}},$$

where $|\cdot|$ here denotes the sup-norm (maximum of the absolute values of the coordinates) of the vector $\boldsymbol{\xi}$ and the matrix A, respectively. This result was originally used by Thue [21] and Siegel [17] to establish measures of irrationality for certain algebraic numbers, specifically to prove the existence of polynomials with small integer coefficients which vanish with prescribed multiplicity at given algebraic points. Since then, Siegel's lemma has found multiple applications in construction of auxiliary polynomials with integer coefficients that vanish at prescribed algebraic points, see the recent book [11] by D. W. Masser for details. The exponent $\frac{M}{N-M}$ in (1.1) is known to be best possible, however this bound lacks invariance under linear transformations: indeed, for any $M \times M$ integer matrix U,

$$(UA)\boldsymbol{\xi} = A\boldsymbol{\xi} = \mathbf{0},$$

however $|UA|$ and $|A|$ can be very different. It therefore makes sense to rephrase this fundamental principle as a result on the existence of points of bounded size in a vector space, i.e. the L-dimensional null-space of the given $N \times M$ linear system, where $L = N - M$. In this context, "size" is usually measured via a suitable *height function*, a standard tool of arithmetic geometry.

The first subspace version of Siegel's lemma over an arbitrary number field k, providing a full small-height basis, was established by Bombieri and Vaaler in [3].

A convenient for our purposes formulation, which easily follows from the results of [3], asserts that, given an L-dimensional subspace $\mathcal{Z}$ of k^N there exists a basis $\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_L$ for $\mathcal{Z}$ such that

$$(1.2) \qquad \prod_{i=1}^{L} H(\boldsymbol{\xi}_i) \leq N^{L/2} \left(\left(\frac{2}{\pi} \right)^{r_2} |\Delta_k| \right)^{\frac{L}{2d}} H(\mathcal{Z}),$$

where $d = [k : \mathbb{Q}]$, r_2 is the number of pairs of complex conjugate embeddings of k, Δ_k is the discriminant of k, and H is a height function (we review all the necessary notation including heights, as defined in [2] and [20], in Section 2).

Our goal is to establish a new version of Siegel's lemma with somewhat different bounds on the maximum of heights of the basis vectors, independent of the field of definition. Throughout this note, we assume that k is an algebraic number field and work in the k-linear space of $N \times 1$ column vectors k^N. More generally, we could also work in an intermediate field K such that $\mathbb{Q} \subseteq K \subseteq \overline{\mathbb{Q}}$ where it may happen that $K/\mathbb{Q}$ is an extension of infinite degree. However, we are interested in simultaneous solutions to finitely many linear equations having algebraic numbers as coefficients. Such systems involve only finitely many algebraic numbers and these generate a finite extension of $\mathbb{Q}$. Several inequalities in the literature (for example, [3, Theorem 8 and Theorem 9] and [20, Theorem 1, and Theorem 2]) that bound the height of solutions to simultaneous systems of linear equations with coefficients in a number field k contain constants that depend on k. Usually these constants depend on the discriminant of k, as in (1.2) above. An exception to this situation can be found in the striking results of Roy and Thunder [13], and [14], on absolute forms of Siegel's Lemma (a similar result follows from the celebrated paper of S. Zhang [22], however the Roy-Thunder notation is more convenient for our purposes). Analogously to (1.2), they prove the existence of a basis $\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_L$ for an L-dimensional subspace $\mathcal{Z} \subset \overline{\mathbb{Q}}^N$ with

$$(1.3) \qquad \prod_{i=1}^{L} H(\boldsymbol{\xi}_i) \leq \left(e^{\frac{L(L-1)}{4}} + \varepsilon \right) H(\mathcal{Z}),$$

for any $\varepsilon > 0$ (the choice of the basis depends on ε). While their bound does not depend on any number field, the vectors $\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_L$ are also not guaranteed to lie over a fixed number field.

Bridging between [3] and the work of Roy and Thunder, we establish the existence of a small-height basis for an L-dimensional subspace of k^N (i.e., the space of solutions to a system of simultaneous linear equations), and the inequalities we prove are free of constants that depend on a number field. While we bound the individual heights of the vectors instead of the product, our basis lies over a fixed number field k and our bound is particularly simple. In fact, we prove more than just the existence of a small-height basis for a subspace. A vector $\boldsymbol{x} \in k^N$ is called *s-sparse* for an integer $1 \leq s \leq N$ if $\boldsymbol{x}$ has no more than s nonzero coordinates. Here is our main result, which exhibits a small-height basis with additional sparsity and "monotonicity" properties.

Theorem 1.1. *Let $\mathcal{Z} = Ak^L \subseteq k^N$ be a subspace of dimension L where $1 \leq L < N$ and A is an $N \times L$ basis matrix for $\mathcal{Z}$. Let B be a full rank $L \times L$ submatrix of A, and write*

$$\left\{ \boldsymbol{\omega}_1, \ldots, \boldsymbol{\omega}_L \right\}$$

for the column vectors of the matrix AB^{-1}. Then $\boldsymbol{\omega}_1, \ldots, \boldsymbol{\omega}_L$ is another basis for $\mathcal{Z}$ over k, which consists of $(N-L+1)$-sparse vectors satisfying the following property: if $I \subseteq \{1, \ldots, L\}$ is a nonempty subset, and

$$(1.4) \qquad \mathcal{Y}_I = \mathrm{span}_k \left\{ \boldsymbol{\omega}_i : i \in I \right\}$$

is the k-linear subspace spanned by the basis vectors that are indexed by the elements in I, then

$$(1.5) \qquad H(\mathcal{Y}_I) \leq H(\mathcal{Z}).$$

In particular, $\max_{1 \leq i \leq L} H(\boldsymbol{\omega}_i) \leq H(\mathcal{Z})$. Moreover, if $I_1 \subsetneq I_2 \subseteq \{1, \ldots, L\}$, then

$$(1.6) \qquad H(\mathcal{Y}_{I_1}) \leq H(\mathcal{Y}_{I_2}).$$

Equality is attained in (1.6) if and only if $\boldsymbol{\omega}_i$ is a standard basis vector for each $i \in I_2 \setminus I_1$.

We illustrate Theorem 1.1 by Example 1 in Section 3. Observe that, while our result does not imply the bounds (1.2) and (1.3), those bounds do not imply our result either. Further, notice that in situations when height of the subspace $\mathcal{Z}$ is dominated by the constant depending on N and k in (1.2) or on L in (1.3) our bound may be better. In fact, the constant in (1.2) has been improved in [20] (we review this result in Section 2), but even this optimal constant depends on a power of Δ_k. Additionally, our bound can be preferable in some applications due to its simplicity. For example, in some situations it is important to obtain a point in a vector space $\mathcal{Z}$ outside of a co-dimension one subspace $\mathcal{Z}'$ (see [5] and references therein for a detailed discussion of Siegel's lemma-type results with avoidance conditions). It is clear that at least one of the basis vectors for $\mathcal{Z}$ is not in $\mathcal{Z}'$, however we do not know which vector it is. If we are to use the Bombieri-Vaaler basis satisfying (1.2), we can only assume that this vector's height is bounded by the right-hand side of that equation: the vector in question may be the largest with respect to height out of the basis, and the rest of the basis vectors may have height equal to 1. On the other, our bound guarantees that this vector has height $\leq H(\mathcal{Z})$, making it better in this situation.

It also makes sense to compare our Theorem 1.1 to a classical result of W. M. Schmidt, specifically Theorem 2(b) of [15]. Translating into our notation, Schmidt's theorem asserts that for any $0 \leq J < L \leq N$ our L-dimensional subspace $\mathcal{Z} \subseteq k^N$ contains a J-dimensional subspace $\mathcal{Y} \subset \mathcal{Z}$ with

$$H(\mathcal{Y}) \leq C H(\mathcal{Z})^{J/L}$$

for some constant C depending on N and the number field k. This inequality is similar to our equation (1.5), where in our result Schmidt's constant C disappears at the expense of a somewhat weaker upper bound for the height, as one might expect.

Our proof of this new form of Siegel's lemma does not use the Dirichlet box principle which was exploited in the earlier work of Baker [1], Roth [12], Siegel [17], and Thue [21]. Our approach also does not use methods from the geometry of numbers which were introduced in [3]; it is based solely on linear algebra. We review the necessary notation and background material in Section 2 and present the proof of Theorem 1.1 in Section 3. Additionally, we present a "relative" version of our Theorem 1.1 in Theorem 3.1: analogously to Theorem 12 of [3], we consider

a finite extension K of the number field k and produce a small-height basis over k for the null-space of a matrix defined over K.

One application of our Theorem 1.1 together with a previous version of Siegel's lemma and certain results on integer sensing matrices (see [7], [9], [10]) is the following observation.

Theorem 1.2. *Let $\mathcal{Z} \subseteq k^N$ be a subspace of dimension L where $1 \leq L < N$. For each integer $M > L$ there exists a collection of vectors*

$$S(M) = \{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_M\} \subset \mathcal{Z}$$

with the following properties:

  (1) *Every subcollection of L vectors from $S(M)$ forms a basis for $\mathcal{Z}$,*
  (2) *For every $\boldsymbol{y}_i \in S(M)$,*

$$H(\boldsymbol{y}_i) \leq L^{3/2}(2M)^{\frac{L-1}{L}} \min\left\{ H(\mathcal{Z})^L, \gamma_k(L)^{L/2} H(\mathcal{Z}) \right\},$$

where $\gamma_k(L)^{1/2}$ is the generalized Hermite's constant discussed in Section 2.

We review the necessary results on integer sensing matrices and prove Theorem 1.2 in Section 4. In the same section, we also prove a similar result for collections of smaller linearly independent sets of bounded-height vectors in a vector space $\mathcal{Z}$ (Proposition 4.1). Finally, we show in Corollary 4.2 that the basis guaranteed by Theorem 1.1 for a subspace spanned over k by rows of an integer sensing matrix must satisfy "strict monotonicity", i.e. inequalities (1.6) become strict. We are now ready to proceed.

2. Notation and heights

Let k be a number field of degree $d := [k : \mathbb{Q}] \geq 1$, as above, and let r_1, r_2 be the numbers of real and conjugate pairs of complex embeddings, respectively, so that

$$d = r_1 + 2r_2.$$

Let Δ_k be the discriminant of k, and write $M(k)$ for the set of places of k. For each $v \in M(k)$ let $d_v = [k_v : \mathbb{Q}_v]$ be the local degree, then for each $u \in M(\mathbb{Q})$, $\sum_{v|u} d_v = d$. We select the absolute values so that $|\ |_v$ extends the usual archimedean absolute value on $\mathbb{Q}$ when $v \mid \infty$, or the usual p-adic absolute value on $\mathbb{Q}$ when $v \nmid \infty$. With this choice, the product formula reads

$$\prod_{v \in M(k)} |\beta|_v^{d_v} = 1,$$

for each nonzero $\beta \in k$. Let $N \geq 2$, and for any place $v \in M(k)$ and $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_N) \in k^N$ define the corresponding sup-norm

$$|\boldsymbol{\beta}|_v = \max\{|\beta_1|_v, \ldots, |\beta_N|_v\}.$$

If $v \mid \infty$, we also define Euclidean norm

$$\|\boldsymbol{\beta}\|_v = \left(\sum_{i=1}^N |\beta_i|_v^2 \right)^{1/2}.$$

Then we can define the projective height function (also known as Arakelov height; see [2], p.66) $H : k^N \to \mathbb{R}_{\geq 0}$ as follows:

$$H(\boldsymbol{\beta}) = \left(\prod_{v \nmid \infty} |\boldsymbol{\beta}|_v^{d_v} \times \prod_{v | \infty} \|\boldsymbol{\beta}\|_v^{d_v} \right)^{1/d}.$$

This height is absolute, meaning that it is the same when computed over any number field k containing coordinates of $\boldsymbol{\beta}$: this is due to the normalizing exponent $1/d$ in the definition.

We often use positive integers L and M that satisfy

$$L + M = N.$$

We assume that A is an $M \times N$ matrix with entries in k and rank $A = M$. Then the null-space associated to A is the subspace

$$(2.1) \qquad \mathcal{Z} = \{ \boldsymbol{\beta} \in k^N : A\boldsymbol{\beta} = \mathbf{0} \}.$$

Obviously $\mathcal{Z} \subseteq k^N$ is a subspace of dimension L. If the vectors in the collection

$$(2.2) \qquad \{ \boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \ldots, \boldsymbol{\xi}_L \}$$

form a k-basis for $\mathcal{Z}$, we write

$$X = \begin{pmatrix} \boldsymbol{\xi}_1 \ \boldsymbol{\xi}_2 \ \cdots \ \boldsymbol{\xi}_L \end{pmatrix}$$

for the corresponding $N \times L$ matrix such that the basis vectors (2.2) are the columns of X. Then the Arakelov height of the subspace $\mathcal{Z}$ is also the Schmidt height (as defined in [3] or in [15]) of the matrix X. These heights are given by

$$(2.3) \qquad H(\mathcal{Z}) = H(X) = H\big(\boldsymbol{\xi}_1 \wedge \boldsymbol{\xi}_2 \wedge \cdots \wedge \boldsymbol{\xi}_L \big),$$

where $\wedge$ stands for the usual wedge product on vectors and the height on the right of (2.3) is defined on the associated Grassmannian as discussed in [2, section 2.8] or in [20, section 2]: it is simply the height H as above of the vector of Grassmann coordinates of $\mathcal{Z}$ identified with the vector space $k^{\binom{N}{L}}$. It follows using the product formula for the normalized absolute values $| \ |_v$, where v is a place of k, that the value of these heights does not depend on the choice of basis. In particular, the Arakelov height on vectors in k^N is well defined on the underlying projective space and a similar remark applies to the height on subspaces. From a duality principle formulated in [8] (see also [2], Corollary 2.8.12 on p.69 for a more modern account), we get the fundamental identity

$$(2.4) \qquad H(\mathcal{Z}) = H(X) = H(A),$$

where $H(A)$ is also the height of the M-dimensional subspace of k^N spanned by the columns of $A^\top$. By a standard convention, we set $H(k^N) = H(\{\mathbf{0}\}) = 1$.

If $1 \leq J < L$ then a generalization of Hadamard's determinant inequality (see [16], Lemma 5F on p.15 and Exercise 8a on p.29) asserts that

$$(2.5) \qquad \begin{aligned} H\big(\boldsymbol{\xi}_1 \wedge \boldsymbol{\xi}_2 \wedge \cdots \wedge \boldsymbol{\xi}_L \big) & \\ \leq H\big(\boldsymbol{\xi}_1 \wedge \boldsymbol{\xi}_2 \wedge \cdots \wedge \boldsymbol{\xi}_J \big) & H\big(\boldsymbol{\xi}_{J+1} \wedge \boldsymbol{\xi}_{J+2} \wedge \cdots \wedge \boldsymbol{\xi}_L \big). \end{aligned}$$

Alternatively, let Y and Z be the $N \times J$ and $N \times (L - J)$ matrices

$$Y = \begin{pmatrix} \boldsymbol{\xi}_1 \ \boldsymbol{\xi}_2 \ \cdots \ \boldsymbol{\xi}_J \end{pmatrix}, \quad \text{and} \quad Z = \begin{pmatrix} \boldsymbol{\xi}_{J+1} \ \boldsymbol{\xi}_{J+2} \ \cdots \ \boldsymbol{\xi}_L \end{pmatrix}.$$

Then (2.5) is also the inequality

$$(2.6) \qquad H(X) = H(Y\ Z) \le H(Y)H(Z)$$

for a partitioned matrix. By repeated application of this inequality we get

$$(2.7) \qquad H(\mathcal{Z}) = H(X) \le \prod_{\ell=1}^{L} H(\boldsymbol{\xi}_\ell).$$

In the special case $L = N$ we have

$$\det\big(\boldsymbol{\xi}_1\ \boldsymbol{\xi}_2\ \cdots\ \boldsymbol{\xi}_N\big) = \boldsymbol{\xi}_1 \wedge \boldsymbol{\xi}_2 \wedge \cdots \wedge \boldsymbol{\xi}_N,$$

and for each $v \mid \infty$,

$$(2.8) \qquad \big|\det\big(\boldsymbol{\xi}_1\ \boldsymbol{\xi}_2\ \cdots\ \boldsymbol{\xi}_N\big)\big|_v \le \prod_{n=1}^{N} \|\boldsymbol{\xi}_n\|_v,$$

which is Hadamard's upper bound for a determinant. As the origin of (2.5) and (2.6) is somewhat obscure, we will follow the terminology used in [4] and refer to all of these as *Hadamard's inequality*.

There is another type of inequality that is satisfied by the Arakelov height on subspaces. Let $\mathcal{Z} \subseteq k^N$ and $\mathcal{Y} \subseteq k^N$ be k-linear subspaces, and let $\langle \mathcal{Z}, \mathcal{Y} \rangle$ by the subspace spanned over k by $\mathcal{Z} \cup \mathcal{Y}$. Of course $\mathcal{Z} \cap \mathcal{Y}$ is also a subspace of k^N. Then these four subspaces satisfy the inequality

$$(2.9) \qquad H\big(\langle \mathcal{Z}, \mathcal{Y} \rangle\big) H(\mathcal{Z} \cap \mathcal{Y}) \le H(\mathcal{Z})H(\mathcal{Y}).$$

This was proved in [18, Theorem 1], and it was proved independently and at about the same time by W. M. Schmidt.

A natural question suggested by the inequality (2.7) is this: if $\mathcal{Z} \subseteq k^N$ is a subspace of dimension L, does there exist a basis for $\mathcal{Z}$ such that there is nearly equality in the inequality (2.7)? As explained in [3], an answer to this question is given by a result that is dual to Siegel's lemma. This was proved as [3, Theorem 8] using the Weil height on vectors. An analogous result for the Arakelov height was proved as [20, Theorem 2]. This later result asserts that there exists a basis $\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \ldots, \boldsymbol{\eta}_L$ for $\mathcal{Z}$ such that

$$(2.10) \qquad \prod_{\ell=1}^{L} H\big(\boldsymbol{\eta}_\ell\big) \le \gamma_k(L)^{L/2} H(\mathcal{Z}),$$

where $\gamma_k(L)^{1/2}$ is a positive constant that depends on the field k and the parameter L. The field constant $\gamma_k(L)^{1/2}$ is a generalization of Hermite's constant. It was first defined and used by J. Thunder in [19].

For each positive integer L and each archimedean place v of k we define positive real numbers

$$r_v(L) = \begin{cases} \pi^{-1/2}\{\Gamma(L/2+1)\}^{1/L} & \text{if } v \text{ is a real place} \\ (2\pi)^{-1/2}\{\Gamma(L+1)\}^{1/2L} & \text{if } v \text{ is a complex place,} \end{cases}$$

and then we define

$$c_k(L) = 2|\Delta_k|^{1/2d} \prod_{v\mid\infty} \{r_v(L)\}^{d_v/d},$$

where Δ_k is the discriminant of k. The constant $c_k(L)$ is the normalized Haar measure of a naturally occurring subset of the L-fold product of the adele ring $k_{\mathbb{A}}^L$,

and appears in applications of the geometry of numbers over $k_{\mathbb{A}}^L$. In particular, it leads to the simple upper bound for $\gamma_k(L)^{1/2}$ given by

$$(2.11) \qquad \gamma_k(L)^{1/2} \le c_k(L).$$

A more complicated lower bound was proved in [19, Theorem 2]. And it was shown in [20, Theorem 3] that the constant $\gamma_k(L)^{L/2}$ which appears in (2.10) is best possible for the inequality (2.10). From the lower bound proved in [19, Theorem 2] it follows that for fixed $d = [k : \mathbb{Q}]$ and fixed L, the value of $\gamma_k(L)^{1/2}$ grows like a small positive power of the absolute discriminant $|\Delta_k|$. In particular, for fixed $d = [k : \mathbb{Q}]$ and fixed L, there are only finitely many fields k such that $\gamma_k(L)^{1/2}$ is less than or equal to a positive real number.

If we arrange the basis vectors $\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \ldots, \boldsymbol{\eta}_L$ in (2.10) so that

$$H(\boldsymbol{\eta}_1) \le H(\boldsymbol{\eta}_2) \le \cdots \le H(\boldsymbol{\eta}_L),$$

we find that

$$H(\boldsymbol{\eta}_1) \le \gamma_k(L)^{1/2} H(\mathcal{Z})^{1/L}.$$

As the Arakelov height of a nonzero vector is always greater than or equal to 1, a trivial consequence of (2.10) is the inequality

$$(2.12) \qquad H(\boldsymbol{\eta}_L) \le \gamma_k(L)^{L/2} H(\mathcal{Z}).$$

However, if one works over the algebraic closure $\overline{\mathbb{Q}}$, then it follows from the results of Roy and Thunder [13], [14], that the dependence on the number field k can be removed. Suppose, for example, that $\mathcal{Z} \subseteq \overline{\mathbb{Q}}^N$ is a $\overline{\mathbb{Q}}$-linear subspace of dimension L. As a basis for $\mathcal{Z}$ requires only finitely many algebraic numbers as coordinates, there exists a number field k and a collection of basis vectors for $\mathcal{Z}$ such that the basis vectors belong to k^N. It follows that the Arakelov height of $\mathcal{Z}$ can be computed as before by using such a number field k. Because of the way we normalize the absolute values $| \ |_v$, where v is a place of k, it follows in a standard manner that the Arakelov height of $\mathcal{Z}$ does not depend on the choice of the field k. For such a subspace $\mathcal{Z} \subseteq \overline{\mathbb{Q}}^N$ and $\varepsilon > 0$, it follows from [14, Theorem 1] that there exists a basis

$$(2.13) \qquad \{\boldsymbol{\zeta}_1, \boldsymbol{\zeta}_2, \ldots, \boldsymbol{\zeta}_L\} \subseteq \mathcal{Z},$$

satisfying the inequality (1.3). Again the basis vectors (2.13) can be arranged so that

$$H(\boldsymbol{\zeta}_1) \le H(\boldsymbol{\zeta}_2) \le \cdots \le H(\boldsymbol{\zeta}_L).$$

Then it follows as before that

$$H(\boldsymbol{\zeta}_1) \le \left(e^{\frac{(L(L-1)}{4}} + \varepsilon\right)^{1/L} H(\mathcal{Z})^{1/L},$$

and

$$H(\boldsymbol{\zeta}_L) \le \left(e^{\frac{(L(L-1)}{4}} + \varepsilon\right) H(\mathcal{Z}).$$

Now, that we have reviewed the properties of heights and the previously known results on small-height bases for subspaces of k^N, we turn to the proof of our main theorem which complements the previously known bounds.

3. Proof of Theorem 1.1

Throughout this section, we let $\mathbf{1}_L$ denote an $L \times L$ identity matrix for $L > 0$ and $\mathbf{0}_{N \times L}$ denote an $N \times L$ matrix of all zeros for positive integers N, L. Additionally, let $[N] = \{1, \ldots, N\}$ and define

$$\mathcal{J}(N, L) = \{I \subseteq [N] : |I| = L\}.$$

For an $N \times L$ matrix A with coefficients in the number field k and for each $I \in \mathcal{J}(N, L)$ define A_I to be the $L \times L$ minor (i.e., $L \times L$ submatrix) of A whose rows are the rows of A indexed by I.

Lemma 3.1. *Let*

$$(3.1) \qquad\qquad A = \begin{pmatrix} \mathbf{1}_{L-1} & \mathbf{0}_{(L-1) \times 1} \\ \mathbf{0}_{1 \times (L-1)} & 1 \\ U & V \end{pmatrix}$$

be an $N \times L$ matrix with coefficients in k where U is an $(N - L) \times (L - 1)$ matrix and V is an $(N - L) \times 1$ matrix over k. Let

$$(3.2) \qquad\qquad A' = \begin{pmatrix} \mathbf{1}_{L-1} \\ \mathbf{0}_{1 \times (L-1)} \\ U \end{pmatrix}.$$

Then

$$(3.3) \qquad\qquad H(A) \geq H(A')$$

with equality if and only if $V = \mathbf{0}_{(N-L) \times 1}$.

Proof. We will first partition the Grassmann coordinates of A' into two sets. Define

$$R' = \{I' \in \mathcal{J}(N, L - 1) : L \notin I'\}$$

and

$$T' = \mathcal{J}(N, L - 1) \setminus R'.$$

Define

$$g(A'_{R'}) = (\det A'_{I'})_{I' \in R'}$$

to be the vector of Grassmann coordinates of A' indexed by R'. Likewise, define

$$g(A'_{T'}) = (\det A'_{I'})_{I' \in T'}$$

to be the vector of Grassmann coordinates of A' indexed by T'. Notice that the L-th row of A' is identically zero, so $\det(A'_{I'}) = 0$ when $L \in I'$. Therefore, for any $I' \in T'$, $\det A'_{I'} = 0$ and $g(A'_{T'}) = \mathbf{0}$. Let $(g(A'_{R'}), g(A'_{T'}))$ denote the vector formed by concatenating the vectors $g(A'_{R'})$ and $g(A'_{T'})$, so that $(g(A'_{R'}), g(A'_{T'}))$ is the vector of Grassmann coordinates of A' (up to a permutations of the Grassmann coordinates, which does not change the height). Since $g(A'_{T'})$ is identically zero and appending additional zeros to a vector does not affect the height, we have

$$H(A') = H(g(A'_{R'}), g(A'_{T'})) = H(g(A'_{R'})).$$

We will similarly partition the Grassmann coordinates of A. Define

$$R = \{I \in \mathcal{J}(N, L) : L \in I\},$$

$$S = \{\{1, 2, \ldots, L - 1, j\} : L + 1 \leq j \leq N\},$$

and

$$T = \mathcal{J}(N, L) \setminus (R \cup S).$$

Notice that for each $I' \in R'$ there exists $I \in R$ such that $I = I' \cup \{L\}$. Therefore, there is a one-to-one correspondence between elements of R' and elements R. In fact, there is also a one-to-one correspondence between the Grassmann coordinates of A' indexed by R' and the Grassmann coordinates of A indexed by R. Let $I'_1 = I' \cap \{1, \ldots, L\}$ be the set of indices of I' less than L and let $I'_2 = I' \cap \{L+1, \ldots, N\}$ be the set indices of I' greater then L, so that

$$A'_{I'} = \begin{pmatrix} A'_{I'_1} \\ A'_{I'_2} \end{pmatrix}.$$

Then A_I is of the form

$$A_I = \begin{pmatrix} A'_{I'_1} & \mathbf{0}_{|I'_1| \times 1} \\ \mathbf{0}_{1 \times (L-1)} & 1 \\ A'_{I'_2} & V \end{pmatrix},$$

and so

$$(3.4) \qquad \qquad \det A'_{I'} = \pm \det A_I.$$

Since multiplying the coordinates of a vector by ± 1 does not affect the height, (3.4) implies that

$$H(g(A'_{R'})) = H(g(A_R)),$$

where analogously $g(A_R) = (\det A_I)_{I \in R}$.

There is also a correspondence between the Grassmann coordinates of A indexed by S and the entries of V. For each $L+1 \le i \le N$ and each $I = \{1, \ldots, L-1, i\} \in S$, A_I is of the form

$$A_I = \begin{pmatrix} \mathbf{1}_{L-1} & \mathbf{0}_{(L-1) \times 1} \\ \boldsymbol{u}_{i-L} & V_{i-L} \end{pmatrix},$$

where $\boldsymbol{u}_{i-L}$ is the $(i-L)$-th row of U and V_{i-L} is the $(i-L)$-th coordinate of V. Thus $\det A_I = V_{i-L}$ and

$$H(g(A_S)) = H(V).$$

Let $(g(A_R), g(A_S), g(A_T))$ be the vector formed by concatenating $g(A_R), g(A_S)$ and $g(A_T)$ so that $H(A) = H(g(A_R), g(A_S), g(A_T))$. Then at each non-archimedian place $v \in M(k)$,

$$
\begin{aligned}
|(g(A_R), g(A_S), g(A_T))|_v &= \max\{|g(A_R)|_v, |g(A_S)|_v, |g(A_T)|_v\} \\
(3.5) &\ge |g(A_S)|_v = |g(A'_{S'})|_v,
\end{aligned}
$$

and at each archimedian place $v \in M(k)$,

$$
\begin{aligned}
\|(g(A_R), g(A_S), g(A_T))\|_v^2 &= \|g(A_R)\|_v^2 + \|g(A_S)\|_v^2 + \|g(A_T)\|_v^2 \\
(3.6) &\ge \|g(A_S)\|_v^2 = \|g(A'_{S'})\|_v^2,
\end{aligned}
$$

with equality in (3.6) if and only if $\|g(A_S)\|_v^2 = \|g(A_T)\|_v^2 = 0$. Combining (3.5) and (3.6) yields (3.3) where equality implies V is identically zero. As it is easy to verify that $H(A) = H(A')$ when $V = \mathbf{0}_{N-L}$, this completes the proof. $\qquad \square$

Corollary 3.2. *Let A be a full rank $N \times L$ matrix with $N > L$ and coefficients in k. Let $\boldsymbol{\omega}_1, \ldots, \boldsymbol{\omega}_L$ be the columns of A so that*

$$A = \begin{pmatrix} \boldsymbol{\omega}_1 & \ldots & \boldsymbol{\omega}_L \end{pmatrix},$$

and for some $1 \le i \le L$ let

$$A' = \begin{pmatrix} \boldsymbol{\omega}_1 & \ldots & \boldsymbol{\omega}_{i-1} & \boldsymbol{\omega}_{i+1} & \ldots & \boldsymbol{\omega}_L \end{pmatrix}$$

be the $N \times (L-1)$ matrix obtained by removing the i-th column of A. Let $I \in \mathcal{J}(N, L)$ and suppose that A_I is a permutation matrix. Then for any $1 \leq i \leq L$

$$H(A) \geq H(A')$$

with equality if and only if $\boldsymbol{\omega}_i$ is a standard basis vector.

Proof. Let Q be an $L \times L$ permutation matrix such that

$$AQ = \begin{pmatrix} \boldsymbol{\omega}_1 & \dots & \boldsymbol{\omega}_{i-1} & \boldsymbol{\omega}_{i+1} & \dots & \boldsymbol{\omega}_L & \boldsymbol{\omega}_i \end{pmatrix} = \begin{pmatrix} A' & \boldsymbol{\omega}_i \end{pmatrix}.$$

Multiplying by an $N \times N$ permutation matrix P_1 on the left we can permute the rows of A so that rows of A indexed by I are swapped with the rows of A indexed by $\{1, \dots, L\}$. Then $P_1 A Q$ is a matrix such that the first L rows form a permutation matrix. Multiplying again by an $N \times N$ permutation matrix P_2 on the left we can further permute the rows so that $P_2 P_1 A Q$ is a matrix so that the first L rows form an $L \times L$ identity matrix and thus $P_2 P_1 A Q$ is of the same form as the matrix in (3.1) and

$$P_2 P_1 A Q = \begin{pmatrix} P_2 P_1 A' & P_2 P_1 \boldsymbol{\omega}_i \end{pmatrix},$$

where $P_2 P_1 A'$ is a matrix of the same form as (3.2). Thus we can apply Lemma 3.1 to show that

$$(3.7) \qquad\qquad H(P_2 P_1 A Q) \geq H(P_2 P_1 A'),$$

with equality if and only if $\boldsymbol{\omega}_i$ is a standard basis vector. Recall that multiplying by a permutation matrix does not change the height of a matrix. Thus, (3.7) implies

$$H(P_2 P_1 A Q) = H(A) \geq H(A') = H(P_2 P_1 A'),$$

which completes the proof. □

We are now ready to prove Theorem 1.1.

Proof of Theorem 1.1. Let A be an $N \times L$ basis matrix for $\mathcal{Z}$ and let $J \in \mathcal{J}(N, L)$ so that A_J is a full rank minor of A. Let $X = A A_J^{-1}$ be another basis matrix for $\mathcal{Z}$ and let $\boldsymbol{\omega}_1, \dots, \boldsymbol{\omega}_L$ be the columns of X. By construction, X_J, the $L \times L$ minor of X indexed by J is a permutation matrix and thus we can apply Corollary 3.2 to X. In fact, for any $I \subseteq \{1, \dots, L\}$ define X^I to the matrix whose columns are the columns of X indexed by I, then X^I will be a matrix such that one of the $|I| \times |I|$ minors of X^I will be a permutation matrix. Therefore, for any $I_1 \subsetneq I_2 \subset \{1, \dots, L\}$ we can construct a sequence of subsets

$$I_1 = \mathcal{G}_1 \subsetneq \mathcal{G}_2 \subsetneq \dots \subsetneq \mathcal{G}_{|I_2 \setminus I_1|} = I_2,$$

where $|\mathcal{G}_{i+1} \setminus \mathcal{G}_i| = 1$ for each $1 \leq i < |I_2 \setminus I_1|$. Then for each $X^{\mathcal{G}_i}$ we can apply Corollary 3.2 to conclude that

$$H(X^{I_1}) = H(X^{\mathcal{G}_1}) \leq \dots \leq H(X^{\mathcal{G}_{|I_2 \setminus I_1|}}) = H(X^{I_2}),$$

with equality at each step if $\boldsymbol{\omega}_{\mathcal{G}_{i+1} \setminus \mathcal{G}_i}$ is a standard basis vector. Since $H(\mathcal{Y}_{I_1}) = H(X^{I_1})$ and $H(\mathcal{Y}_{I_2}) = H(X^{I_2})$ this implies (1.6) which in turn implies (1.5).

Finally, from the above argument, we see that the permutation matrix $X_J = A_J A_J^{-1}$, a submatrix of the basis matrix $X = A A_J^{-1}$ indexed by J, is in fact the $L \times L$ identity matrix. This means that each column of X has at least $L - 1$ zero coordinates. Since the vectors $\boldsymbol{\omega}_1, \dots, \boldsymbol{\omega}_L$ are the columns of X, the basis we constructed consists of $(N - L + 1)$-sparse vectors. This completes the proof of the theorem. □

Example 1. *We demonstrate Theorem 1.1 on a simple example. Let $k = \mathbb{Q}$, $N = 4$, $L = 3$, and take*

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 3 & 1 \\ 5 & 2 & 1 \\ 2 & 1 & 3 \end{pmatrix}.$$

The corresponding vector of Grassmann coordinates is $-18 \cdot (1, 1, 1, 1)$, and hence the height of the subspace $\mathcal{Z} = A\mathbb{Q}^3 \subset \mathbb{Q}^4$ is $H(\mathcal{Z}) = 2$. Take the indexing set $J = \{1, 2, 3\} \subset \mathcal{J}(4, 3)$ and consider the corresponding nonsingular minor

$$A_J = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 3 & 1 \\ 5 & 2 & 1 \end{pmatrix},$$

then we obtain the new basis matrix

$$X = AA_J^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & -1 & 1 \end{pmatrix}.$$

The column vectors $\boldsymbol{\omega}_1, \boldsymbol{\omega}_2, \boldsymbol{\omega}_3$ of X are 2-sparse and of height $\sqrt{2}$ each. Further, the 2-dimensional subspace spanned by any two of these column vectors has height $\sqrt{3}$. For instance, take

$$\mathcal{Y}_1 = \operatorname{span}_{\mathbb{Q}}\{\boldsymbol{\omega}_1\}, \ \mathcal{Y}_{1,2} = \operatorname{span}_{\mathbb{Q}}\{\boldsymbol{\omega}_1, \boldsymbol{\omega}_2\},$$

and observe that

$$H(\mathcal{Y}_{\{1\}}) = \sqrt{2} < H(\mathcal{Y}_{\{1,2\}}) = \sqrt{3} < H(\mathcal{Z}) = 2.$$

We can additionally prove a relative version of Theorem 1.1, analogous to Theorem 12 of [3]. Let K be a finite extension of the number field k with $[K : k] = r \geq 2$. Let F be a number field such that $k \subseteq K \subseteq F$, F is a Galois extension of k with Galois group $G(F/k)$ and F is a Galois extension of K with Galois group $G(F/K)$. Then $G(F/K)$ is a subgroup of $G(F/k)$ of index r. Let $\sigma_1, \ldots, \sigma_r \in G(F/k)$ be a set of distinct representatives of the cosets of $G(F/K)$ in $G(F/k)$. If A is an $M \times N$ matrix with coefficients a_{mn} in K for $1 \leq m \leq M, 1 \leq n \leq N$, define $\sigma_i(A) = (\sigma_i(a_{mn}))$. Then define

$$(3.8) \qquad \mathcal{A} = \begin{pmatrix} \sigma_1(A) \\ \vdots \\ \sigma_r(A) \end{pmatrix}.$$

Theorem 3.1. *Let F, K, k be number fields as above. Let A be an $M \times N$ matrix with $rM < N$ and coefficients in K and define $\mathcal{A}$ as above. Suppose that $\operatorname{rank}\mathcal{A} = rM$ and let $L = N - rM$. Let A_m denote the m-th row of A. Then $\mathcal{Z} := \ker(A) \cap k^N$ is an L-dimensional k-vector space so that*

$$(3.9) \qquad H(\mathcal{Z}) = H(\mathcal{A}) \leq \prod_{m=1}^{M} H(A_m)^r.$$

Moreover, there exists a basis

$$\{\boldsymbol{\omega}_1, \boldsymbol{\omega}_2, \ldots, \boldsymbol{\omega}_L\}$$

for $\mathcal{Z}$ with the following property: if $I \subseteq \{1, \ldots, L\}$ is a nonempty subset, and

$$\mathcal{Y}_I = \mathrm{span}_k \left\{ \boldsymbol{\omega}_i : i \in I \right\}$$

is the k-linear subspace spanned by the basis vectors indexed by I, then

$$H(\mathcal{Y}_I) \leq \prod_{m=1}^{M} H(A_m)^r.$$

In particular, $\max_{1 \leq i \leq L} H(\boldsymbol{\omega}_i) \leq \prod_{m=1}^{M} H(A_m)^r$. Moreover, if $I_1 \subsetneq I_2 \subseteq \{1, \ldots, L\}$, then

$$H(\mathcal{Y}_{I_1}) \leq H(\mathcal{Y}_{I_2}).$$

Proof. Let $\{\boldsymbol{\zeta}_1, \ldots, \boldsymbol{\zeta}_r\}$ be a k-basis for K. Let $A = (a_{mn})$ for $1 \leq m \leq M, 1 \leq n \leq N$. We can express each coefficient a_{mn} in terms of this basis as

$$a_{mn} = \sum_{j=1}^{r} a_{mn}^{(j)} \boldsymbol{\zeta}_j,$$

where $a_{mn}^{(j)} \in k$ for all m, n, j. For each $1 \leq j \leq r$, let $A^{(j)} = (a_{mn}^{(j)})$, $1 \leq m \leq M, 1 \leq n \leq N$ be an $M \times N$ matrix, and let

$$A' = \begin{pmatrix} A^{(1)} \\ \vdots \\ A^{(r)} \end{pmatrix}.$$

Then $\boldsymbol{x} \in k^N$ solves $A\boldsymbol{x} = \boldsymbol{0}$ if and only if it solves $A'\boldsymbol{x} = \boldsymbol{0}$, hence

$$V = \ker_k(A) = \ker_k(A').$$

Further, A' is a full-rank matrix with coefficients in k so

$$H(\mathcal{Z}) = H(A').$$

Now let

$$\Omega = (\sigma_i(\boldsymbol{\omega}_j)\mathbf{1}_M)_{\substack{i=1,\ldots,r \\ j=1,\ldots,r}}$$

be an $rM \times rM$ matrix organized into $M \times M$ blocks so that the (i, j) block is $\sigma_i(\boldsymbol{\omega}_j)\mathbf{1}_M$. Then

$$\Omega A' = \mathcal{A}.$$

Notice that Ω is a full-rank matrix, and so
(3.10)

$$H(\mathcal{A}) = H(\Omega A') = \left(\prod_{v \in M(K)} |\det(\Omega)|_v^{[K_v:\mathbb{Q}_v]} \right)^{\frac{1}{[K:\mathbb{Q}]}} H(A') = H(A') = H(\mathcal{Z}),$$

by the product formula. Writing $\mathcal{A}_i$ for the i-th row of $\mathcal{A}$, we see that (2.7) gives

$$(3.11) \qquad\qquad H(\mathcal{A}) \leq \prod_{i=1}^{rM} H(\mathcal{A}_i) = \prod_{i=1}^{M} \prod_{j=1}^{r} H(\sigma_j(A_i)).$$

Since $H(\sigma_j(A_i)) = H(A_i)$ combining (3.10) with (3.11) gives (3.9). We can now complete the proof by applying Theorem 1.1 to $\mathcal{A}$. $\qquad\square$

4. Proof of Theorem 1.2

To prove Theorem 1.2, let us first recall some results on the existence of integer sensing matrices. An $L \times M$ integer matrix A, $L < M$, is called an *integer sensing matrix for L-sparse signals* if every $L \times L$ submatrix of A is nonsingular. This notation comes from the study of sparse signal recovery in the compressive sensing signal processing paradigm: the defining condition for A ensures that for a vector $\boldsymbol{x} \in \mathbb{Z}^M$ with no more than L nonzero coordinates (L-sparse), $A\boldsymbol{x} = \boldsymbol{0}$ if and only if $\boldsymbol{x} = \boldsymbol{0}$, which allows for a unique recovery of the original sparse signal from its image under A. The problem of existence of such matrices $A = (a_{ij})_{1 \leq i \leq L, 1 \leq j \leq M}$ with bounded sup-norm

$$|A| := \max\{|a_{ij}| : 1 \leq i \leq L, 1 \leq j \leq M\}$$

and M as large as possible as a function of L and $|A|$ has been considered recently in [7] with further improvements in [9] and [10]. In particular, Theorem 2.2 of [7] establishes the existence of an $L \times M$ integer sensing matrix A for L-sparse signals with $|A| = T$ and

$$(4.1) \qquad\qquad M \geq CL\sqrt{T}$$

for an absolute constant C. Theorem 1.3 of [10] then establishes existence of such matrices with

$$(4.2) \qquad\qquad M \geq \max\left\{T + 1, \frac{T^{\frac{L}{L-1}}}{2}\right\},$$

which is an improvement on (4.1) when $L = o(\sqrt{T})$. Combining (4.1) with (4.2), we obtain an $L \times M$ integer matrix A with

$$(4.3) \qquad\qquad |A| = T \leq \min\left\{\left(\frac{M}{CL}\right)^2, (2M)^{\frac{L-1}{L}}, M - 1\right\},$$

so that any subcollection of L column vectors of A is linearly independent.

Proof of Theorem 1.2. Now, let the notation be as in the statement of Theorem 1.2 and let $\boldsymbol{\omega}_1, \ldots, \boldsymbol{\omega}_L$ be a basis for $\mathcal{Z}$, a specific choice will be made later. Dividing by a nonzero coordinate, if necessary, we can assume that each $\boldsymbol{\omega}_i$ has a coordinate equal to 1: this operation does not change the height of these basis vectors due to the product formula. This implies that for each $1 \leq i \leq L$,

$$(4.4) \qquad\qquad \prod_{v \in M(K)} |(1, \boldsymbol{\omega}_i)|_v^{\frac{d_v}{d}} = \prod_{v \in M(K)} |\boldsymbol{\omega}_i|_v^{\frac{d_v}{d}} \leq H(\boldsymbol{\omega}_i).$$

Write $W = (\boldsymbol{\omega}_1 \ \ldots \ \boldsymbol{\omega}_L)$ for the corresponding basis matrix and let $B = WA$, where A is an integer sensing matrix for L-sparse signals as described above. Let $S(M) = \{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_M\} \subset \mathcal{Z}$ be the set of column vectors of B. Then conclusion (1) of Theorem 1.2 are automatically satisfied, and we only need to prove (2). Notice that for each $1 \leq i \leq M$,

$$\boldsymbol{y}_i = \sum_{j=1}^{L} a_{ij} \boldsymbol{\omega}_j,$$

where all $a_{ij} \in \mathbb{Z}$ with $|a_{ij}| \leq T$. Then Lemma 2.1 of [5] asserts that

$$H(\boldsymbol{y}_i) \leq L^{3/2} \max_{1 \leq j \leq L} |a_{ij}| \prod_{v \in M(K)} |(1, \boldsymbol{\omega}_i)|_v^{\frac{d_v}{d}}.$$

Combining this observation with (4.4) implies that for each $1 \leq i \leq M$,

$$(4.5) \qquad\qquad H(\boldsymbol{y}_i) \leq L^{3/2} T \prod_{i=1}^{L} H(\boldsymbol{\omega}_i).$$

Making a choice of the basis $\boldsymbol{\omega}_1, \ldots, \boldsymbol{\omega}_L$ as in Theorem 1.1, we obtain a bound $H(\boldsymbol{y}_i) \leq L^{3/2} T H(\mathcal{Z})^L$. On the other hand, making a choice of the basis as in (2.10), we obtain a bound $H(\boldsymbol{y}_i) \leq L^{3/2} \gamma_k(L)^{L/2} T H(\mathcal{Z})$. The result of Theorem 1.2 now follows by combining these observations with (4.3), where we are choosing the bound $(2M)^{\frac{L-1}{L}}$ for T since its the smallest of the three for large M. $\qquad \square$

There are further results on $L \times M$ integer sensing matrix for s-sparse signals where $s < L$. Specifically, Theorem 1.1 of [6] asserts that for all sufficiently large L there exist $L \times M$ integer sensing matrices A for s-sparse signals with $1 \leq s \leq L-1$ such that $|A| = 2$ and

$$M \geq \left(\frac{L+2}{2} \right)^{1 + \frac{2}{3s-2}}.$$

Using the same reasoning as in our argument above with $T = 2$, we immediately obtain the following observation.

Proposition 4.1. *For sufficiently large integers $L < N$, any integer $1 \leq s \leq L-1$, and an L-dimensional subspace $\mathcal{Z} \subseteq k^N$, there exists a collection of vectors*

$$S = \{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_M\} \subset \mathcal{Z}$$

with the following properties:

(1) *For every $\boldsymbol{y}_i \in S$,*

$$H(\boldsymbol{y}_i) \leq 2L^{3/2} \min \left\{ H(\mathcal{Z})^L, \gamma_k(L)^{L/2} H(\mathcal{Z}) \right\},$$

   where $\gamma_k(L)^{1/2}$ is the generalized Hermite's constant discussed in Section 2,

(2) *The cardinality of the set S is*

$$M \geq \left(\frac{L+2}{2} \right)^{1 + \frac{2}{3s-2}},$$

(3) *Every subcollection of s vectors from S is linearly independent.*

Finally, we can show that the monotonicity property of the basis constructed in Theorem 1.1 becomes strict if the subspace in question is generated by rows of a sensing matrix.

Corollary 4.2. *Let A be an $L \times M$ integer sensing matrix for L-sparse signals with $L < M$. Let $\mathcal{Z} \subset k^M$ be the L-dimensional subspace spanned over the field k by the columns of $A^\top$ and let $\boldsymbol{\omega}_1, \ldots, \boldsymbol{\omega}_L$ be the basis for $\mathcal{Z}$ constructed in Theorem 1.1. Let $\mathcal{Y}_I$ be as in (1.4) for each $I \subseteq \{1, \ldots, L\}$. Then for every $I_1 \subsetneq I_2 \subset \{1, \ldots, L\}$*

$$H(\mathcal{Y}_{I_1}) < H(\mathcal{Y}_{I_2}).$$

Proof. As demonstrated in the proofs of Lemmas 3.1 and 3.2, if $H(\mathcal{Y}_{I_1}) = H(\mathcal{Y}_{I_2})$ then there must exist an $L \times L$ minor of $A^\top$ (and thus of A) of rank less than L. This contradicts the fact that A is an integer sensing matrix for L-sparse signals. $\square$

Conflict of interest statement: There are no conflicts of interest to be reported.

Data availability statement: Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

REFERENCES

[1] A. Baker, Linear forms in the logarithms of algebraic numbers, I, II, III, IV, *Mathematika*, 13 (1966), 204–216; 14 (1967), 102–107, 220–228; 15 (1968), 204–216.

[2] E. Bombieri and W. Gubler, *Heights in Diophantine Geometry*, Cambridge U. Press, New York, 2006.

[3] E. Bombieri and J. D. Vaaler, On Siegel's lemma, *Invent. Math.*, 73 (1983), 11–32.

[4] E. Fischer, Über den Hadamardschen Determinantensatz, *Arch. Math. (Basel)*, 13, (1908), 32–40.

[5] L. Fukshansky. Algebraic points of small height missing a union of varieties. *J. Number Theory*, 130(10):2099–2118, 2010.

[6] L. Fukshansky and A. Hsu. Covering point-sets with parallel hyperplanes and sparse signal recovery. *Discrete Comput. Geom.*, 2022, https://doi.org/10.1007/s00454-022-00375-y.

[7] L. Fukshansky, D. Needell, and B. Sudakov. An algebraic perspective on integer sparse recovery. *Appl. Math. Comput.*, 340:31–42, 2019.

[8] P. Gordan, Über einige Anwendungen diophantischer Approximationen, *Math. Annalen*, 7, (1873), 443–448.

[9] S. V. Konyagin. On the recovery of an integer vector from linear measurements. *Mat. Zametki*, 104(6):863–871, 2018.

[10] S. V. Konyagin and B. Sudakov. An extremal problem for integer sparse recovery. *Linear Algebra Appl.*, 586:1–6, 2020.

[11] D. W. Masser, *Auxiliary Polynomials in Number Theory*, Cambridge Tracts in Mathematics, 207. Cambridge U. Press, Cambridge, 2016

[12] K. F. Roth, Rational approximations to algebraic numbers, *Mathematika,* 2 (1955), 1–20; corrigendum, *ibid* 168.

[13] D. Roy and J. L. Thunder, An absolute Siegel's Lemma, *J. reine angew. Math.*, 476 (1996), 1–26.

[14] D. Roy and J. L. Thunder, Addendum and erratum to An absolute Siegel's Lemma, *J. reine angew. Math.*, 508 (1999), 47–51.

[15] W. M. Schmidt, On heights of algebraic subspaces and Diophantine approximations, *Ann. Math.*, 85, (1967), 430–472.

[16] W. M. Schmidt. Diophantine approximations and Diophantine equations. Lecture Notes in Mathematics, 1467. Springer-Verlag, Berlin, 1991.

[17] C. L. Siegel, Über einige Anwendungen diophantischer Approximationen, *Abh. der Preuß, Akad. der Wissenschaften.* Phys.-math. Kl. Nr. 1 (1929) (= *Ges. Abh.*, I, 209–266).

[18] T. Struppeck and J. D. Vaaler, Inequalities for heights on algebraic subspaces and the Thue-Siegel principle, *Analytic Number Theory, Proc. Conf. in honor of P. T. Bateman,* (B. C. Berndt, H. G. Diamond, H. Halberstam, A. Hildebrand, eds.), Birkhäuser, Boston, Mass., 1990.

[19] J. L. Thunder Higher dimensional analogues of Hermite's constant *Mich. Math. Jour.*, 45 (1998), 301–314

[20] J. D. Vaaler, The Best Constant in Siegel's Lemma, Monatsh. Math. 140, (2003), 71–89.

[21] A. Thue, Über Annäherungswerte algebraischer Zahlen, *J. reine angew. Math.*, 135, (1909), 284–305.

[22] S. Zhang, Positive line bundles on arithmetic varieties, *J. Amer. Math. Soc.*, 8(1), (1995), 187–221.

Institute of Mathematical Sciences, Claremont Graduate University, Claremont, CA 91711 USA

Email address: `maxwell.forst@cgu.edu`

Department of Mathematics, 850 Columbia Avenue, Claremont McKenna College, Claremont, CA 91711 USA

Email address: `lenny@cmc.edu`