BOUNDEDNESS THEOREMS FOR THE RELAXATION METHOD

EDOARDO AMALDI* AND RAPHAEL HAUSER†

Abstract. A classical theorem by Block and Levin says that certain variants of the relaxation method for solving systems of linear inequalities produce bounded sequences of intermediate solutions even when running on inconsistent input data. Using a new approach, we prove a more general version of this result and answer an old open problem of quantifying the bounds as a function of the input data.

AMS subject classifications. Primary 15A39,05A20, 52B55. Secondary 90C05, 62M45,68T01.

Key words. Relaxation method, perceptron boundedness, condition number, linear inequalities.

1. Introduction. The relaxation method [1, 12] for finding a point in an open convex set $\mathcal{K}$ of a real inner product space $(V, \langle\cdot,\cdot\rangle)$ is a conceptual algorithm based on a separation oracle: given $x_i \in V$, the oracle either provides $a_i \in V, b_i \in \mathbb{R}$ such that

$$\langle a_i, x\rangle > b_i \qquad \forall x \in C, \tag{1.1}$$

$$\langle a_i, x_i\rangle \leq b_i, \tag{1.2}$$

or else a certificate that $x_i \in C$. Given a starting point x_0, the method proceeds by updating x_i to $x_{i+1} = x_i + \eta_i a_i$ in the first case and by stopping when a certificate of membership is found. Sometimes $\mathcal{K}$ is assumed closed, in which case $>$ is replaced by $\geq$ in (1.1) and $\leq$ by $<$ in (1.2), but assuming that $\mathcal{K}$ is open is more convenient for our purposes.

In situations where $\mathcal{K}$ is implicitly described as the intersection of a set of half-planes

$$\mathcal{K} = \bigcap_{(a,b)\in A} \{x \in V : \langle a, x\rangle > b\} \tag{1.3}$$

the relaxation idea can usually be turned into a practical algorithm by constructing an oracle that samples the linear inequalities in A. In particular, the polyhedra that play a role in linear programming are usually described by linear inequalities. In radiation therapy planning a large number of linear inequalities occur naturally in a mathematical formulation of the problem of delivering a maximum dose of X-rays inside a tumour subject to the constraint of keeping the exposure of surrounding tissue below a threshold [5]. The situation is similar in certain approaches to neural networks based on finding linear classifiers that should correctly classify a set of training points [14].

Many different versions of the relaxation method have been proposed, depending on how the step size η_i and the points (a_i, b_i) defining the separating plane are chosen. The perceptron algorithm [14] is based on choosing $\eta_i \equiv 1$, the cyclic projection algorithm [3] checks all the inequalities in A in a cyclic manner and takes a step

*Dipartimento di Elettronica e Informazione, Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133 Milano, Italy, (amaldi@elet.polimi.it).

†Oxford University Computing Laboratory, Wolfson Building, Parks Road, Oxford, OX1 3QD, United Kingdom, (hauser@comlab.ox.ac.uk). This author was supported through grant NAL/00720/G from the Nuffield Foundation and through grant GR/M30975 from the Engineering and Physical Sciences Research Council of the UK.

length $\eta_i = (b_i - \langle a_i, x_i\rangle)/\langle a_i, a_i\rangle$. Also related are Cimmino's method [6], the random Bregman projection method [2] and various types of subgradient algorithms [15] such as the bundle method [9].

When the relaxation method is applied to a finite feasible data set A (one that leads to $\mathcal{K} \neq \emptyset$), then x_i becomes a member of $\mathcal{K}$ after a finite number of iterations for all interesting variants. A typical convergence result is that of the perceptron method which finds a feasible point in a number of iterations that is proportional to the square of Goffin's condition number, see [8]. When the input data are rational, the worst-case number of iterations needed til termination is exponential in the input size. The relaxation method can be greatly improved by rescaling the underlying space at each iteration [16]. This insight led to the development of the ellipsoid method [17] which became the first algorithm for linear programming shown to terminate in polynomial time [10].

A special situation occurs when $\mathcal{K} = \emptyset$ but $A \neq \emptyset$ (we say that A is an infeasible problem instance): in this case the algorithm will never stop, as the oracle will always find a violated inequality. In this paper we are going to investigate this situation for the case where A is a finite set. Although the problem of finding a point in the empty set is an impossible undertaking, the application of the relaxation method to infeasible input data is interesting in contexts where one wishes to satisfy as large a number of inequalities as possible but where satisfying all of them is usually impossible, as is typically the case in radiation therapy planning, computer learning and neural networks. In this situation one would usually like to have information about the spatial distribution of the points x_i generated by the algorithm. Our paper answers a question of this type by showing that for an important family of relaxation methods the sequence $(x_i)_{\mathbb{N}_0}$ remains bounded within a ball of quantifiable radius.

We will now describe the technical setup of our problem in some detail. Let $A \subset \mathbb{R}^n \setminus \{0\}$ be a finite set of nonzero vectors, $x_0 \subset \mathbb{R}^n$ a given starting point, $(\eta_i)_{\mathbb{N}_0} \in \mathbb{R}_+$ a sequence of nonnegative numbers and $(a_i)_{\mathbb{N}_0} \subseteq A$ an infinite sequence of vectors from A. Suppose that the sequence of points defined iteratively by $x_{i+1} = x_i + \eta_i a_i$ satisfies the condition

$$a_i^{\mathrm{T}} x_i \leq 0 \tag{1.4}$$

for all $i \in \mathbb{N}_0$. This situation occurs when the relaxation method with step lengths η_i is applied to an infeasible system of homogeneous linear inequalities $A^{\mathrm{T}}x > 0$. Here, as elsewhere below, we identify the set A with the matrix obtained by collecting its elements as column vectors in some order.

Nilson and Beyer (see references in [7]) conjectured that in the case where $\eta_i \equiv 1$ (which corresponds to the perceptron algorithm) the sequence of iterates $(x_i)_{\mathbb{N}_0}$ remains bounded in norm. A proof of this conjecture, known today as the Perceptron Boundedness Theorem, was developed in a series of papers by Effron [7], Minsky–Papert [11] and Block–Levin [4]. Their proof is based on the notion of *strict F-chains* as a basic analytic tool to enable an induction over the problem dimension. Using the same technique, it can easily be established that the result extends to the case where there exist constants L, U such that

$$0 < L < \eta_i < U \qquad \forall i \in \mathbb{N}_0. \tag{1.5}$$

How general are these assumptions? Clearly, in the absence of an upper bound on the η_i, $(x_i)_{\mathbb{N}_0}$ does not have to remain bounded. Hence, the upper bound is necessary.

On the other hand, the requirement that the lower bound be strictly positive may seem like a minor technical detail, and one is led to suspect that it should be possible to replace (1.5) by the less restrictive condition

$$0 \leq \eta_i < U \qquad \forall i \in \mathbb{N}_0 . \tag{1.6}$$

Indeed, proving perceptron boundedness under the more general condition (1.6) is one of the main goals of this paper, see Theorem 6.1 below.

Unfortunately, the strict positivity of the lower bound in (1.5) is a key requirement in arguments based on F-chains. A F-chain free approach is thus needed to establish the result under the more general assumption (1.6). Our proof of Theorem 6.1 is based on such an approach, which is not only conceptually new and less restrictive but also easier to understand geometrically than F-chains. Moreover, while earlier proofs showed only the existence of a bound on $(x_i)_{\mathbb{N}}$, our Theorem 6.1 explicitly quantifies the bound as a function of the input data A. This solves an old open problem which was first mentioned by Block and Levin [4].

Both the fact that one can relax the lower bound on the step size and the availability of an explicit expression for the bound as a function of the input data provide interesting mathematical tools. In a forthcoming paper we will show how both facts play a key role in the analysis of convergence rates for randomised relaxation schemes.

In Section 7 we further generalise the boundedness theorem to include inhomogeneous inequalities, that is, we show that relaxation methods with step lengths obeying (1.6) and applied to systems $A^{\mathrm{T}}x > b$ produce a bounded sequence of intermediate solutions $(x_0)_{\mathbb{N}}$, with a bound that can be quantified as a function of (A, b).

A few remarks about notation: we endow $\mathbb{R}^n$ with the canonical Euclidean product, denoted by the $\cdot$ product, and we write $\|\cdot\|$ for the associated norm. Any other norms used are explicitly constructed as the gauge functions associated with convex, balanced, absorbing, bounded sets (see [13]) and are denoted by ξ. If $x \in \mathbb{R}^n$ is a nonzero vector, then $\hat{x} = x/\|x\|$ denotes its normalisation. If $A = (a^{[1]}, \ldots, a^{[k]})$ is an ordered set of vectors in $\mathbb{R}^n$, then we identify A with the matrix obtained by collecting its elements as column vectors in the same order

$$(a^{[1]}, \ldots, a^{[k]}) \equiv \begin{pmatrix} a^{[1]} & \ldots & a^{[k]} \end{pmatrix}.$$

Any subset $B \subset A$ inherits the same order of elements and is identified with the corresponding submatrix of A. When $A = \{a^{[1]}, \ldots, a^{[k]}\}$ is an unordered set then the same identification holds modulo ordering of the elements, that is, we think of A as a matrix obtained by collecting its elements as column vectors in an arbitrary order.

2. The Underlying Geometric Insight. Before analysing the general case, it is helpful to illustrate the geometric ideas that drive our approach in $\mathbb{R}^2$ and $\mathbb{R}^3$. From here on and until the beginning of Section 7 we consider homogeneous systems $A^{\mathrm{T}}x > 0$. In this and a few other sections we will assume that $\|a\| = 1$ for all $a \in A$. This implies no loss of generality, since we can replace η_i by $\eta_i\|a_i\|$, U by $U \cdot \max_{a \in A} \|a\|$ and a_i by $a_i/\|a_i\|$.

We start by discussing the case $n = 2$. Let $\overline{B_1}(0)$ be the closed unit disk and $S^1 = \partial\overline{B_1}(0)$ the unit circle around the origin. We will remove the set

$$\mathcal{M} := \{x \in S^1 : \exists a \in A \text{ s.t. } a \perp x\}$$

from $\overline{B_1}(0)$ via a finite number of convex cuts. For $a \in A$ and $r \geq 0$ let

$$C_{\{a\}^\perp}(r) := \{x \in \mathbb{R}^2 : \|\pi_{\{a\}^\perp} x\| \leq r\},$$

where $\pi_{\{a\}^\perp} x = x - (a \cdot x)a$ is the projection of x into $\mathrm{span}(\{a\})^\perp = \{y \in \mathbb{R}^2 : a \cdot y = 0\}$ along a. $C_{\{a\}^\perp}(r)$ is thus a sandwich of thickness $2r$ parallel to a. We adopt this slightly cumbersome notation with later generalisations in mind. Since $\mathcal{M}$ consists

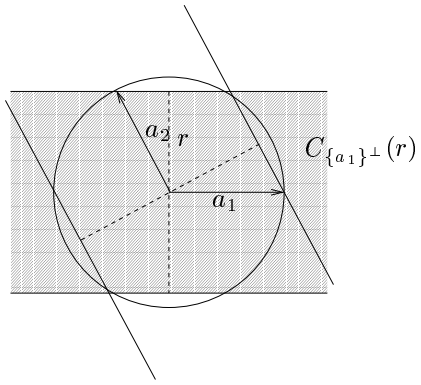

FIG. 2.1. *Geometric interpretation of* $C_{\{a_1\}^\perp}(r)$

of finitely many isolated points on S^1, it is possible to choose radii $r_{\{a\}} \in (0,1)$ sufficiently close to 1 for all $a \in A$ such that

$$\overline{B_1}(0) \cap C^c_{\{a\}^\perp}(r_{\{a\}^\perp}) \cap C^c_{\{\tilde{a}\}^\perp}(r_{\{\tilde{a}\}^\perp}) = \emptyset \quad \forall a, \tilde{a} \in A \text{ with } a \neq \tilde{a}, \tag{2.1}$$

see Figure 2.1. For example, it suffices to choose $r_{\{a\}^\perp} = \cos\theta/2$ for all $a \in A$, where

$$\cos\theta = \max\{|a \cdot \tilde{a}| : a, \tilde{a} \in A, a \neq \tilde{a}\}.$$

Let us now consider the set

$$C = \overline{B_1}(0) \cap \bigcap_{a \in A} C_{\{a\}^\perp}(r_{\{a\}^\perp})$$

obtained from the closed unit ball by removing $\mathcal{M}$ via the convex cuts defined by the sandwiches $C_{\{a\}^\perp}(r_{\{a\}^\perp})$, see Figure 2.2. Note that

$$r_{\{a\}} := \left(1 - r^2_{\{a\}^\perp}\right)^{\frac{1}{2}} > 0.$$

Equation (2.1) implies that there exists a number $\delta = \min\{r_{\{a\}} : a \in A\} > 0$ such that

$$a \in A, x \in C, a^{\mathrm{T}} x \leq 0, 0 \leq \eta \leq \delta \Rightarrow x + \eta a \in C. \tag{2.2}$$

We will prove this claim in the more general setting of Lemma 5.4. For the time being, let us convince that this property holds in $\mathbb{R}^2$ by inspection of Figure 2.2.

The set C is bounded, closed and convex. It is also balanced, that is, $x \in C \Leftrightarrow -x \in C$, and absorbing, that is, for every $x \in \mathbb{R}^2$ there exists a $\lambda \geq 0$ such that $x \in \lambda C$. Combined together, these properties imply that the *gauge function*

$$\xi : x \mapsto \min\{\lambda \geq 0 : x \in \lambda C\}$$

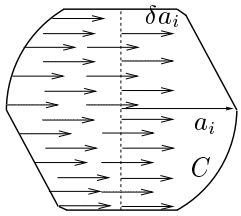

FIG. 2.2. *Moving by δ along a_i in the unit ball of the gauge*

associated with C is a norm on $\mathbb{R}^2$, see for example [13]. Relation (2.2) implies that whenever $\xi(x_i) > U/\delta$ we have $\xi(x_{i+1}) \leq \xi(x_i)$. Therefore, the gauge cannot grow larger than $O(U/\delta)$. Since all norms on $\mathbb{R}^2$ are equivalent, and since the sequence $(x_i)_{\mathbb{N}}$ is bounded in the norm ξ, it is also bounded in the Euclidean norm.

Why could our proof not have worked with the Euclidean norm in place of ξ and the unit disk $\overline{B_1}(0)$ in lieu of C? Note that with these replacements, for all $x \notin \mathcal{M}$ the key identity (2.2) can be made to hold as long as δ is small enough. This is clear when x is an interior point of the disk, and for $x \in \partial\overline{B_1}(0)$, the condition $a^{\mathrm{T}}x < 0$ implies that the line $x + ta$ cuts into the interior of $\overline{B_1}(0)$ for small positive t. The problem is that now δ is a function of x and goes to zero near points $x \in \mathcal{M}$. Thus, the bound U/δ is arbitrarily large and the argument breaks down.

Of course, we intend to generalise the construction of the gauge ξ and its associated closed unit ball C to higher dimensions, and then extra cuts are needed because $\mathcal{M}$ no longer consists of isolated points. In $\mathbb{R}^3$ for example, this set is obtained as a union of grand circles

$$\mathcal{M} = \bigcup_{a \in A} \{x \in S^2 : x \cdot a = 0\},$$

We define $C_{\{a\}^\perp}(r_{\{a\}^\perp})$ as before, although $\{a\}^\perp$ is now the orthogonal complement of $\mathrm{span}(\{a\})$ in $\mathbb{R}^3$, $\overline{B_1}(0)$ is the closed Euclidean unit ball in $\mathbb{R}^3$ and $S^2 = \partial\overline{B_1}(0)$ is the unit sphere, see Figure 2.3.

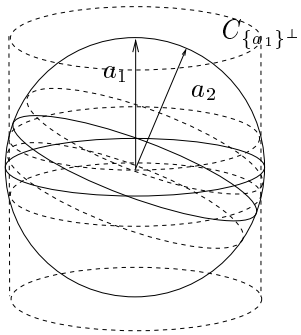

FIG. 2.3. *Cuts in $\mathbb{R}^3$*

If we want (2.2) to be satisfied for two distinct vectors $a, \tilde{a} \in A$, problems occur near intersection points of the two grand circles associated with a and $\tilde{a}$: as we saw above, the purpose of the cuts $C_{\{a\}^\perp}(r_{\{a\}^\perp})$ is to flatten the gauge unit ball C in the direction of a, so that it is possible to move by a finite amount in the direction of

a from any point within C where the inequality $a^{\mathrm{T}}x > 0$ is violated. However, in a neighbourhood of points $x \in \operatorname{span}(\{a, \tilde{a}\})^{\perp} \cap \partial C$, the flattening produced by the cut $C_{\{a\}^{\perp}}(r_{\{a\}^{\perp}})$ is destroyed by the cut $C_{\{\tilde{a}\}^{\perp}}(r_{\{\tilde{a}\}^{\perp}})$ and vice versa, see the "diamond" shaped areas in Figure 2.4.

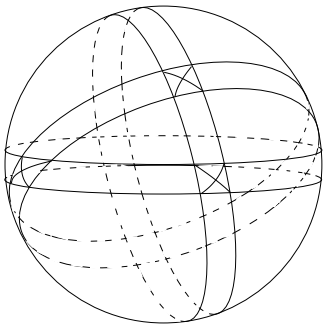

FIG. 2.4. *Interference of cuts in diamond shaped region*

The problem can be fixed by applying further cuts of the form

$$C_{\{a,\bar{a}\}^{\perp}}(r_{\{a,\bar{a}\}^{\perp}}) := \left\{x \in \mathbb{R}^3 : \|\pi_{\{a,\bar{a}\}^{\perp}} x\| \leq r_{\{a,\bar{a}\}^{\perp}}\right\},$$

where $\pi_{\{a,\bar{a}\}^{\perp}} x$ is the projection of x into $\operatorname{span}(a, \tilde{a})^{\perp}$ along $\operatorname{span}(a, \tilde{a})$, which implies that $C_{\{a,\bar{a}\}^{\perp}}(r_{\{a,\bar{a}\}^{\perp}})$ is a sandwich parallel to the hyperplane $\operatorname{span}(a, \tilde{a})$. Note that if a and $\tilde{a}$ are almost collinear, then the diamond shaped region becomes very elongated and potentially large, unless $r_{\{a\}^{\perp}}$ and $r_{\{\tilde{a}\}^{\perp}}$ are chosen close to 1. Moreover, $r_{\{a,\bar{a}\}}$ must be chosen small enough for $C_{\{a,\bar{a}\}^{\perp}}(r_{\{a,\bar{a}\}^{\perp}})$ to cut off the diamond shaped region, but large enough so that it does not interfere with the flattening produced by other cuts of the same type. Finally, three different vectors $a, \tilde{a}, \bar{a} \in A$ could be linearly dependent, and then the three corresponding grand circles intersect in two points. This situation does not create the need for introducing a new type of cut, but it does generate additional constraints on the cut-depths.

In $\mathbb{R}^n$ the situation is further exacerbated by the fact that $n-1$ different types of cuts will be needed, one type each for intersections of $k = 1, \ldots, n-1$ of the grand circles $\{x \in S^{n-1} : a \cdot x = 0\}$, $(a \in A)$ corresponding to k linearly independent subsets of A, where $S^{n-1} = \partial\overline{B_1}(0)$ is the $(n-1)$-dimensional unit sphere and $\overline{B_1}(0)$ the n-dimensional Euclidean closed unit ball. There exists a complicated combinatorial set of constraints linking the depth of the different cuts, and a quantitative expression for the radii depends on the angles between subspaces $\operatorname{span}(P)$ and $\operatorname{span}(Q)$ for different subsets $P, Q \subset A$. Moreover, we need an adequate generalisation of (2.1) to be able to guarantee that (2.2) holds.

Luckily, there exists a simple algebraic condition on the radii that involves only very few of the exponentially many cut-depth constraints and guarantees that the remaining constraints are satisfied, see (4.2). Moreover, the dependence of the radii on the angles between the exponentially many subspaces spanned by the vectors of A can be captured in a single condition number $\kappa(A)$ which will be developed in Section 3.

3. A Condition Number to Quantify the Bounds. Let us now attack the problem in $\mathbb{R}^n$ and generalise some of the concepts introduced above. We start with an object that will allow us to quantify the bounds in our main theorems.

Let A be a finite ordered set of vectors in $\mathbb{R}^n \setminus \{0\}$. As mentioned earlier, we identify A and any of its subsets with the matrices obtained by collecting their elements as column vectors in the specified order. Let $m := \operatorname{rank}(A)$, $B \subseteq A$ and $k := \#B$. Let $B = Q(B)R(B)$ be the thin QR decomposition of B. That is, $Q(B) \in \mathbb{R}^{n\times k}$ satisfies $Q(B)^{\mathrm{T}}Q(B) = \mathrm{I}_k$ (it is a Stiefel matrix), and $R(B) \in \mathbb{R}^{k\times k}$ is an upper triangular matrix with strictly positive diagonal, $\operatorname{diag}(R(B)) > 0$. We write

$$|\det_k(B)| := \det(R(B)).$$

Note that because of the way $R(B)$ was chosen, this number is always positive. We endow all subspaces V of $\mathbb{R}^n$ with the restriction of the canonical Euclidean inner product of $\mathbb{R}^n$ to V. With this convention in place it is the case that $|\det_k(B)|$ expresses the standard k-dimensional volume of the parallelepiped

$$\Big\{\sum_{i=1}^{k} \eta_i b_i : \eta_i \in [0,1](i = 1, \dots, k)\Big\} \subset \operatorname{span}(B),$$

where $B = (b_1, \dots, b_k)$. We define

$$\kappa(A)^{-1} := \min\big\{|\det_k(B)| : \ (k = 1, \dots, m), B \subseteq A, \#B = k, |\det_k(B)| \neq 0\big\}. \tag{3.1}$$

Lemma 3.1 sums up a few simple properties of $\kappa(A)$. Part iv) and its proof are inspired by the localisation lemma for the ellipsoid method [10]. It provides an efficiently computable upper bound on $\kappa(A)$ when A consists of rational input data. This is useful since a direct computation of $\kappa(A)$ via (3.1) is generally impractical. Let us explain how the bit-length $\mathscr{D}$ that appears in part iv) is defined: every rational number $r = \pm p/q$ can be represented as a triplet (p, q, s), where p and q are mutually coprime natural numbers and $s = \pm 1$ is the sign of r. We assume that p and q are expressed in the shortest possible binary expansion, that is, if the string is nonempty then the leading digit is a 1. The empty string represents the number 0. Finally, s can be represented by a single bit of information. Since $q \neq 0$, the shortest possible representation of a number takes two digits, corresponding to $r = 0$ represented by $(\emptyset, 1, 0)$. The bit-length $\mathscr{D}$ is then defined as the sum of the bit-lengths of all components appearing in the vectors of A.

Lemma 3.1. *Let $A \subset \mathbb{R}^n \setminus \{0\}$ be a finite ordered set of* $\operatorname{rank}(A) = m \leq n$. *Then the following statements hold true:*

i) *$\kappa(A)$ is invariant under reordering of the elements of A.*

ii) *If $B = \begin{pmatrix} B_1 & B_2 \end{pmatrix} \subseteq \hat{A}$ is of rank m where $B_1 = \begin{pmatrix} \hat{a}_1 & \dots & \hat{a}_{k-1} \end{pmatrix}$ and $B_2 = \begin{pmatrix} \hat{a}_k & \dots & \hat{a}_m \end{pmatrix}$, and if $\bar{B} = \begin{pmatrix} F & B_2 \end{pmatrix}$ where $F = \begin{pmatrix} f_1 & \dots & f_{k-1} \end{pmatrix}$ is an orthonormal basis of* $\operatorname{span}(B_1)$, *then*

$$\big|\det_{m-k+1}(B_2)\big| \geq \big|\det_m(\bar{B})\big| \geq \big|\det_m(B)\big|.$$

In particular, this implies that only matrices B of rank m need to be considered in the definition of $\kappa(A)$.

iii) *If $\hat{A} := \{\hat{a} : a \in A\}$, where $\hat{a} = a/\|a\|$, then*

$$\kappa(\hat{A}) \leq \kappa(A) \cdot \min_{a\in A} \|a\|^{-m}.$$

iv) if A consists of rational input vectors of total bit length $\mathscr{D}$ then

$$\kappa(A) \leq 2^{\mathscr{D}}.$$

Proof. i) Let $B \in \mathbb{R}^{n\times k}$ be a rank k matrix, let $P \in \mathcal{S}_k$ be a $k \times k$ permutation matrix and $\tilde{B} = BP$. All we need to show is that $|\det_k B| = |\det_k \tilde{B}|$. Let $G = \begin{pmatrix} B & G_2 \end{pmatrix} \in \mathbb{R}^{n\times n}$ be such that G_2 has column vectors consisting of an orthonormal basis of $\operatorname{span}(B)^{\perp}$. Then the QR decomposition of G is $\begin{pmatrix} Q_1 & G_2 \end{pmatrix}\begin{pmatrix} R_{11} & 0 \\ 0 & \mathrm{I} \end{pmatrix}$, where $Q_1 R_{11}$ is the thin QR decomposition of B. Likewise, the matrix $\tilde{G} := G\begin{pmatrix} P & 0 \\ 0 & \mathrm{I} \end{pmatrix} = \begin{pmatrix} \bar{B} & G_2 \end{pmatrix}$ has QR decomposition $\tilde{G} = \begin{pmatrix} \bar{Q}_1 & G_2 \end{pmatrix}\begin{pmatrix} \bar{R}_{11} & 0 \\ 0 & \mathrm{I} \end{pmatrix}$, where $\tilde{Q}_1 \tilde{R}_{11}$ is the thin QR decomposition of $\tilde{B}$. Therefore,

$$|\det_k \tilde{B}| = \det \tilde{R}_{11} = |\det \tilde{G}| = |\det G| = \det R_{11} = |\det_k B|.$$

ii) Let $\begin{pmatrix} Q_1 & Q_2 \end{pmatrix}\begin{pmatrix} R_{11} & R_{12} \\ 0 & R_{22} \end{pmatrix}$ and $\begin{pmatrix} \bar{Q}_1 & \bar{Q}_2 \end{pmatrix}\begin{pmatrix} \bar{R}_{11} & \bar{R}_{12} \\ 0 & \bar{R}_{22} \end{pmatrix}$ be the thin QR decompositions of B and $\bar{B}$ respectively, where the blocking is consistent with $\begin{pmatrix} B_1 & B_2 \end{pmatrix}$. Then $Q_2 = \bar{Q}_2$, $R_{22} = \bar{R}_{22}$, and $\bar{R}_{11} = \mathrm{I}$. Moreover, the i-th column of R_{11} is a vector of length $\|\hat{a}_i\| = 1$, which implies that all diagonal elements of R_{11} are ≤ 1. The second inequality now follows from

$$\left|\det_m(\bar{B})\right| = \det(\bar{R}_{11})\det(\bar{R}_{22}) \geq \det(R_{11})\det(R_{22}) = \left|\det_m(B)\right|.$$

Now let $\begin{pmatrix} \breve{Q}_1 & \breve{Q}_2 \end{pmatrix}\begin{pmatrix} \breve{R}_{11} & \breve{R}_{12} \\ 0 & \breve{R}_{22} \end{pmatrix}$ be the thin QR decomposition of $\breve{B} := \begin{pmatrix} B_2 & F \end{pmatrix}$. Then $\breve{Q}_1 \breve{R}_{11}$ is the thin QR decomposition of B_2, and by the same argument as above, the diagonal elements of $\breve{R}_{22}$ lie in $(0, 1]$. Therefore,

$$\left|\det_{m-k+1}(B_2)\right| = \det(\breve{R}_{11}) \geq \det(\breve{R}_{11})\det(\breve{R}_{22}) = \left|\det_m(\breve{B})\right| = \left|\det_m(\bar{B})\right|,$$

where the last equality follows from part i).

iii) It follows from part ii) that we only need to consider rank-m subsets $B = (b_1, \ldots, b_m) \subseteq A$. Let $B = QR$ and $\hat{B} = \hat{Q}\hat{R}$ be the thin QR decompositions of B and $\hat{B} = \{\hat{b} : b \in B\}$, where $\hat{b} = b/\|b\|$ denotes the normalisation of b. Then $Q = \hat{Q}$ and $R = \hat{R} \cdot \operatorname{diag}(\|b_i\|)$, and therefore,

$$|\det_m B| = \det R = \det \hat{R} \cdot \prod_{i=1}^{m} \|b_i\| \geq |\det_m \hat{B}| \cdot \min_{a\in A} \|a\|^m.$$

The claim now follows easily.

iv) Let $B = (b_1, \ldots, b_m)$ be a subset of A on which the minimum in (3.1) is achieved. Then an ordered subset $F = (f_{m+1}, \ldots, f_n)$ of the canonical unit vectors in $\mathbb{R}^n$ can be chosen to complete B into a basis $\bar{B} = \begin{pmatrix} B & F \end{pmatrix}$ of $\mathbb{R}^n$. If $\begin{pmatrix} Q_1 & Q_2 \end{pmatrix}\begin{pmatrix} R_{11} & R_{12} \\ 0 & R_{22} \end{pmatrix}$ is the QR decomposition of $\bar{B}$, then the thin QR decomposition of B is $Q_1 R_{11}$. Moreover, since the i-th column vector of $\begin{pmatrix} R_{11} & R_{12} \\ 0 & R_{22} \end{pmatrix}$ has norm $\|f_i\| = 1$ for $(i = m+1, \ldots, n)$, the diagonal entries of R_{22} are all ≤ 1. It follows that

$$|\det_m(B)| = \det(R_{11}) \geq \det(R_{11})\det(R_{22}) = |\det(\bar{B})|.$$

It suffices therefore to show that $|\det(\bar{B})| \geq 2^{-\mathscr{D}}$. For $(i = 1, \ldots, m)$ let $\mathscr{D}(b_i)$ be the sum of the bit lengths of the components of b_i. Let $\operatorname{scm}(b_i)$ be the smallest common

multiple of the components of b_i. Then $\operatorname{scm}(b_i)$ is smaller than the product of the denominators of the components in b_i, and this is smaller than $2^{\mathscr{D}(b_i)}$. Moreover, $\bar{B} \cdot \left(\begin{smallmatrix} \operatorname{diag} \operatorname{scm}(b_i) \\ & \mathrm{I} \end{smallmatrix} \right)$ is an integer matrix and hence has integer determinant. Therefore,

$$|\det(\bar{B})| = \frac{\left|\det\left(\bar{B} \cdot \left(\begin{smallmatrix} \operatorname{diag} \operatorname{scm}(b_i) \\ & \mathrm{I} \end{smallmatrix} \right)\right)\right|}{\prod_{i=1}^{k} \operatorname{scm}(b_i)} \geq \frac{1}{2^{\sum_{i=1}^{k} \mathscr{D}(b_i)}} \geq 2^{-\mathscr{D}},$$

which completes the proof. □

DEFINITION 3.2. *For an unordered set A of nonzero vectors from $\mathbb{R}^n$ we write $\kappa(A)$ for the number κ obtained under* (3.1) *for an arbitrary ordering of the elements of A. Part i) of Lemma 3.1 shows that this notion is well-defined. $\kappa(A)$ plays the role of a condition number for the input data set A in quantifying the bounds of Sections 6 and 7.*

4. Selecting Convex Cuts for the Construction of a Gauge. In all of this section we take A to be a set of unit vectors. For $P \subset \mathbb{R}^n$ let π_P be the orthogonal projection of $\mathbb{R}^n$ onto $\operatorname{span}(P)$ along $\operatorname{span}(P)^\perp$, and let $\pi_{P^\perp} = id - \pi_P$. For $r \geq 0$ let us define the sets

$$C_P(r) := \{x \in \mathbb{R}^n : \|\pi_P x\| \leq r\},$$
$$C_{P^\perp}(r) := \{x \in \mathbb{R}^n : \|\pi_{P^\perp} x\| \leq r\}.$$

Note that since any vector in $\mathbb{R}^n$ is orthogonal to all elements of the empty set, we have in particular $\pi_{\emptyset^\perp} = id$, and hence we consider $\pi_\emptyset = 0$ to be the projection onto $\{0\}$. This slight abuse of notation is convenient, for we get $C_{\emptyset^\perp}(1) = \overline{B_1}(0)$ and $C_\emptyset(r) = \mathbb{R}^n$ for all $r \geq 0$. Note also that if $\operatorname{span}(P) = \mathbb{R}^n$ then $C_{P^\perp}(r) = \mathbb{R}^n$ for all $r \geq 0$.

Next we will define an algorithm that defines a set of radii $\{r_P : P \subseteq A\} \subset [0,1]$ which is associated with another set $\{r_{P^\perp}\} \subset [0,1]$ via the relation

$$r_{P^\perp} = \sqrt{1 - r_P^2} \qquad \forall P \subseteq A. \tag{4.1}$$

We write $\mathcal{P}_k(A) = \{P : P \subseteq A, \operatorname{rank}(P) = k\}$, $(k = 1, \ldots, m)$, where $m = \operatorname{rank}(A)$. Let $n_k := \#\mathcal{P}_k(A)$, and let $\mathcal{P}_k = \{P_1^{[k]}, \ldots, P_{n_k}^{[k]}\}$ be an enumeration of the elements of $\mathcal{P}_k$. Also, let $n_{ki} := \#P_i^{[k]}$ and let $P_i^{[k]} = \{a_1^{[ki]}, \ldots, a_{n_{ki}}^{[ki]}\}$ be an enumeration of the elements of $P_i^{[k]}$.

ALGORITHM 4.1.

S0 *Set $r_\emptyset = 0$, $k = m$ and $r_P = 1/2$ $(P \in \mathcal{P}_m(A))$, $r_Q = 1$ $(Q \in \bigcup_{l=1}^{m-1} \mathcal{P}_l(A))$.*
S1 *For $i = 1, \ldots, n_k$*
  for $j = 1, \ldots, n_{ki}$
   if $\operatorname{rank}\big(P_i^{[k]} \setminus \{a_j^{[ki]}\}\big) = k - 1$
    $r_{P_i^{[k]} \setminus \{a_j^{[ki]}\}} \leftarrow \min\left(r_{P_i^{[k]} \setminus \{a_j^{[ki]}\}}, 2^{-3/2} r_{P_i^{[k]}} \sqrt{1 - \|\pi_{P_i^{[k]} \setminus \{a_j^{[ki]}\}} a_j^{[ki]}\|^2}\right)$;
    $r_{\{a_j^{[ki]}\}} \leftarrow \min\left(r_{\{a_j^{[ki]}\}}, 2^{-3/2} r_{P_i^{[k]}} \sqrt{1 - \|\pi_{P_i^{[k]} \setminus \{a_j^{[ki]}\}} a_j^{[ki]}\|^2}\right)$
  end

end
end.
S2 *If* $k > 2$
$k \leftarrow k - 1$
elseif $k = 2$
stop
end.

LEMMA 4.2. *Let A be a finite set of unit vectors in $\mathbb{R}^n$ and let the set of radii $\{r_P : P \subseteq A\}$ be generated by Algorithm 4.1. Then the following properties hold true:*
i) $r_\emptyset = 0$,
ii) $r_P \in (0, 1]$ *for all* $P \subseteq A$ *such that* $P \neq \emptyset$,
iii) for all $P \subset A$ *and* $a \in A$ *such that* $P \neq \emptyset$ *and* $a \notin \operatorname{span}(P)$,

$$r^2_{P\cup\{a\}} \geq \frac{4\left(r^2_{\{a\}} + r^2_P\right)}{1 - \|\pi_P a\|^2}. \tag{4.2}$$

Proof. Condition i) is satisfied because $r_\emptyset$ is set to zero in step **S0** and subsequently never changed. Condition ii) is satisfied because for all $P \subseteq A$ such that $P \neq \emptyset$, r_P is set in $[1/2, 1]$ in step **S0**, and whenever r_P is changed subsequently it is replaced by a smaller but strictly positive value. Since r_P is changed only finitely many times, the final value is in $(0, 1]$. Step **S1** guarantees that condition iii) holds for $(P \cup \{a\}) \in \mathcal{P}_k(A)$ at the end of iteration k, where we count iterations backward from m down to 2. Since $r_{P\cup\{a\}}$ does not change post completion of iteration $k + 1$ and never if $k = m$, and since r_P and $r_{\{a\}}$ can only further decrease after their values are set to satisfy (4.2) in iteration k, the algorithm stops with condition iii) holding for all $(P\cup\{a\}) \in \bigcup_{k=2}^m \mathcal{P}_k$. Moreover, if $(P\cup\{a\}) \in \mathcal{P}_1$ and $a \notin \operatorname{span}(P)$ then $P = \emptyset$, and we don't need to check (4.2). ∎

LEMMA 4.3. *Let $A \subset \mathbb{R}^n$ be a finite set of unit vectors and let the set of radii $\{r_P : P \subseteq A\}$ be generated by Algorithm 4.1 for some specific choice of enumeration of the elements of $\mathcal{P}_k$ and likewise of their elements for $(k = 1, \ldots, m)$. Then $\{r_P : P \subseteq A\}$ and the associated set of radii $\{r_{P^\perp} : P \subseteq A\}$ satisfy the following properties:*
i) $\min\{r_{P^\perp} : P \subseteq A\} \geq 1/2$,
ii) $\min\{r_{\{a\}} : a \in A\} \geq \kappa(A)^{-1} 2^{-\frac{3n+2}{2}}$.

Proof. i) Note that if $k < m$ then every $Q \in \mathcal{P}_k(A)$ arises as $Q = P \setminus \{a\}$ for some $P \in \mathcal{P}_{k+1}(A)$ and a vector $a \in A$ such that $a \notin \operatorname{span}(Q)$. Therefore, if at the beginning of iteration $k+1$ we have $r_P \leq \frac{1}{2}$ for all $P \in \mathcal{P}_{k+1}$ then r_Q will be changed to a value $\leq r_P 2^{-\frac{3}{2}} < \frac{1}{2}$ in iteration $k + 1$. This then implies that at the beginning of iteration k we have $r_Q \leq \frac{1}{2}$ for all $Q \in \mathcal{P}_k$. Since at the beginning of iteration m we have $r_P \leq \frac{1}{2}$ for all $P \in \mathcal{P}_m$, induction on k shows that by the time the algorithm stops we have $r_P \leq \frac{1}{2}$ for all $P \subseteq A$, including $P = \emptyset$ for which we have $r_P = 0$. Therefore, $r_{P^\perp} = \sqrt{1 - r_P^2} \geq \frac{\sqrt{3}}{2} \geq \frac{1}{2}$ for all $P \subseteq A$.

ii) Let $a \in A$. The proof of part i) shows that $r_{\{a\}} \leq \frac{1}{2}$ and hence $r_{\{a\}}$ is changed at least once during the run of Algorithm 4.1. Therefore, there exists an index $k \in \{2, \ldots, m\}$ and a subset $P^{[k]} \in \mathcal{P}_k$ such that $a \in P^{[k]}$, $\operatorname{rank}(P^{[k]} \setminus \{a\}) = k - 1$ and where $P^{[k]}$ and a play the role of $P_i^{[k]}$ and $a_j^{[ki]}$ when $r_{\{a\}}$ is changed for the last

time. That is, we have

$$r_{\{a\}} = r_{P^{[k]}} \cdot \sqrt{1 - \|\pi_{P^{[k]}\setminus\{a^{[k]}\}} a^{[k]}\|^2} \cdot 2^{-\frac{3}{2}}.$$

Moreover, $r_{P^{[k]}}$ is at its final value at this stage because this radius does not change after iteration $k+1$ (again counting backwards from m down to 2). Unless $k = m$, $r_{P^{[k]}}$ too changed at least once, but only finitely many times and only during iteration $k+1$. Hence there was a last time when this radius was changed to

$$r_{P^{[k]}} = r_{P^{[k+1]}} \cdot \sqrt{1 - \|\pi_{P^{[k+1]}\setminus\{a^{[k+1]}\}} a^{[k+1]}\|^2} \cdot 2^{-\frac{3}{2}}$$

for some $P^{[k+1]} \in \mathcal{P}_{k+1}$ and a vector $a^{[k+1]} \in P^{[k+1]}$ which is linearly independent of $P^{[k]}$ and such that $P^{[k+1]} = P^{[k]} \cup \{a^{[k+1]}\}$. Writing $P^{[k-1]}$ for $P^{[k]} \setminus a^{[k]}$ and continuing this construction by induction, we find that there exists a set $P^{[m]} \in \mathcal{P}_m$, and vectors $\{a^{[k]}, \ldots, a^{[m]}\} \subset A$ such that

$$\begin{aligned} P^{[m]} &= P^{[k-1]} \cup \{a^{[k]}, \ldots, a^{[m]}\}, \\ P^{[s]} &:= P^{[k-1]} \cup \{a^{[k]}, \ldots, a^{[s]}\} \in \mathcal{P}_s \quad (s = k, \ldots, m), \end{aligned}$$

and

$$\begin{aligned} r_{\{a\}} &= r_{P^{[m]}} \cdot \prod_{s=k}^{m} \sqrt{1 - \|\pi_{P^{[s-1]}} a^{[s]}\|^2} \cdot 2^{-\frac{3}{2}(m-k+1)} \\ &= \frac{1}{2} \cdot \left| \det_m \left(f_1, \ldots, f_{k-1}, a^{[k]}, \ldots, a^{[m]} \right) \right| \cdot 2^{-\frac{3}{2}(m-k+1)} \\ &\overset{\text{Lem3.1 ii)}}{\geq} \kappa(A)^{-1} \cdot 2^{-\frac{3n+2}{2}}, \end{aligned}$$

where $\{f_1, \ldots, f_{k-1}\}$ is an orthogonal basis of $P^{[k-1]}$. ∎

5. The Gauge Unit Ball and its Properties. In this section we continue to assume that A is a set of unit vectors.

LEMMA 5.1. *Let A be a finite set of unit vectors in $\mathbb{R}^n$, and let $P \subset A$ and $a \in A$ be such that $P \neq \emptyset$ and $a \notin \operatorname{span}(P)$. Finally, let $r_{\{a\}}$, r_P,$r_{\{a\}\cup P} \in (0,1]$ be such that* (4.2) *holds. Then*

$$C_{\{a\}\cup P}\left(\frac{r_{\{a\}\cup P}}{\sqrt{2}}\right) \supset C_{\{a\}}(r_{\{a\}}) \cap C_P(r_P).$$

Proof. Any $x \in \mathbb{R}^n$ can be written as

$$x = \lambda a + \mu \tilde{a} + x^{\perp},$$

where $x^{\perp} \perp \operatorname{span}(P \cup \{a\})$, and where $\tilde{a} \in \operatorname{span}(P)$ is such that $\|\tilde{a}\| = 1$. Then $x \in C_{\{a\}}(r_{\{a\}})$ and $x \in C_P(r_P)$ imply

$$\begin{aligned} \lambda^2 + \mu^2 (a \cdot \tilde{a})^2 + 2\lambda\mu (a \cdot \tilde{a}) &= (a \cdot x)^2 = \|\pi_{\{a\}} x\|^2 \leq r_{\{a\}}^2, \\ \mu^2 + \lambda^2 (a \cdot \tilde{a})^2 + 2\mu\lambda (a \cdot \tilde{a}) &= (\tilde{a} \cdot x)^2 = \|\pi_{\{\tilde{a}\}} x\|^2 \leq \|\pi_P x\|^2 \leq r_P^2. \end{aligned}$$

Adding the two inequalities we get

$$r_{\{a\}}^2 + r_P^2 \geq (\lambda^2 + \mu^2)(1 + (\tilde{a} \cdot a)^2) + 4\lambda\mu(a \cdot \tilde{a}). \tag{5.1}$$

On the other hand,

$$\|\pi_{\{a\}\cup P} x\|^2 = \|\lambda a + \mu\tilde{a}\|^2 = \lambda^2 + \mu^2 + 2\lambda\mu(a \cdot \tilde{a}).$$

Therefore,

$$\begin{aligned} 2\|\pi_{\{a\}\cup P} x\|^2 &= (\lambda^2 + \mu^2)(1 + (\tilde{a} \cdot a)^2) + 4\lambda\mu(a \cdot \tilde{a}) + (\lambda^2 + \mu^2)(1 - (\tilde{a} \cdot a)^2) \\ &\overset{(5.1)}{\leq} r_{\{a\}}^2 + r_P^2 + (\lambda^2 + \mu^2)(1 - (\tilde{a} \cdot a)^2). \end{aligned} \tag{5.2}$$

Without loss of generality we may assume that $0 \leq \tilde{a} \cdot a < 1$. In fact, if $\tilde{a} \cdot a < 0$ then we can replace $\tilde{a}$ by $-\tilde{a}$ and μ by $-\mu$. Moreover, $\|a\| = \|\tilde{a}\| = 1$ and the Cauchy–Schwartz inequality imply that $0 \leq \tilde{a} \cdot a \leq 1$, and if $\tilde{a} \cdot a = 1$ then $a = \tilde{a} \in \operatorname{span}(P)$, contradicting our assumption.

Note that $2|\lambda\mu| \leq \lambda^2 + \mu^2$ by quadratic completion. Hence,

$$4|\lambda\mu|(a \cdot \tilde{a}) \leq 2(\lambda^2 + \mu^2)(a \cdot \tilde{a}),$$

and therefore, (5.1) shows that

$$\begin{aligned} r_{\{a\}}^2 + r_P^2 &\geq (\lambda^2 + \mu^2)(1 + (\tilde{a} \cdot a)^2 - 2(\tilde{a} \cdot a)) \\ &= (\lambda^2 + \mu^2)(1 - \tilde{a} \cdot a)^2, \end{aligned}$$

that is,

$$(\lambda^2 + \mu^2)(1 - (\tilde{a} \cdot a)^2) \leq \frac{\left(r_{\{a\}}^2 + r_P^2\right)(1 - (\tilde{a} \cdot a)^2)}{(1 - \tilde{a} \cdot a)^2} = \frac{\left(r_{\{a\}}^2 + r_P^2\right)(1 + \tilde{a} \cdot a)}{(1 - \tilde{a} \cdot a)}.$$

Now (5.2) implies

$$\begin{aligned} \|\pi_{\{a\}\cup P} x\|^2 &\leq \frac{r_{\{a\}}^2 + r_P^2}{2} \cdot \left(\frac{1 + \tilde{a} \cdot a}{1 - \tilde{a} \cdot a} + 1\right) = \frac{r_{\{a\}}^2 + r_P^2}{1 - \tilde{a} \cdot a} \leq \frac{r_{\{a\}}^2 + r_P^2}{1 - \|\pi_P a\|} \\ &= \frac{\left(r_{\{a\}}^2 + r_P^2\right)(1 + \|\pi_P a\|)}{(1 - \|\pi_P a\|)(1 + \|\pi_P a\|)} \leq \frac{2\left(r_{\{a\}}^2 + r_P^2\right)}{1 - \|\pi_P a\|^2} \\ &\overset{(4.2)}{\leq} \frac{r_{P\cup\{a\}}^2}{2}. \end{aligned}$$

This shows the claim ∎

Lemma 5.2. *Let A be a finite set of unit vectors in $\mathbb{R}^n$, let $P, Q \subseteq A$ be such that $\operatorname{span}(P) \subset \operatorname{span}(Q)$ but $\operatorname{span}(P) \neq \operatorname{span}(Q)$, and let $r_P < r_Q/\sqrt{2}$. Then*

$$C_Q\left(\frac{r_Q}{\sqrt{2}}\right) \cap C_{Q^\perp}(r_{Q^\perp}) \subset C_{P^\perp}(r_{P^\perp}).$$

Proof. Let $x \in C_Q\left(\frac{r_Q}{\sqrt{2}}\right) \cap C_{Q^\perp}(r_{Q^\perp})$. Then

$$\|x\|^2 = \|\pi_Q x\|^2 + \|\pi_{Q^\perp} x\|^2 \leq \frac{r_Q^2}{2} + r_{Q^\perp}^2 = \frac{r_Q^2}{2} + (1 - r_Q^2) = 1 - \frac{r_Q^2}{2}.$$

But then

$$\|\pi_{P^\perp} x\|^2 \leq \|x\|^2 \leq 1 - \frac{r_Q^2}{2} < 1 - r_P^2 = r_{P^\perp}^2.$$

□

COROLLARY 5.3. *Let $P \subset A$ and $a \in A$ be such that $a \notin \operatorname{span}(P)$. If (4.2) holds and $r_\emptyset = 0$ then*

$$\begin{aligned} C_{\{a\}^\perp}(r_{\{a\}^\perp}) \cap C_{\{a\}}(r_{\{a\}}) \cap C_{P^\perp}(r_{P^\perp}) \\ \supseteq C_{\{a\}^\perp}(r_{\{a\}^\perp}) \cap C_{\{a\}}(r_{\{a\}}) \cap C_{(\{a\}\cup P)^\perp}(r_{(\{a\}\cup P)^\perp}). \end{aligned} \tag{5.3}$$

Proof. If $P = \emptyset$ then

$$C_{P^\perp}(r_{P^\perp}) = C_{\emptyset^\perp}(1) = \overline{B_1}(0)$$

and $C_{(\{a\}\cup P)^\perp}(r_{(\{a\}\cup P)^\perp}) = C_{\{a\}^\perp}(r_{\{a\}^\perp})$, and since $C_{\{a\}^\perp}(r_{\{a\}^\perp}) \cap C_{\{a\}}(r_{\{a\}}) \subset \overline{B_1}(0)$, (5.3) is trivially true. We may therefore assume that $P \neq \emptyset$. Let

$$x \in C_{\{a\}^\perp}(r_{\{a\}^\perp}) \cap C_{\{a\}}(r_{\{a\}}) \cap C_{(\{a\}\cup P)^\perp}(r_{(\{a\}\cup P)^\perp})$$

and assume to the contrary of our claim that

$$x \notin C_{P^\perp}(r_{P^\perp}). \tag{5.4}$$

Since $C_{\{a\}^\perp}(r_{\{a\}^\perp}) \cap C_{\{a\}}(r_{\{a\}}) \subset \overline{B_1}(0)$, we have $\|x\| \leq 1$. Therefore, (5.4) implies that

$$1 \geq \|x\|^2 = \|\pi_{P^\perp} x\|^2 + \|\pi_P x\|^2 \overset{(5.4)}{>} r_{P^\perp}^2 + \|\pi_P x\|^2,$$

and this implies that

$$\|\pi_P x\|^2 < 1 - r_{P^\perp}^2 = r_P^2,$$

that is, $x \in C_P(r_P)$. Since x is also in $C_{\{a\}}(r_{\{a\}})$ and (4.2) holds, Lemma 5.1 implies that $x \in C_{\{a\}\cup P}(r_{\{a\}\cup P}/\sqrt{2})$. Since we chose x in $C_{(\{a\}\cup P)^\perp}(r_{(\{a\}\cup P)^\perp})$ and

$$\frac{r_{\{a\}\cup P}^2}{2} \geq \frac{r_{\{a\}}^2 + r_P^2}{1 - \|\pi_P a\|} \geq r_P^2,$$

Lemma 5.2 applied to P and $Q = \{a\} \cup P$ implies that $x \in C_{P^\perp}(r_{P^\perp})$, contrary to assumption (5.4). □

LEMMA 5.4. *Let A be a finite set of unit vectors in $\mathbb{R}^n$, let $\{r_P : P \subseteq A\} \subset [0,1]$ be a set of radii for which the properties of Lemma 4.2 hold, and let*

$$C := \bigcap_{P \subseteq A} C_{P^\perp}(r_{P^\perp}).$$

Then the following implications holds true:

$$a \in A, x \in C \Rightarrow x - (x \cdot a)a \in C,$$
$$a \in A, x \in C, x \cdot a \leq 0, 0 \leq \eta \leq r_{\{a\}} \Rightarrow x + \eta a \in C.$$

Proof. We need to show that $x - (x \cdot a)a, x + \eta a \in C_{P^\perp}(r_{P^\perp})$ for all $P \subseteq A$. Case 1: if $a \in \operatorname{span}(P)$ then the assumption $x \in C_{P^\perp}(r_{P^\perp})$ implies

$$\|\pi_{P^\perp}(x + \lambda a)\| = \|\pi_{P^\perp} x\| \leq r_{P^\perp} \quad \forall \lambda \in \mathbb{R},$$

and hence, $x - (x \cdot a)a, x + \eta a \in C_{P^\perp}(r_{P^\perp})$. Case 2: $a \notin \operatorname{span}(P)$. If follows from case 1 that

$$x - (x \cdot a)a \in C_{\{a\}^\perp}(r_{\{a\}^\perp}) \cap C_{\{a\}}(r_{\{a\}}) \cap C_{(P \cup \{a\})^\perp}(r_{(P \cup \{a\})^\perp}).$$

Equation (5.3) therefore implies that $x - (x \cdot a)a \in C_{P^\perp}(r_{P^\perp})$. Also, together with case 1, the assumptions $x \cdot a \leq 0$ and $0 \leq \eta \leq r_{\{a\}}$ imply that either

$$x + \eta a \in C_{\{a\}^\perp}(r_{\{a\}^\perp}) \cap C_{\{a\}}(r_{\{a\}}) \cap C_{(P \cup \{a\})^\perp}(r_{(P \cup \{a\})^\perp}),$$

and then (5.3) implies that $x + \eta a \in C_{P^\perp}(r_{P^\perp})$, or else $(x + \eta a) \cdot a \leq -r_{\{a\}}$. But then we have $0 \leq \eta \leq \eta + r_{\{a\}} \leq -x \cdot a$, which shows that $x + \eta a$ is a convex combination of x and $x - (x \cdot a)a$, both of which are in $C_{P^\perp}(r_{P^\perp})$. The convexity of $C_{P^\perp}(r_{P^\perp})$ therefore implies that $x + \eta a \in C_{P^\perp}(r_{P^\perp})$. □

6. Relaxation Boundedness for Homogeneous Systems. We are finally ready to start attacking the main results of this paper which generalise and strengthen the classical perceptron boundedness theorem:

THEOREM 6.1. *Let A be a finite set of nonzero vectors in $\mathbb{R}^n$, $x_0 \in \mathbb{R}^n$ a given starting point, $(\eta_i)_{\mathbb{N}_0} \in [0, U]$ and $(a_i)_{\mathbb{N}_0} \subseteq A$ such that the sequence of points defined iteratively by $x_{i+1} = x_i + \eta_i a_i$ satisfies the condition $x_i \cdot a_i \leq 0$ for all $i \in \mathbb{N}_0$. Then for all i, $\|x_i\|$ is bounded by*

$$M(A) := 2 \max \left(\|x_0\|, U \max_{a \in A} \|a\| \cdot \left[\min_{a \in A} \|a\|^{\operatorname{rank}(A)} \kappa(A) 2^{3n/2} + 1 \right] \right)$$

Proof. Let $\{r_{P^\perp} : P \subseteq A\}$, $\{r_P : P \subseteq A\}$ and C be chosen as in Lemma 5.4 when it is applied to $\hat{A} = \{a/\|a\| : a \in A\}$. The set C is a finite intersection of convex, balanced, absorbing sets, whence C inherits these properties. Moreover, C is bounded since $C \subset C_{\emptyset^\perp}(r_{\emptyset^\perp}) = \overline{B_1}(0)$. The gauge

$$\xi : x \mapsto \min\{\lambda \geq 0 : x \in \lambda C\}$$

associated with C is therefore a norm on the finite-dimensional space $\mathbb{R}^n$, see e.g. [13]. Let us define

$$\Lambda_s := U \cdot \max_{a \in A} \|a\| \cdot \max_{\hat{a} \in \hat{A}} r_{\{\hat{a}\}}^{-1},$$
$$\Lambda_l := \Lambda_s + U \cdot \max_{a \in A} \|a\| \cdot \max_{P \subseteq \hat{A}} r_{P^\perp}^{-1}.$$

Suppose $\xi(x_i) \geq \Lambda_s$ for some $i \in \mathbb{N}$. Note that the definition of ξ implies that $x_i \in \xi(x_i)C$. Moreover, since $x_i \cdot a \leq 0$ and

$$\frac{\eta_i \cdot \|a_i\|}{\xi(x_i)} \leq U \cdot \max_{a \in A} \|a\| \cdot \Lambda_s^{-1} \leq r_{\{\hat{a}_i\}},$$

Lemma 5.4 shows that

$$x_i + \eta_i a_i = \xi(x_i) \left(\frac{x_i}{\xi(x_i)} + \frac{\eta_i \cdot \|a_i\|}{\xi(x_i)} \hat{a}_i \right) \in \xi(x_i)C,$$

and hence, $\xi(x_i + \eta_i a_i) \leq \xi(x_i)$. That is, as long as $\xi(x_i) \geq \Lambda_s$, the gauge value can only decrease in subsequent iterations. On the other hand, if $\xi(x_i) \leq \Lambda_s$ then $\|\pi_{P^\perp} x_i\| \leq \Lambda_s r_{P^\perp}$ for all $P \subseteq \hat{A}$, and hence,

$$\|\pi_{P^\perp}(x_i + \eta_i a_i)\| \leq \|\pi_{P^\perp} x_i\| + \eta_i \|\pi_{P^\perp} a_i\| \leq \Lambda_s r_{P^\perp} + U \cdot \|a_i\| \leq \Lambda_l r_{P^\perp}.$$

Since this is true for all $P \subseteq \hat{A}$, we have $\xi(x_{i+1}) \leq \Lambda_l$. In summary,

$$\xi(x_{i+1}) \leq \begin{cases} \Lambda_l & \text{if } \xi(x_i) \leq \Lambda_s, \\ \xi(x_i) & \text{if } \xi(x_i) \geq \Lambda_s. \end{cases} \tag{6.1}$$

Iterative application of (6.1) shows that $\xi(x_i) \leq \max(\xi(x_0), \Lambda_l)$ for all i. But note that $\xi(x) \geq \|x\|$ for all $x \in \mathbb{R}^n$, because the unit ball of ξ is contained in the unit ball of $\|\cdot\|$. Therefore, $(x_i)_\mathbb{N}$ is bounded in the Euclidean norm by $\max(\xi(x_0), \Lambda_l)$.

It remains only to quantify Λ_l. It follows from Lemma 4.3 i) that $\xi(x_0) \leq 2\|x_0\|$. Part ii) of the same result shows that $\Lambda_s \leq U \cdot \max_{a \in A} \|a\| \cdot \kappa(\hat{A}) 2^{\frac{3n+2}{2}}$. Finally, another application of part i) in conjunction with part iii) of Lemma 3.1 yields

$$\Lambda_l \leq 2U \max_{a \in A} \|a\| \cdot \left[\min_{a \in A} \|a\|^{\operatorname{rank}(A)} \kappa(A) 2^{3n/2} + 1 \right].$$

Combining the bounds on $\xi(x_0)$ and Λ_l, one obtains the claimed bound. □

The explicit formula for $M(A)$ as a function of the input data A solves an old open problem raised by Block and Levin [4]. Since $M(A)$ depends on A through $\kappa(A)$, the bound can be very large when A contains subsets that are close to rank-deficient, but not. Note that to compute $\kappa(A)$ one usually has to evaluate an exponential number of determinants. The situation is greatly simplified in the case where A consists of rational input data, as the following corollary shows:

COROLLARY 6.2. *If A consists of rational data of total bit length $\mathscr{D}$ then*

$$M(A) \leq 2 \max \left(\|x_0\|, U \cdot \left[2^{(n+2)\mathscr{D} + 3n/2} + 2^{\mathscr{D}} \right] \right).$$

Proof. Arguing as in the proof of Lemma 3.1, it is easy to see that $\max_{a \in A} \|a\| \leq 2^{\mathscr{D}}$. Using $m \leq n$ and part iv) of Lemma 3.1, the claim is now an immediate consequence of Theorem 6.1. □

7. Extension to Inhomogeneous Systems. Theorem 6.1 provides information about the behaviour of the relaxation method applied to the system of homogeneous linear inequalities $A^{\mathrm{T}}x > 0$. We will now extend this result so that it is valid for inhomogeneous systems $A^{\mathrm{T}}x > b$ too. This can be achieved through the standard process of homogenisation and the careful construction of a new gauge.

THEOREM 7.1. *Let $A = (a^{[1]}, \dots, a^{[k]})$ be an ordered set of k nonzero vectors from $\mathbb{R}^n$, $b \in \mathbb{R}^k$, $x_0 \in \mathbb{R}^n$, $(\eta_i)_{\mathbb{N}_0} \subset [0, U]$, and $(j_i)_{\mathbb{N}_0} \subseteq \{1, \dots, k\}$ such that the sequence of points defined iteratively by $x_{i+1} = x_i + \eta_i a^{[j_i]}$ satisfies the condition $x_i \cdot a^{[j_i]} \le b_{j_i}$ for all $i \in \mathbb{N}_0$. Let $\tilde{A} := \left(\begin{smallmatrix} A & 0 \\ -b^{\mathrm{T}} & 1 \end{smallmatrix}\right)$. Then for all i, $\|x_i\|$ is bounded by*

$$M(A,b) := 2\max\left(\left\|\left(\begin{smallmatrix} x_0 \\ 1 \end{smallmatrix}\right)\right\|, 1 + U \max_{\tilde{a}\in\tilde{A}} \|\tilde{a}\| \cdot \left[\min_{\tilde{a}\in\tilde{A}} \|\tilde{a}\|^{\mathrm{rank}(\tilde{A})} \kappa(\tilde{A}) 2^{3(n+1)/2} + 1\right]\right) \tag{7.1}$$

Moreover, if A and b consist of rational data of total bit length $\mathscr{D}$ then

$$M(A,b) \le 2\max\left(\left\|\left(\begin{smallmatrix} x_0 \\ 1 \end{smallmatrix}\right)\right\|, 1 + U \cdot \left[2^{(n+3)\mathscr{D} + 3(n+1)/2} + 2^{\mathscr{D}}\right]\right). \tag{7.2}$$

Proof. Note that $a^{[i]} \cdot x > b_i$ if and only if $(\tilde{a}^{[i]})^{\mathrm{T}}\left(\begin{smallmatrix} x \\ 1 \end{smallmatrix}\right) > 0$ where $\tilde{a}^{[i]}$ is the i-th column vector of $\tilde{A}$ ($i = 1, \dots, k$). Moreover, $(\tilde{a}^{[k+1]})^{\mathrm{T}}\left(\begin{smallmatrix} x \\ 1 \end{smallmatrix}\right) = 1 > 0$ is always satisfied, and $j_i \ne k+1$ for all i. There is a one-one correspondence between the points in $\mathbb{R}^n$ and those in the affine hyperplane $H := \left\{\left(\begin{smallmatrix} x \\ z \end{smallmatrix}\right) \in \mathbb{R}^{n+1} : z = 1\right\}$ via the projection $\Psi : \left(\begin{smallmatrix} x \\ 1 \end{smallmatrix}\right) \mapsto x$. Since $\|x\| < \|\Psi^{-1}x\|$, it suffices to show that the sequence $(\|\Psi^{-1}x_i\|)_{\mathbb{N}_0}$ is bounded by $M(A,b)$. But note that $\Psi^{-1}x_{i+1} = \Phi\, \tilde{x}_{i+1}$, where $\tilde{x}_{i+1} = \left(\begin{smallmatrix} x_i \\ 1 \end{smallmatrix}\right) + \eta_i \tilde{a}^{[j_i]}$ and $\Phi : \left(\begin{smallmatrix} x \\ z \end{smallmatrix}\right) \mapsto \left(\begin{smallmatrix} x \\ 1 \end{smallmatrix}\right)$ is the projection of $\mathbb{R}^{n+1}$ into H along the $(n+1)$-th coordinate axis. Let C be the gauge unit ball corresponding to $\hat{A} := \{\tilde{a}/\|\tilde{a}\| : \tilde{a} \in \tilde{A}\}$ under the construction of Lemma 5.4, and let ξ be the associated gauge on $\mathbb{R}^n$. Let us set

$$\Lambda_s := \max(U,1) \cdot \max_{\tilde{a}\in\tilde{A}} \|\tilde{a}\| \cdot \max_{\hat{a}\in\hat{A}} r_{\{\hat{a}\}}^{-1},$$
$$\Lambda_l := \Lambda_s + \max(U,1) \cdot \max_{\tilde{a}\in\tilde{A}} \|\tilde{a}\| \cdot \max_{P\subseteq\hat{A}} r_{P^\perp}^{-1} + 2.$$

The same arguments as those leading to (6.1) demonstrate that

$$\xi(\tilde{x}_{i+1}) \le \begin{cases} \Lambda_l & \text{if } \xi(\Psi^{-1}x_i) < \Lambda_s, \\ \xi(\Psi^{-1}x_i) & \text{if } \xi(\Psi^{-1}x_i) \ge \Lambda_s. \end{cases} \tag{7.3}$$

Lemma 5.4 shows that

$$\xi\left(\left(\begin{smallmatrix} x_{i+1} \\ 0 \end{smallmatrix}\right)\right) = \xi\left(\tilde{x}_{i+1} - (\tilde{x}_{i+1} \cdot \tilde{a}^{[k+1]})\tilde{a}^{[k+1]}\right) \le \xi(\tilde{x}_{i+1})$$

Therefore, if $\xi\left(\left(\begin{smallmatrix} x_{i+1} \\ 0 \end{smallmatrix}\right)\right) \ge \Lambda_s$, then Lemma 5.4 implies

$$\xi(\Psi^{-1}x_{i+1}) = \xi\left(\left(\begin{smallmatrix} x_{i+1} \\ 0 \end{smallmatrix}\right) + 1 \cdot \tilde{a}^{[k+1]}\right) \le \xi\left(\left(\begin{smallmatrix} x_{i+1} \\ 0 \end{smallmatrix}\right)\right) \le \xi(\tilde{x}_{i+1}). \tag{7.4}$$

On the other hand, if $\xi\left(\left(\begin{smallmatrix} x_{i+1} \\ 0 \end{smallmatrix}\right)\right) < \Lambda_s$, then the triangular inequality of the norm ξ and the inequality $\xi \le 2\|\cdot\|$ – which follows from Lemma 4.3 i) – imply that

$$\xi(\Psi^{-1}x_{i+1}) \le \xi\left(\left(\begin{smallmatrix} x_{i+1} \\ 0 \end{smallmatrix}\right)\right) + \xi\left(\left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right)\right) < \Lambda_s + 2 < \Lambda_l. \tag{7.5}$$

The combination of (7.3), (7.4) and (7.5) now show

$$\xi(\Psi^{-1}x_{i+1}) \leq \max\big(\Lambda_l, \xi(\Psi^{-1}x_i)\big). \tag{7.6}$$

Applying (7.6) iteratively, we find

$$\|x_i\| < \|\Psi^{-1}x_i\| \leq \xi(\Psi^{-1}x_i) \leq \max\big(\xi(\Psi^{-1}x_0), \Lambda_l\big).$$

It remains to bound $\xi(x_0)$ and Λ_l, which can be done as in the proof of Theorem 6.1, leading to (7.1). Finally, progressing as in the proof of Corollary 6.2, one obtains (7.2). □

REFERENCES

[1] S. Agmon. The relaxation method for linear inequalities. *Canadian Journal of Mathematics*, 6:382–392, 1954.

[2] H.H. Bauschke and J.M. Borwein. Legendre functions and the method of random bregman projections. *Journal of Convex Analysis*, 4:27–67, 1997.

[3] H.H. Bauschke, J.M. Borwein, and A. Lewis. The method of cyclic projections for closed convex sets in hilbert space. *Contemporary Mathematics*, 204:1–38, 1997.

[4] H.D. Block and S.A. Levin. On the boundedness of an iterative procedure for solving a system of linear inequalities. *Proc. Amer. Math. Soc.*, 26:229–235, 1970.

[5] Y. Censor and S.A. Zenios. *Parallel Optimization: Theory, algorithms and application.* Oxford University Press, 1997.

[6] G. Cimmino. Calcolo approssimato per soluzioni dei sistemi di equazioni lineari. *La Ricerca Scientifica XVI, Series II, Anno IX*, 1:326–333, 1938.

[7] B. Effron. The perceptron correction procedure in nonseparable situations, 1964. Rome Air Development Center Technical Documentary Report RADC-TDR-63-533.

[8] J.-L. Goffin. The relaxation method for solving systems of linear inequalities. *Mathematics of Operations Research*, 5:388–414, 1980.

[9] J.-B. Hiriart-Urruty and C. Lemaréchal. *Convex Analysis and Minimization Algorithms.* Grundlehren der mathematischen Wissenschaften 305, Springer Verlag, 1991.

[10] L.G. Khachiyan. A polynomial algorithm in linear programming. *Dokl. Akad. Nauk SSSR*, 244:1093–1096, 1979. (In Russian, English translation in *Soviet Math. Dokl.*, 20:191–194, 1979.).

[11] M.L. Minsky and S.A. Papert. *Perceptrons.* MIT Press, 1969.

[12] T. Motzkin and I.Y. Schönberg. The relaxation method for linear inequalities. *Canadian Journal of Mathematics*, 6:393–404, 1954.

[13] A.P. Robertson and W.J. Robertson. *Topological Vector Spaces.* Cambridge University Press, 1973.

[14] R. Rosenblatt. *Principles of neurodynamics: Perceptrons and the theory of brain mechanisms.* Spartan Books, 1962.

[15] N.Z. Shor. *On the structure of algorithms for the numerical solution of optimal planning and design problems (in Russian).* PhD thesis, Dissertaion, Cybernetics Institute, Academy of Sciences of the Ukrainian SSR, Kiev, 1964.

[16] N.Z. Shor. Utilization of the operation of space dilatation in the minimization of convex functions (in russian). *Kibernetika (Kiev)*, 1:6–12, 1970.

[17] D.B. Yudin and A.S. Nemirovskii. Informational complexity and efficient methods for the solution of convex extremal problems (in russian). *Èkonomika i Matematicheskie Metody*, 12:357–369, 1976.