SIAM J. DISCRETE MATH.
Vol. 30, No. 4, pp. 2064–2085

TROPICAL COMPLEXITY, SIDON SETS, AND DYNAMIC PROGRAMMING*

STASYS JUKNA†

Abstract. Many dynamic programming algorithms for discrete 0-1 optimization problems are just special (recursively constructed) tropical (min,+) or (max,+) circuits. A problem is homogeneous if all its feasible solutions have the same number of ones. Jerrum and Snir [*J ACM* 29 (1982), pp. 874–897] proved that tropical circuit complexity of homogeneous problems coincides with the monotone arithmetic circuit complexity of the corresponding polynomials. So, lower bounds on the monotone arithmetic circuit complexity of these polynomials yield lower bounds on the tropical complexity of the corresponding optimization problems. But the situation with nonhomogeneous problems is entirely different: here the gap between their tropical and arithmetic complexities can be even exponential. In this paper, we improve two classical lower bounds for monotone arithmetic circuits—Schnorr's bound and Hyafil–Valiant's bound—and use these improvements to derive general lower bounds for the tropical circuit complexity of nonhomogeneous optimization problems. In particular, we show that optimization problems, whose sets of feasible solutions are cover free, have large tropical complexity.

Key words. tropical circuits, arithmetic circuits, dynamic programming, cover-free sets, Sidon sets, lower bounds

AMS subject classifications. 68Q17, 68R05, 05C35

DOI. 10.1137/16M1064738

1. Introduction.

Let $\mathbb{N} = \{0, 1, 2, \ldots\}$. Every finite set $A \subset \mathbb{N}^n$ of vectors (of *feasible solutions*) defines two natural optimization problems: given an assignment $x \in \mathbb{N}^n$ of nonnegative integer weights, compute the minimum or the maximum weight $\langle a, x\rangle = a_1x_1 + \cdots + a_nx_n$ of a feasible solution $a \in A$.

For example, if $A \subset \{0,1\}^{n^2}$ is the set of all characteristic 0-1 vectors of perfect matchings in a complete bipartite $n \times n$ graph $K_{n,n}$, then the corresponding optimization problem on A is, given an assignment of weights to the edges of $K_{n,n}$, compute the minimum or the maximum weight of a perfect matching.

Every such problem can be solved by a *tropical*[1] $(\min, +)$ or $(\max, +)$ circuit. Such a circuit is a directed acyclic graph with n source (indegree zero) nodes $x_1, \ldots, x_n$. Every other node (called a *gate*) has indegree two, and computes either the sum or minimum/maximum of the values computed at its two predecessors. The *size* of a circuit is the total number of its gates. Given a set $A \subset \mathbb{N}^n$ of feasible solutions, let $\mathrm{Min}(A)$ and $\mathrm{Max}(A)$ denote, respectively, the minimum size of a $(\min, +)$ and $(\max, +)$ circuit solving the corresponding optimization problem on A. We will refer to these measures as the *tropical complexity* of A.

Motivation. Besides being interesting in their own right, the importance of tropical circuits stems form their intimate connection with dynamic programming (DP)

*Received by the editors March 8, 2016; accepted for publication (in revised form) August 10, 2016; published electronically November 1, 2016.

http://www.siam.org/journals/sidma/30-4/M106473.html

Funding: This research was supported by DFG grant SCHN 503/6-1.

†Institute of Computer Science, Goethe University Frankfurt, Frankfurt am Main, Germany; affiliated with Institute of Mathematics and Informatics, Vilnius University, Vilnius, Lithuania (jukna@thi.informatik.uni-frankfurt.de, http://www.thi.cs.uni-frankfurt.de/%7Ejukna/).

[1]The adjective "tropical" is not in contrast with "polar geometry." It was coined by French mathematicians in honor of Imre Simon who lived in Sao Paulo (south tropic). Tropical algebra and geometry are now actively studied topics in mathematics.

algorithms. Many of these algorithms are "pure" in that their recursion equations only use Min and Sum or Max and Sum operations. Prominent examples of pure DP algorithms are the well-known Bellman–Ford–Moore DP algorithm for the single-source shortest paths [2, 9, 25], the Floyd–Warshall DP algorithm for the all-pairs shortest paths problem [8, 35], the Held–Karp DP algorithm for the traveling salesman problem [14], and others. It is clear that every pure DP algorithm is just a special, recursively constructed tropical circuit. So, lower bounds on the size of tropical circuits show limitations of pure DP. For example, the Held–Karp DP algorithm gives a tropical circuit of size $O(n^2 2^n)$ solving the traveling salesman problem. On the other hand, Jerrum and Snir [17] have shown that $\Omega(n^2 2^n)$ gates are also necessary in any such circuit solving this problem. This implies that the Held–Karp DP algorithm is *optimal* among all pure DP algorithms for this problem.

Jerrum and Snir [17] observed that tropical complexity is related to a more tractable measure of Minkowski complexity of vector sets. Recall that the *sumset* or the *Minkowski sum* of two sets $X, Y \subseteq \mathbb{N}^n$ of vectors is the set

$$X + Y = \{x + y \colon x \in X \ \text{ and } \ y \in Y\},$$

where $x + y = (x_1 + y_1, \ldots, x_n + y_n)$ is the componentwise addition of vectors.

A *Minkowski circuit* is a directed acyclic graph with n source (indegree zero) nodes holding single-element sets $\{e_1\}, \ldots, \{e_n\}$, where e_i is a 0-1 vector with exactly one 1 in the ith position. Every other node, a *gate*,[2] has indegree two, and performs either the set-theoretic union or the sumset operation on its two inputs (two gates entering this gate). The *Minkowski complexity* of a vector set A, which we denote by $\mathrm{L}(A)$, is the minimum number of gates in a Minkowski circuit producing this set. So, the Minkowski complexity of A is the minimum number of set-theoretic union $X \cup Y$ and sumset $X + Y$ operations required to create the set A when starting from single-element sets $\{e_1\}, \ldots, \{e_n\}$.

Remark 1. Minkowski complexity of sets of vectors is intimately related to the monotone *arithmetic* $(+, \times)$ circuit complexity of monotone multivariate polynomials. There is a natural homomorphism from the semiring of monotone multivariate polynomials to the semiring $(2^{\mathbb{N}^n}, \cup, +)$ of finite subsets of vectors which maps every polynomial

$$f(x_1, \ldots, x_n) = \sum_{a \in A_f} c_a \prod_{i=1}^{n} x_i^{a_i} \tag{1}$$

with $A_f \subset \mathbb{N}^n$ and positive integer coefficients c_a to the set A_f of its exponent vectors. In particular, each variable x_i is mapped to $A_{x_i} = \{e_i\}$. That this is indeed a homomorphism follows from easily verifiable equalities $A_{f+h} = A_f \cup A_h$ and $A_{f \cdot h} = A_f + A_h$. This, in particular, implies that the Minkowski complexity $\mathrm{L}(A_f)$ is a lower bound on the number of gates in any monotone arithmetic $(+, \times)$ circuit computing any multivariate polynomial f of the form (1) with arbitrary positive coefficients c_a. In fact, all known lower bounds on the monotone arithmetic $(+, \times)$ circuit complexity of polynomials f, including [29, 31, 34, 17, 11, 33, 30, 12, 28, 15], are lower bounds on the Minkowski complexity $\mathrm{L}(A_f)$ of their sets A_f of exponent vectors; see also recent surveys [32, 3].

[2]The term "gate" comes from electronic engineering, and is only used to stress that a node has its associated operation.

Suppose now we have a tropical circuit solving an optimization problem on a given set $A \subset \mathbb{N}^n$ of feasible solutions. If we replace every input variable x_i by the single-element set $\{e_i\}$, every min/max gate by the union gate, and every sum gate by the sumset gate, then the resulting Minkowski circuit will create some set $B \subset \mathbb{N}^n$.

It is clear that this (unknown to us) set B does not need to coincide with A. The only information about this set is that it must define the same optimization problem as the original set A. That is, for every vector $x \in \mathbb{N}^n$ of weights, the minimum/maximum of $\langle b, x\rangle$ over all $b \in B$ must coincide with the minimum/maximum of $\langle a, x\rangle$ over all $a \in A$. Still, the following lemma shows that such a set B is not completely "unknown."

We say that a vector $a \in \mathbb{R}^n$ *contains* a vector $b \in \mathbb{R}^n$ if $a \geq b$ holds, that is, if $a_i \geq b_i$ holds for all $i = 1, \ldots, n$. A set A of vectors is an *antichain* if no vector of A contains another vector of A. For sets $A, B \subseteq \mathbb{N}^n$ of vectors, we say that

- B *lies above* A if every vector of B contains at least one vector of A;
- B *lies below* A if every vector of B is contained in at least one vector of A.

LEMMA 1 (structural lemma). *If $A \subseteq \{0,1\}^n$ is an antichain, then*

$$\mathrm{Min}(A) = \min\{\mathrm{L}(B)\colon A \subseteq B \subset \mathbb{N}^n \text{ and } B \text{ lies above } A\}\,;$$
$$\mathrm{Max}(A) = \min\{\mathrm{L}(B)\colon A \subseteq B \subseteq \{0,1\}^n \text{ and } B \text{ lies below } A\}\,.$$

The lemma (for $\mathrm{Min}(A)$) is a special version of a more general result proved by Jerrum and Snir [17, Theorem 2.8] using a basic separation result in convexity theory—the Farkas theorem. An elementary proof (without any use of Farkas' theorem) was given in [19, Lemmas 6,7]; for completeness, we recall the proof for $\mathrm{Min}(A)$ in Appendix A.

The structural lemma implies that, at least in the case of 0-1 optimization problems (when the underlying set A of feasible solutions consists of 0-1 vectors) their tropical complexity can be lower bounded by the Minkowski complexity of appropriate subsets of A.

To be more specific, define the *degree* of a vector $a \in \mathbb{N}^n$ to be its Manhattan norm, that is, the sum $a_1 + \cdots + a_n$ of its entries. A set of vectors is *homogeneous* if all its vectors have the same degree. The *lower envelope* A_{le} of a set A is the set of all vectors in A of minimum degree, and the *upper envelope* A_{ue} is the set of all vectors in A of maximum degree. Note that both these sets are homogeneous. Also, if A itself is homogeneous, then $A_{\mathrm{le}} = A_{\mathrm{ue}} = A$. As observed in [17], by appropriately discarding some of the edges entering union ($\cup$) gates, one can easily show that

$$\mathrm{L}(A) \geq \max\{\mathrm{L}(A_{\mathrm{le}}), \mathrm{L}(A_{\mathrm{ue}})\}\,.$$

Using the structural lemma, this observation can be extended to tropical circuits.

LEMMA 2 (reduction lemma). *If $A \subseteq \{0,1\}^n$ is an antichain then* $\mathrm{Min}(A) \geq \mathrm{L}(A_{\mathrm{le}})$ *and* $\mathrm{Max}(A) \geq \mathrm{L}(A_{\mathrm{ue}})$. *In particular, if A itself is homogeneous, then*

$$\mathrm{Min}(A) = \mathrm{Max}(A) = \mathrm{L}(A)\,.$$

Indeed, by the structural lemma, we have $\mathrm{Min}(A) = \mathrm{L}(B)$ for some set $B \subset \mathbb{N}^n$ of vectors such that $A \subseteq B$ and B lies above A. These two latter conditions on B imply that $B_{\mathrm{le}} = A_{\mathrm{le}}$. Thus, $\mathrm{Min}(A) = \mathrm{L}(B) \geq \mathrm{L}(B_{\mathrm{le}}) = \mathrm{L}(A_{\mathrm{le}})$. The proof of $\mathrm{Max}(A) \geq \mathrm{L}(A_{\mathrm{ue}})$ is similar.

The second claim of the reduction lemma is important: it shows that pure DP algorithms for 0-1 optimization problems with *homogeneous* sets of feasible solutions

are no more powerful than monotone arithmetic circuits! This explains why we do not have efficient DP algorithms for homogeneous optimization problems whose arithmetic complexity is large.

But what about *nonhomogeneous* optimization problems?

Since we only consider optimization problems with *nonnegative* weights, the question is not trivial: for some nonhomogeneous sets of vectors, the gap between their tropical and Minkowski complexities can even be exponential.

Example 1. Take any homogeneous set $A \subset \{0,1\}^n$ whose Minkowski complexity is superpolynomial in n, like those given in Examples 3 and 5. Since the set is homogeneous, the reduction lemma implies that then the tropical complexity $\mathrm{Min}(A)$ of A is also superpolynomial. Consider now the extended set $B = A \cup \{e_1, \ldots, e_n\}$. Since A is the upper envelope of B, the reduction lemma implies that the Minkowski complexity $\mathrm{L}(B) \geq \mathrm{L}(A)$ remains superpolynomial. But the tropical complexity $\mathrm{Min}(B)$ drops down to n: just compute the minimum $f(x) = \min\{x_1, \ldots, x_n\}$. Since the weight are nonnegative, $f(x)$ is exactly the minimum of $\langle b, x\rangle$ over all vectors $b \in B$.

Example 2. To give a less artificial example of a nonhomogeneous set A exhibiting an exponential gap between tropical and Minkowski complexities, let A be the set of all characteristic 0-1 vectors of paths in K_n from node 1 to node n; here and throughout, K_n denotes the complete undirected graph on nodes $1, \ldots, n$. As shown in [17], we have $\mathrm{L}(A) = 2^{\Omega(n)}$. But the well-known Bellman–Ford DP algorithm for the shortest s-t path problem yields $\mathrm{Min}(A) = O(n^3)$.

Our main results are general lower bounds on the size of *tropical* circuits solving optimization problems which are *not* necessarily homogeneous (Theorems C and E). We obtain these results by extending to tropical circuits two classical lower bounds for monotone *arithmetic* circuits—Schnorr's bound and Hyafil–Valiant's bound (Theorems A, B, and D).

Remark 2 (role of the domain). It is clear that the larger the domain $D \subseteq \mathbb{R}$ of allowed weights $x_1, \ldots, x_n$ is, the easier is the task of proving lower bounds on the tropical complexity of optimization problems. The most difficult case is $D = \{0,1\}$. In this case, we arrive at monotone *boolean* $(\vee, \wedge)$ circuits because then $x \wedge y = \min\{x, y\}$ and $x \vee y = \min\{1, x + y\}$. For these circuits, only a few nontrivial lower bounds are known, and their proofs are rather involved (see, for example, [18, Chapter 9]). On the other hand, if we allow negative weights, say, $D = \mathbb{Z}$, then we arrive at monotone *arithmetic* circuits: in this case, both $\mathrm{Min}(A)$ and $\mathrm{Max}(A)$ coincide with $\mathrm{L}(A)$ (see [17, Corollary 2.7] or [19, Lemma 5]). Thus, tropical circuits over the domains $D = \{0,1\}$ or $D = \mathbb{Z}$ do not constitute any new model of computation: we then essentially have monotone *boolean* or monotone *arithmetic* circuits. But the case $D = \mathbb{N}$ (considered in this paper) is already interesting because then tropical circuits can be exponentially more powerful than arithmetic circuits (Examples 1, 2), and can be exponentially weaker than monotone boolean circuits [19, section 7].

Motivating Example. In the well-known *assignment problem* problem, we obtain an assignment of nonnegative integer weights to the edges of a complete bipartite $n \times n$ graph $K_{n,n}$, and the goal is to compute the minimum weight of a perfect matching. The arithmetic version of this problem is to compute the permanent of an $n \times n$ matrix, and it is long known that its monotone *arithmetic* complexity is exponential in n. This was first shown by Jerrum and Snir [17]; see also Example 5 below. Since the problem is homogeneous (every perfect matching has the same number n of edges)

the reduction lemma implies that the *tropical* complexity of the assignment problem is exponential as well.

Let us now slightly perturb the problem by adding some other subgraphs of $K_{n,n}$ with less than n edges as feasible solutions. The arithmetic complexity will then remain exponential because the upper envelope of the new problem will remain the same—the set of all perfect matchings. What can then be said about the tropical complexity of the modified *minimization* problem? If we add all single edges as feasible solutions then this complexity drops down from exponential to linear (in the number n^2 of variables): just compute the minimum of weights of single edges.

But what happens if we add, say, 3-stars $K_{1,3}$ as feasible solutions: will the min-plus complexity then also drop down exponentially? If $n \geq 4$, then the reduction lemma cannot yield any lower bound larger than $n\binom{n}{3} = \Theta(n^4)$ since the lower envelope of the (modified) problem consists of 3-stars. Still, one of our lower bounds (Theorem D) implies that the min-plus complexity of the modified problem remains $2^{\Omega(n)}$. And, as we will see, the (somewhat unexpected) reason for this to happen is that no 3-star can be contained in a union of two perfect matchings.

2. Known results. In this section, we recall the most basic lower bounds on the Minkowski (and hence, also arithmetic) complexity $\mathrm{L}(A)$ of sets of vectors $A \subset \mathbb{N}^n$. It is clear that $\mathrm{L}(A) \leq d|A|$ holds for every set, where d is the maximum degree of (sum of entries of) a vector in A. So, what sets A have large Minkowski complexity, near to $|A|$? The first set of these bounds shows that such are the well-known cover-free and Sidon sets.

A set $A \subset \mathbb{N}^n$ of vectors is a *Sidon set* if it has the following property: if we know the sum of two vectors of A, we know which vectors were added. That is, if $a+b = c+d$ holds for some vectors of A then $\{c,d\} = \{a,b\}$. For example, in the case $n = 1$, $A = \{1,2,5,7\}$ is a Sidon set, but $B = \{1,2,4,5,7\}$ is not a Sidon set because, for example, $2+4 = 1+5$. Sidon sets $A \subset \mathbb{N}$ are also known as *Golomb rulers*.

The interpretation of Sidon sets in terms of *graphs* is the following. Associate with a set $A \subset \mathbb{N}^n$ of vectors a bipartite graph G_A whose nodes are vectors in $\mathbb{N}^n$, and two nodes x and y are adjacent precisely when $x+y \in A$. Then A is a Sidon set if and only if the graph G_A contains no copy of a complete bipartite 2×2 graph; see Lemma 10 in Appendix B.

The term "Sidon sets" was coined by Erdős in honor of Fourier analyst Simon Sidon who introduced these sets in order to solve a problem in harmonic analysis. In one of the first papers on these sets, Erdős and Turán [7] have shown that the maximum size of a Sidon subset of $\{1\ldots,m\}$ is asymptotically equal to $m^{1/2}$. In larger dimensions, Lindström [24] and Cilleruelo [4] have shown that the maximum size of a Sidon set in $\{1,\ldots,m\}^n$ is asymptotically equal to $m^{n/2}$. A comprehensive bibliography on Sidon sets and their explicit constructions is given by O'Bryant [26].

An important special case of Sidon sets are *cover-free* sets A: if $a+b \geq c$ holds for some vectors of A, then $c \in \{a,b\}$. These sets were first introduced in 1964 by Kautz and Singleton [20] to investigate nonrandom superimposed binary codes. Cover-free families have been considered for many cryptographic problems. Erdős, Frankl, and Füredi [5] have shown that the maximum number $m = |A|$ of vectors in a cover-free set $A \subseteq \{0,1\}^n$ satisfies $1.134^n < m < 1.25^n$.

2.1. Schnorr's combinatorial bound. One of the first (if not the first) general lower bound on the Minkowski complexity of vector sets (and, hence, also on the monotone arithmetic circuit complexity of polynomials) was proved by Schnorr [29].

A set $B \subseteq A \subset \mathbb{N}^n$ is *cover-free inside* A if for every $a, b \in B$ and $c \in A$, $a + b \geq c$ implies $c \in \{a, b\}$.

THEOREM 1 (Schnorr [29]). *If $B \subseteq A$ is cover-free inside A, then* $\mathrm{L}(A) \geq |B|$.

Explicit homogeneous cover-free sets $A \subseteq \{0,1\}^n$ of size $|A| = 2^{\Omega(n)}$ can be constructed using error-correcting codes; see Appendix C. The next example gives a smaller but combinatorially "cleaner" explicit cover-free set.

Example 3 (Schnorr [29]). Let $n = \binom{m}{2}$, and let $A \subset \{0,1\}^n$ be the set of all $|A| = \binom{m}{k}$ characteristic vectors of k-cliques (complete graphs on k nodes) in K_n; we view a k-clique as the set of its $\binom{k}{2}$ edges. To show that A is cover free, assume contrariwise that the union of some two k-cliques contains some third k-clique. Since all k-cliques have exactly k nodes, the latter clique must then have a node u not in the first clique and a node v not in the second clique. If $u = v$ then the node u is not covered, and if $u \neq v$ then the edge $\{u, v\}$ is not covered by the first two cliques, a contradiction. Thus, A is cover free. Together with the reduction lemma, Theorem 1 implies that both $\mathrm{Min}(A)$ and $\mathrm{Max}(A)$ must be at least $|A| = \binom{m}{k}$. For $k = m/2$, the bound has the form $2^{\Omega(\sqrt{n})}$.

Important in Schnorr's result is that the set A itself does not need to be cover free: it is enough that some of its large subset is cover free *inside* A.

Example 4. Consider the following optimization problem: given an assignment of nonnegative integer weights to the edges of a complete graph K_n on n nodes, find the minimum (or maximum) weight of a subgraph which is either a triangle or a 3-matching (a set of three disjoint edges). Let A be the set of characteristic 0-1 vectors of triangles and 3-matchings; hence, each vector of A has $\binom{n}{2}$ positions (one for each edge of K_n), and exactly three of them are ones. The set A itself is not cover free because a union of two 3-matchings may contain a third 3-matching. But the set B of triangles is already cover free inside A: a union of no two triangles can contain a 3-matching since it contains at most two disjoint edges, and it cannot contain a new triangle because any two edges of a triangle uniquely determine the triangle. Since our problem is homogeneous, Schnorr's theorem and the reduction lemma imply that its tropical complexity is $\Omega(n^3)$.

Using different arguments, Gashkov [11] and Gashkov and Sergeev [12] extended a special case of Schnorr's bound (when A itself is cover-free) to Sidon sets.

THEOREM 2 (Gashkov and Sergeev [12]). *If A is a Sidon set, then* $\mathrm{L}(A) \geq |A|$.

It is clear that every cover-free set is also a Sidon set, but there are Sidon sets which are not cover free; see Appendix D for an example.

In fact, Gashkov and Sergeev [12] proved a more general lower bound. For parameters $1 \leq k \leq l$, a set $A \subset \mathbb{N}^n$ is (k,l)-*sparse* if $|X| \leq k$ or $|Y| \leq l$ holds for any two sets $X, Y \subset \mathbb{N}^n$ of vectors such that $X + Y \subseteq A$. In other words, a set $A \subset \mathbb{N}^n$ is (k,l)-sparse if given any set $X \subset \mathbb{N}^n$ of $|X| = k + 1$ vectors, and any $l + 1$ distinct vectors $y_1, \ldots, y_{l+1}$ in $\mathbb{N}^n$, at least one of the translates $X + y_i = \{x + y_i \colon x \in X\}$ is not contained in A.

The interpretation of sparse sets $A \subset \mathbb{N}^n$ in terms of *graphs* is the following. Associate with A a bipartite graph G_A whose nodes are vectors in $\mathbb{N}^n$, and two nodes x and y are adjacent precisely when $x + y \in A$. Then A is (k,l)-sparse if and only if the graph G_A contains no copy of a complete bipartite $(k+1) \times (l+1)$ graph.

Sparse sets were introduced by Erdős and Harzheim in [6], and they constitute a natural generalization of Sidon sets: every Sidon set is $(1,1)$-sparse. To see this,

suppose that a set A is not $(1,1)$-sparse. Then $\{x,x'\}+\{y,y'\}\subseteq A$ must hold for some vectors $x\neq x'$ and $y\neq y'$. The sum $a+b$ of the two vectors $a=x+y$ and $b=x'+y'$ of A is then equal to the sum $c+d$ of the two vectors $c=x+y'$ and $d=x'+y$ of A. But $c\neq a$ because $y\neq y'$, and $c\neq b$ because $x\neq x'$. So, A is then not a Sidon set.

Thus, Theorem 2 is a special case of the following more general bound.

THEOREM 3 (see [12]). *If A is (k,l)-sparse then* $\mathrm{L}(A)\geq |A|/\max\{k^3,l^2\}$.

Remark 3. An important aspect of this more general theorem is that, together with known constructions of large (k,l)-sparse sets $A\subset\{0,1\}^n$, it yields an almost maximal lower bound $\mathrm{L}(A)\geq 2^{n-o(n)}$ on their Minkowski complexity and, hence, also on the monotone arithmetic circuit complexity of the corresponding multilinear polynomials. Namely, Kollár, Rónyai, and Szabó [21] used so-called norm graphs to construct explicit $(t,t!)$-sparse sets $A\subset\{0,1\}^n$ of size $|A|\geq 2^{n-n/t}$; we sketch their construction in Appendix E.

2.2. Hyafil–Valiant's bound. Another classical lower bound on the Minkowski complexity of vector sets is based on the observation that sets of small Minkowski complexity can be covered by a small number of "balanced" sumsets. This fact (in the context of arithmetic circuits) was first proved by Hyafil [16]; a different and elementary proof was given by Valiant [34]. Various versions of this fact (with different notions of being balanced) were also proved by other authors, including Jerrum and Snir [17] (implicitly), Raz and Yehudayoff [28], Hrubes and Yehudayoff [15].

A sumset $X+Y$ is *balanced* if every vector of X has degree between $m/3$ and $2m/3$, where m is the minimum degree of a vector in $X+Y$. Recall that the *degree* of a vector is the sum of its entries, and that a set of vectors is *homogeneous* of degree m if all its vectors have the same degree m.

THEOREM 4 (Hyafil [16], Valiant [34]). *Let $A\subset\mathbb{N}^n$ be a homogeneous set of degree $m\geq 2$. Then A is a union of at most* $\mathrm{L}(A)$ *balanced sumsets.*

Thus, in order to show that $\mathrm{L}(A)\geq |A|/h$ for a homogeneous set A, it is enough to show that no balanced sumset lying in A can have more than h vectors. There are many applications of this bound. In particular, Grigoriev and Koshevoy [13] have recently used it prove a lower bound on the tropical complexity of Schur polynomials.

Example 5. Let us demonstrate this bound right now on an important example of the assignment problem. Namely, let us show that if $A\subset\{0,1\}^{n^2}$ is the set of characteristic vectors of perfect matchings in $K_{n,n}$ then both $\mathrm{Min}(A)$ and $\mathrm{Max}(A)$ are $2^{\Omega(n)}$. Since the set A is homogeneous (of degree n), it is enough (by the reduction lemma) to show that $\mathrm{L}(A)=2^{\Omega(n)}$. By Theorem 4, it is enough to show the following claim.

- If $X+Y$ is a balanced sumset, and if $X+Y\subseteq A$, then there is an integer $n/3\leq r\leq 2n/3$ such that $|X+Y|\leq n!/\binom{n}{r}$.

To show this, fix arbitrary two vectors $x\in X$ and $y\in Y$. These vectors correspond to matchings such that x has r edges for some $n/3\leq r\leq 2n/3$, and $x+y$ is the characteristic vector of a perfect matching. Since all vectors in $x+Y$ contain all ones (edges) of the matching x and must be perfect matchings themselves, we have $|Y|=|x+Y|\leq (n-r)!$. Similarly, since the matching y has $n-r$ edges, we have $|X|=|X+y|\leq r!$. Thus, $|X+Y|\leq |X|\cdot|Y|\leq r!(n-r)!=n!/\binom{n}{r}$, as claimed.

We stress the simplicity of the proof: by using tighter arguments, Jerrum and Snir [17] have proved that $\mathrm{L}(A)\geq n(2^{n-1}-1)$.

3. Our results. We prove general lower bounds for tropical circuits solving *nonhomogeneous* optimization problems. We obtain these bounds by extending the classical lower bounds on the Minkowski complexity reviewed in the previous section.

3.1. Tropical version of Schnorr's combinatorial bound. Our first results (Theorem B and its special case, Theorem A) extend both Schnorr's and Gashkov–Sergeev's results to larger classes of vector sets; these extensions will be essential in proving over first lower bounds for tropical complexity (see Remark 9).

Say that a subset $B \subseteq A$ of $A \subset \mathbb{N}^n$ is a *Sidon set inside* A if $a+b=c+d$ with $a,b \in B$ and $c,d \in A$ imply $\{c,d\}=\{a,b\}$. That is, now only the sums of vectors in B must be "unique" within A: sums of other elements of A need not be unique.

THEOREM A. *If $B \subseteq A$ is a Sidon set inside A then* $\mathrm{L}(A) \geq |B|/2$.

Remark 4. Numerically, our bound is slightly worse than that in Theorem 1 and 2 (by an additional factor of 1/2). However, our bound holds for a more general Minkowski circuit, where *any* sets $\{x\}$ with $x \in \mathbb{N}^n$ (not only n sets $\{e_1\},\dots,\{e_n\}$) can be used as inputs. In particular, in this case, $\mathrm{L}(A) \leq |A|-1$ is a trivial upper bound: just take a union of all inputs $\{x\}$ with $x \in A$.

Theorem A *properly* extends Theorems 1 and 2 because there are many non-Sidon sets with large Sidon sets inside them.

Example 6. Let $S \subset \mathbb{N}^n$ be any Sidon set, and $T \subset \mathbb{N}^n$ any set which is not a Sidon set. Assume for simplicity that neither S nor T contains the all-0 vector 0. Let $A = B \cup C$ be the union of two sets of vectors in $\mathbb{N}^{2n}$, where B consists of all vectors (x,x) with $x \in S$, and C consists of all vectors $(y,0)$ with $y \in T$. Then B is a Sidon set because S is such a set, but the entire set A is not a Sidon set because already T is not such a set. We claim that B is a Sidon set *inside* A.

To show this, assume contrariwise that a sum $(x,x)+(z,z)$ of two vectors in B is equal to a sum of some other two vectors in A. Since B is a Sidon set, at least one of these two other vectors must belong to C, that is, must be of the form $(y,0)$ with $y \neq 0$. So we have $(x,x)+(z,z)=(y,0)+(u,v)$. If $(u,v) \notin B$ then $v=0$, and we obtain that $x+z=0$, which is not possible because x and z are nonzero vectors. If $(u,v) \in B$ then $v=u$, and we obtain that $x+z=y+u$ and $x+z=0+u$, which is also not possible because $y \neq 0$.

In fact, what we will prove is an extension of Theorem 3 to sets which themselves are not sparse enough but contain large sparse subsets. For this, we introduce the notion of "projection" of sets of vectors onto sumsets.

The *projection* of a set $B \subset \mathbb{N}^n$ of vectors onto the first part X of a sumset $X+Y$ is the set

$$X_B = \{x \in X : x+y \in B \text{ for some } y \in Y\}$$

of all vectors in X that can be extended to a vector in B by adding some vector from the second part Y. The projection Y_B of B onto the second part Y is defined similarly. Note that the "reduced" sumset X_B+Y_B contains all vectors of B which belonged to the "larger" sumset $X+Y$.

For parameters $1 \leq k \leq l$, call a subset $B \subseteq A$ (k,l)-*sparse inside* A if for any two sets $X,Y \subset \mathbb{N}^n$, $X+Y \subseteq A$ implies $|X_B| \leq k$ or $|Y_B| \leq l$.

In Appendix B we will show (see Lemma 10) that a set B is $(1,1)$-sparse inside A if and only if B a Sidon set inside A. So, Theorem A is a special case of the following more general result.

THEOREM B. *If B is (k,l)-sparse inside A then* $\mathrm{L}(A) \geq |B|/2lk^2$.

Our proof is amazingly simple, and is entirely different from those in [29, 12]. The idea—inspired by Pippenger's paper [27]—is to analyze the "progress" made along the source-output paths of a circuit until particular "bottlenecks" are found. We believe that the insights in this "bottleneck counting" argument might be of independent interest.

Our next result is the following extension of Schnorr's lower bound to tropical circuits. Call a subset $B \subseteq A$ of a set $A \subset \mathbb{N}^n$ *noncoverable inside* A if for all $a, b \in A$ and $c \in B$, $a + b \geq c$ implies $c \in \{a, b\}$. Recall that B is *cover free inside* A if this implication holds for all $a, b \in B$ and $c \in A$.

THEOREM C. *Let $A \subset \{0,1\}^n$ be an antichain, and $B \subseteq A$.*
(i) *If B is cover free inside A, then* $\mathrm{Min}(A) \geq |B|/2$.
(ii) *If B is noncoverable inside A, then* $\mathrm{Max}(A) \geq |B|/2$.

Remark 5. Again, these lower bounds hold for a more general model of tropical circuits, where instead of single variables $x_1, \ldots, x_n$ any linear functions $a_1x_1 + \cdots + a_nx_n$ with nonnegative integer coefficients a_i can be used as inputs. In this model, $|A| - 1$ is a trivial upper bound on both $\mathrm{Min}(A)$ and $\mathrm{Max}(A)$; cf. Remark 4.

Example 7. Let $2\sqrt{n} \leq k < n/2$ and consider the following minimization problem: given an assignment of nonnegative integer weights to the edges of K_n, find the minimum weight of a subgraph which is either a k-clique K_k or a star $K_{1,n-1}$. Let A be the set of characteristic 0-1 vectors of these subgraphs. Since $n-1 < \binom{k}{2}$, the lower envelope of A is the set of n stars. So, the reduction lemma cannot yield any larger than n lower bound on $\mathrm{Min}(A)$. Still, Theorem B yields an almost optimal lower bound $\mathrm{Min}(A) = n^{\Omega(k)}$. For this, it is enough to show that the set of (characteristic vectors of) k-cliques is cover free inside A.

To show this, take a union of two k-cliques. Since $2(k-1) < n-1$ and no star in K_k can have more than $k-1$ edges, the union cannot contain a star $K_{1,n-1}$. Also, as we have shown in Exercise 3, the union cannot contain a third k-clique. Thus, the set of k-cliques is cover free inside A, and Theorem C yields $\mathrm{Min}(A) \geq \binom{n}{k}/2$.

Example 8. Let $n > 4$ be a prime power, and consider the complete bipartite $n \times n$ graph $K_{n,n}$ with parts $U = V = \mathrm{GF}(n)$. The graph of a polynomial $g(x)$ is a subgraph of $K_{n,n}$ consisting of n edges $(i, g(i))$ with $i \in \mathrm{GF}(n)$. A *double star* is a $K_{2,n}$ subgraph of $K_{n,n}$. Let $A = B \cup C \subseteq \{0,1\}^{n^2}$, where B is the set of all n^d characteristic vectors of graphs of polynomials of degree at most $d-1$ over $\mathrm{GF}(n)$, and C is the set of all $\binom{n}{2}$ characteristic vectors of double stars. We want to lower bound $\mathrm{Max}(A)$. The upper envelope of A is the set of doubles tars. So, the reduction lemma cannot yield any larger than an $\binom{n}{2}$ lower bound on $\mathrm{Max}(A)$. Still, for every $d \leq n/2$, Theorem B yields an almost optimal lower bound $\mathrm{Max}(A) = n^{\Omega(d)}$. For this, it is enough to show that the set $B \subset A$ of (characteristic vectors) of graphs of polynomials is noncoverable inside A.

To show this, suppose $a + b \geq c$ holds for some $a, b \in A$ and $c \in B$. Hence, c is a (graph of a) polynomial $g(x)$. Since vector c has n ones, it must share at least $n/2$ ones with at least one of the vectors a and b; let it be vector a. This vector cannot be the characteristic vector of a double star because every double star has only $2 < n/2$ non-isolated nodes in U. So, a must be a graph of some polynomial $h(x)$. Since no two distinct polynomials of degree at most $d-1$ can share d or more values in common, and since $n/2 \geq d$, we have that $g = h$ and, hence, also $c = a$. Thus, the set B is noncoverable inside A, and Theorem B yields $\mathrm{Max}(A) \geq |B|/2 = n^d/2$, as desired.

3.2. Tropical version of Hyafil–Valiant's bound. By going deeper into the structure of Minkowski circuits, we will extend Theorem 4 to nonhomogeneous sets.

By a *norm measure* we will mean any assignment of nonnegative real numbers to vectors in $\mathbb{N}^n$ such that every unit 0-1 vector gets norm at most 1, and the norm is subadditive in that the norm of a sum of two vectors does not exceed the sum of their norms. In particular, the degree of vectors (sum of all entries) or their *length* (number of nonzero positions) or their standard ℓ_p-norms are norm measures.

THEOREM D. *Let $m \geq 2$ and $1/m \leq \epsilon < 1$. Suppose that a set $A \subset \mathbb{N}^n$ can be created by a Minkowski circuit with t sumset gates. Then there exist t sumsets $X + Y \subseteq A$ with the following property:*

(∗) *for every norm measure $\mu : \mathbb{N}^n \to \mathbb{R}_+$, and for every vector $a \in A$ of norm $\mu(a) \geq m$, at least one of these sumsets $X + Y$ contains vectors $x \in X$ and $y \in Y$ such that $x + y = a$ and $\epsilon m/2 \leq \mu(x) \leq \epsilon m$.*

By taking one norm measure for all vectors $a \in A$, we obtain the following handy version of this theorem. A sumset $X + Y$ is (m, ϵ)-*balanced* (with respect to a given norm measure of vectors) if every vector in X has norm between $\epsilon m/2$ and ϵm.

COROLLARY 5. *Let $m \geq 2$ and $1/m \leq \epsilon < 1$. Suppose that a set $A \subset \mathbb{N}^n$ can be created by a Minkowski circuit with t sumset gates. Then there exist at most t (m, ϵ)-balanced sumsets $X + Y \subseteq A$ such that every vector in A of norm at least m belongs to at least one of these sumsets.*

Proof. Fix a norm measure, and a set of at most t sumsets $X+Y \subseteq A$ guaranteed by Theorem D (for this one fixed measure). To make these sumsets (m, ϵ)-balanced, remove from X all vectors whose norm is either smaller than $\epsilon m/2$ or is larger than ϵm. ∎

The main difference of Corollary 5 from Theorem 4 is that now the set A needs not be homogeneous (vectors of A do not need to have the same norm), and we still have information about the norms of *individual* vectors in one part of the sumsets.

Using Corollary 5, we will prove our second lower bound on tropical complexity. A sumset $R = X+Y$ is *orthogonal* if $\langle x, y\rangle = 0$ holds for all vectors $x \in X$ and $y \in Y$. A sumset is *strongly* (m, ϵ)-*balanced* if it is orthogonal, and all vectors in X have the same length lying between $\epsilon m/2$ and ϵm. Recall that the *length* of a vector is the number of its nonzero entries.

THEOREM E. *Let $m \geq 2$ and $1/m \leq \epsilon < 1$. Let $A \subseteq \{0, 1\}^n$ be an antichain, and let $B \subseteq A$ be the set of all vectors of A with exactly m ones.*

(i) *If no vector of A with fewer than m ones is contained in any vector of the sumset $B + B$, then B is a union of at most* $\mathrm{Min}(A)$ *strongly (m, ϵ)-balanced sumsets.*

(ii) *If no vector of A with more than m ones shares $\epsilon m/2$ or more ones with any vector of B, then B is a union of at most* $\mathrm{Max}(A)$ *strongly (m, ϵ)-balanced sumsets.*

Example 9. Consider the following *minimization* problem: given an assignment of nonnegative integer weights to the edges of $K_{n,n}$, find the minimum weight of a subgraph which is either a perfect matching or a 3-star $K_{1,3}$. Let A be the set of characteristic 0-1 vectors of these subgraphs. Since the *lower* envelope of A is the set of 3-stars, the reduction lemma cannot yield any larger lower bound on $\mathrm{Min}(A)$ than $n\binom{n}{3}$. On the other hand, since no 3-star can be contained in a union of two perfect matchings, the first condition (i) of Theorem E—with B being the set of characteristic

vectors of perfect matchings—is fulfilled. Together with Example 5, Theorem E yields $\mathrm{Min}(A) = 2^{\Omega(n)}$.

Example 10. Consider the following *maximization* problem: given an assignment of nonnegative integer weights to the edges of $K_{n,n}$ for $n > 6$, find the maximum weight of a subgraph which is either a perfect matching or a double star $K_{2,n}$. Let A be the set of characteristic 0-1 vectors of these subgraphs. Since the *upper* envelope of A is the set of double stars, the reduction lemma cannot yield any larger lower bound on $\mathrm{Max}(A)$ than $\binom{n}{2}$. On the other hand, since no double star can share more than $2 < n/3$ edges with a perfect matching, the second condition (ii) of Theorem E (with B being the set of characteristic vectors of perfect matchings) is fulfilled. Together with Example 5, Theorem E yields $\mathrm{Max}(A) = 2^{\Omega(n)}$.

4. Structure of Minkowski circuits. The goal of this section is to establish some basic structural properties of Minkowski $(\cup, +)$ circuits. At a node v of a given Minkowski circuit, some subset $X_v \subset \mathbb{N}^n$ of vectors is *created* in a natural way:

- $X_v = \{e_i\}$ if v is an input node holding a set $\{e_i\}$;
- $X_v = X_u \cup X_w$ if $v = u \cup w$ is a union gate;
- $X_v = X_u + X_w$ if $v = u + w$ is a sumset gate.

The set created by the entire circuit is the set created at its output gate. The *size* of the circuit is the total number of its gates. For a set $A \subset \mathbb{N}^n$, $\mathrm{L}(A)$ denotes the minimum size of a Minkowski circuit producing A.

4.1. Contents of nodes and edges. Fix a Minkowski $(\cup, +)$ circuit, and let $A \subset \mathbb{N}^n$ be the set of vectors created by it. We associate with every node v of the circuit the following three sets of vectors.

- X_v is the set of vectors created at gate v (as defined above).
- $Y_v = \{y \in \mathbb{N}^n : x + y \in A \text{ for all } x \in X_v\}$ is the *complement* of gate v.
- $A_v = X_v + Y_v$ is the *content* of gate v.

Note that X_v doesn't need to lie in A, but at least one of its translates $X_v + y$ must already lie in A; the complement Y_v of node v collects all such vectors y. Speaking informally, once a vector has been created, it must find its way into the final result A. This property of Minkowski circuits severely limits their power, and was essentially exploited in all known proofs of lower bounds on $\mathrm{L}(A)$.

At a source node v holding a singleton $\{e_i\}$, the set X_v is small and Y_v is large: $X_v = \{e_i\}$ and $Y_v = \{a - e_i : a \in A, a_i \neq 0\}$. At the output gate v, the situation is just the opposite: then $X_v = A$ is large and $Y_v = \{(0, \ldots, 0)\}$ is small. The following simple lemma describes the "dynamics" of contents of gates as computation proceeds.

LEMMA 3 (content lemma).

- *If $v = u \cup w$ is a union gate, then $X_v = X_u \cup X_w$ and $Y_v = Y_u \cap Y_w$.*
- *If $v = u + w$ is a sumset gate, then $X_v = X_u + X_w$,*

$$X_w + Y_v \subseteq Y_u, \quad \text{and} \quad X_u + Y_v \subseteq Y_w\,. \tag{2}$$

Proof. Only the inclusions (2) need a certification. To show $X_w + Y_v \subseteq Y_u$, suppose contrariwise that there are vectors $x \in X_w$ and $y \in Y_v$ for which $x + y$ does not belong to Y_u. By the definition of the complement Y_u of gate u, there must be a vector $x' \in X_u$ for which $x' + (x + y)$ does not belong to A. But this is impossible because $x' + x$ belongs to the set $X_v = X_u + X_w$ created at the gate v, and y is in the complement Y_v of this gate. ∎

An example in Figure 1 shows that the converse inclusions in (2) do *not* need to hold.

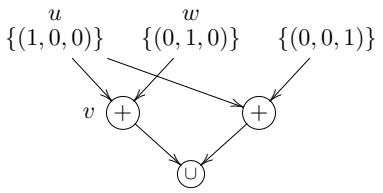

FIG. 1. *An example showing that the inclusions in* (2) *cannot be reversed. The set created by the circuit is* $A = \{(1,1,0),(1,0,1)\}$. *Nodes* u *and* w *are source nodes. The complements of gates* u *and* v *are* $Y_u = \{(0,1,0),(0,0,1)\}$ *and* $Y_v = \{(0,0,0)\}$. *But the vector* $(0,0,1)$ *of* Y_u *does not belong to* $X_w + Y_v = \{(0,1,0)\}$.

Our definition of the contents A_e of *edges* $e = (u, v)$ is already "gate sensitive": it depends on the operation (union or sumset) associated with its head v.

- If v is a sumset gate, then let $A_e = X_v + Y_v$.
- If v is a union gate, then let $A_e = X_u + Y_v$.

Note that $A_e \subseteq A$ also holds in this latter case because then $Y_v \subseteq Y_u$ and $A_u = X_u + Y_u \subseteq A$. The reason for this "asymmetry" in our definition of contents of edges is explained by the following lemma.

LEMMA 4 (content propagation lemma). *Let* v *be gate, and* a *be a vector in its content.*

- *If* v *is a union gate, then vector* a *belongs to the content of* at least one *edge entering* v*, as well as to the content of the tail of this edge.*
- *If* v *is a sumset gate, then vector* a *belongs to the contents of* both *edges entering* v*, as well as to the contents of the tails of these edges.*

Proof. First let $v = u \cup w$ be a union gate. By our assumption, the vector a belongs to the content $A_v = X_v + Y_v = (X_u \cup X_w) + Y_v$ of gate v. This immediately implies that a must belong at least one of the contents $X_u + Y_v$ or $X_w + Y_v$ of the edges entering gate v. Also, in this case we have that $Y_v = Y_u \cap Y_w$. Thus, a must then also belong to the content $A_u = X_u + Y_u$ or $A_w = X_w + Y_w$ of the tail u or w of the corresponding edge.

Now let $v = u + w$ be a sumset gate. Since our vector a belongs to the content A_v of gate v, it must also belong to the contents of both edges entering this gate (just because these contents coincide with A_v in this case). Moreover, (2) yields

$$A_v = X_v + Y_v = X_u + (X_w + Y_v) \subseteq X_u + Y_u = A_u$$

and, similarly, $A_v \subseteq A_w$. Since vector a belongs to A_v, this implies that it must also belong to the contents of both gates entering v. ∎

4.2. Traces of vectors. Take a Minkowski $(\cup, +)$ circuit, and let $A \subset \mathbb{N}^n$ be the set of vectors created at its output gate. A *trace* in the circuit is a subgraph of the underlying directed acyclic graph obtained by removing exactly one edge entering each union gate (see Figure 2). Traces can be also defined inductively as follows. The output gate is included in every trace. Let v be a gate already included in a trace.

- If v is a union $(\cup)$ gate, then *exactly one* of its inputs is included in the trace.
- If v is a sumset $(+)$ gate, then *both* its inputs are included in the trace.

If we add an empty set as a missing input to each union gate in a trace, then the obtained Minkowski circuit will create exactly one vector. Intuitively, a trace of a vector charts the creation of this vector by the circuit. If a vector has more than one trace, this means that this vector was created in more than one way.

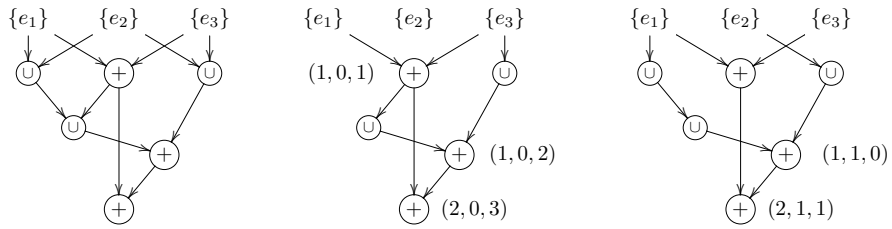

FIG. 2. *A Minkowski circuit with three source nodes holding the unit vectors* $e_1 = (1,0,0)$, $e_2 = (0,1,0)$, *and* $e_3 = (0,0,1)$, *and two its traces creating the vectors* $(2,0,3)$ *and* $(2,1,1)$. *The circuit itself, creates the set* $A = \{(2,1,1),(1,2,1),(2,1,2),(2,0,2),(1,1,2),(2,0,3)\}$.

Remark 6. The concept of traces was introduced by Jerrum and Snir [17] who named them "parse trees." If a vector is a 0-1 vector, then its trace must indeed be a tree, but this does not need to hold for vectors having entries larger than 1. This is why we prefer the term "trace."

We say that a trace in the circuit is a *trace of* a vector $a \in A$, if this vector belongs to the contents of all edges and all gates of the trace.

LEMMA 5 (trace lemma). *If A is the set of vectors created by a Minkowski circuit, then every vector of A has its trace in the circuit.*

Proof. Take a vector $a \in A$. A trace T_a for a can be constructed by using the content propagation lemma: start at the output gate of the circuit (whose content coincides with A and, hence, contains vector a), and construct T_a backwards by the following rule. If a sumset gate v is included in T_a, then both its inputs are included T_a. If a union gate v is included in T_a, then include in T_a that of its inputs whose content contains vector a; if vector a belongs to the contents of both inputs, then include any of them. ∎

Remark 7. By the trace lemma, we can associate with every vector $a \in A$ a trace T_a in the circuit G such that the content A_e of each edge e of this trace contains vector a. This yields $|\{a \in A\colon e \in T_a\}| \le |A_e|$ for every edge e of G. So, if the circuit G has t edges, then double counting gives

$$|A| \cdot \min_{a\in A} \sum_{e\in T_a} \frac{1}{|A_e|} \le \sum_{a\in A} \sum_{e\in T_a} \frac{1}{|A_e|} = \sum_{e\in G} \sum_{a\in A\colon\, e\in T_a} \frac{1}{|A_e|} \le \sum_{e\in G} 1 = t\,.$$

In particular, if at least one trace of each vector $a \in A$ has an edge e with $|A_e| \le h$, then the circuit must have $t \ge |A|/h$ edges.

We now have all we need to prove our main results, Theorems B–E.

5. Proof of Theorem B. Let $A \subset \mathbb{N}^n$, and let $B \subseteq A$ be (k,l)-sparse inside A. Take a Minkowski $(\cup,+)$ circuit producing the set A. Our goal is to prove that then the circuit must have $t \ge |B|/2lk^2$ gates. We can assume that $|B| > k$, for otherwise there would be nothing to prove.

The proof idea is to use the sparseness of B inside A to associate with every vector $a \in B$ an edge e in the circuit such that the content A_e of e contains vector a, but does not contain "too many" vectors of B. This implies that there must be many edges and, hence, also gates in the circuit.

Recall that the content A_e of an edge $e = (u,v)$ is the sumset $A_e = X_u + Y_v$ if v is a union $(\cup)$ gate, and is the sumset $A_e = X_v + Y_v$ (the content of gate v) if

v is a sumset (+) gate. Here, as before, X_v is the set of vectors created at v, and $Y_v = \{y \in \mathbb{N}^n : X_v + y \subseteq A\}$ is the complement of gate v.

Call an edge e *light* if its content A_e contains at most lk^2 vectors of B.

LEMMA 6. *Every vector of B is contained in the content of at least one light edge.*

Note that this lemma already yields Theorem B: since the content of no light edge can have more than lk^2 vectors of B, there must be at least $|B|/lk^2$ edges in the circuit and, hence, at least $|B|/2lk^2$ gates, because each gate has indegree two.

Proof (of Lemma 6). For a gate v, let X'_v and Y'_v be the projections of the set B onto the corresponding parts of the sumset $X_v + Y_v$. That is, X'_v consists of all vectors $x \in X_v$ such that $x + y \in B$ holds for at least one vector $y \in Y_v$, and similarly for Y'_v. Define the *cost* of a node v to be the number $|X'_v|$ of vectors in the projection X'_v of B onto X_v. Call a node *cheap* if its cost is at most k, and *expensive* otherwise.

For the ith source node v, we have $X_v = \{e_i\}$. Since clearly $|X'_v| \le |X_v| = 1 \le k$, every source node is cheap. On the other hand, for the output gate w, we have $X_w = A \supseteq B$ and $Y_w = \{0\}$; hence, $X'_w = B$. Since we assumed that $|B| > k$, the output gate is expensive.

Fix now a vector $a \in B$. We have to show that a must belong to the content of at least one light edge. By the trace lemma, there is a trace in the circuit, the contents A_e of *all* whose edges e contain this vector a. So, it is enough to show that at least one edge of the trace must be light.

Start at the output gate of the trace (which is also the output gate of the entire circuit), and traverse a path P in the trace by going backwards and using the following rule, where v is the last already reached gate:

- If v is a union gate, then go to the (unique) gate entering v in the trace.
- If v is a sumset gate, then go to any of its two inputs if they both are expensive or both are cheap, and go to the expensive input if the second input is cheap.

Since the output gate is expensive and every source node is cheap, we will eventually reach some source node (in this backward run). Since the first (source) node of the corresponding source-output path P in the circuit is cheap, and the last one is expensive, there must be an edge $e = (u, v)$ in P such that $|X'_u| \le k$ but $|X'_v| > k$. It remains to show that the edge e must be a light edge.

Since $X_v + Y_v \subseteq A$ and $|X'_v| > k$, the (k, l)-sparseness of B inside A implies that $|Y'_v| \le l$ must hold. Thus, we have the following information about the found edge $e = (u, v)$:

$$|X'_u| \le k \quad \text{and} \quad |Y'_v| \le l\,. \tag{3}$$

We now consider two cases depending on what operation is performed at gate v.

Case 1: $v = u \cup w$ is a union gate. In this case, the content of edge e is $A_e = X_u + Y_v$. By (3), it is enough to show that then $B \cap A_e$ is contained in the sumset $X'_u + Y'_v$, which yields $|B \cap A_e| \le |X'_u + Y'_v| \le |X'_u| \cdot |Y'_v| \le kl$, meaning that the edge e is light.

To show the inclusion $B \cap A_e \subseteq X'_u + Y'_v$, take a vector $a \in B$ which belongs to the content $A_e = X_u + Y_v$ of edge e. Then $a = x + y$ for some $x \in X_u \subseteq X_v$ and $y \in Y_v \subseteq Y_u$. But then we also have that $x \in X'_u$ (since $y \in Y_u$) and $y \in Y'_v$ (since $x \in X_v$). Thus, vector a belongs to the reduced content $X'_u + Y'_v$ of the edge e, as desired.

Case 2: $v = u + w$ is a sumset gate. In this case, the content of edge e is $A_e = X_v + Y_v = X_u + X_w + Y_v$. Since the gate u is cheap (it has $|X'_u| \le k$), step 2

in the construction of the path implies that the second node w entering v must be also cheap, that is, $|X'_w| \le k$ must hold as well. It remains therefore to show that $B \cap A_e$ is contained in $X'_u + X'_w + Y'_v$. Together with (3) and $|X'_w| \le k$, this will yield $|B \cap A_e| \le k^2 l$, as desired.

To show the inclusion $B \cap A_e \subseteq X'_u + X'_w + Y'_v$, take a vector $a \in B$ which belongs to the content $A_e = X_v + Y_v = X_u + X_w + Y_v$ of edge e. Then $a = x_u + x_w + y$ for some $x_u \in X_u$, $x_w \in X_w$, and $y \in Y_v$. By the inclusions (2), we have that $x_w + y \in Y_u$ and $x_u + y \in Y_w$. Hence, $x_u \in X'_u$ and $x_w \in X'_w$. Since clearly, $x_u + x_w \in X_v$, we also have that $y \in Y'_v$. Thus, the vector $a = x_u + x_w + y$ belongs to the reduced content $X'_u + X'_w + X'_v$ of the edge e, as desired.

This completes the proof of Lemma 6 and, thus, also the proof of Theorem B. ∎

Remark 8. Note that we actually proved a stronger result: the lower bound holds for more general $(\cup, +)$ circuits, where *any* sets $X \subset \mathbb{N}^n$ of $|X| \le k$ vectors can be used as inputs. This holds because the source nodes are then still cheap.

6. Proof of Theorem C. Let $A \subset \{0,1\}^n$ be an antichain and $B \subseteq A$. The proof idea in both cases (minimization and maximization) is similar: the structural lemma gives us a set F of vectors (with particular structural properties) whose Minkowski complexity $\mathrm{L}(F)$ coincides with the tropical complexity of A. We then use the properties of the subset B (cover free inside A or noncoverable inside A) to show that the set B must be a Sidon set inside F. Theorem A then yields the desired lower bound $\mathrm{L}(F) \ge |B|/2$ on the Minkowski complexity of F and, hence, also on the tropical complexity of A.

Proof for $(\min, +)$ *circuits.* Suppose that B is cover free inside A, that is, for any $a, b \in B$ and $c \in A$, $a + b \ge c$ implies $c \in \{a, b\}$. Our goal is to show that then $\mathrm{Min}(A) \ge |B|/2$. By the structural lemma, we know that $\mathrm{Min}(A) = \mathrm{L}(F)$ must hold for some set $F \subset \mathbb{N}^n$ of vectors such that $A \subseteq F$ and F lies above A. In particular, we also have $B \subseteq F$. By Theorem A, it is enough to show that then the set B must be a Sidon set inside F.

Suppose contrariwise that there are vectors $a, b \in B$ and $c, d \in F$ such that $a + b = c + d$ but $c, d \notin \{a, b\}$. Since F lies above A, there must be vectors $x, y \in A$ such that $c \ge x$ and $d \ge y$. As B is cover free inside A, $a + b \ge c \ge x$ implies $x \in \{a, b\}$. Assume without loss of generality (w.l.o.g.) that $x = a$; then $c \ge a$. Since $c \notin \{a, b\}$, we have that $c > a$ (a proper inequality). Together with $a + b = c + d$, this yields $d < b$ and, hence, also $y \le d < b$. But then one vector y of A is contained in another vector b of A, a contradiction with A being an antichain. ∎

Remark 9. Note that B needs not be *cover free* inside F. Since, by our assumption, B is a cover free inside A, $a + b \ge c \ge x$ with $a, b \in B$, $c \in F$, and $x \in A$ imply that $x \in \{a, b\}$. But vector c itself doesn't need to belong to $\{a, b\}$: it could then be any vector satisfying $a < c \le a + b$. Thus, usage of Theorem A was essential: the Schnorr's bound itself does not yield Theorem C.

Proof for $(\max, +)$ *circuits.* The argument is similar to that for the $(\min, +)$ case. Suppose that B is noncoverable inside A, that is, for any $x, y \in A$ and $a \in B$, $x + y \ge a$ implies $a \in \{x, y\}$. Our goal is to show that then $\mathrm{Max}(A) \ge |B|/2$. By the structural lemma, we know that $\mathrm{Max}(A) = \mathrm{L}(F)$ must hold for some set $F \subset \{0,1\}^n$ of vectors such that $A \subseteq F$ and F lies below A. By Theorem A, it is enough to show that then B must be a Sidon set inside F.

Suppose contrariwise that B is not a Sidon set inside F. Then there are vectors $a, b \in B$ and $c, d \in F$ such that $a + b = c + d$ but $c, d \notin \{a, b\}$. Since F lies below

A, there must be vectors $x, y \in A$ such that $c \leq x$ and $d \leq y$. As B is noncoverable inside A, $x + y \geq a$ implies $a \in \{x, y\}$. Assume w.l.o.g. that $a = x$; then $a \geq c$. Since $c \notin \{a, b\}$, this yields a strong inequality $a > c$ and, hence, also $b < d \leq y$. But then one vector y of A contains another vector b of A, a contradiction with A being an antichain. ∎

7. Proof of Theorem D. Let $m \geq 2$ and $1/m \leq \epsilon < 1$. Suppose that a set $A \subset \mathbb{N}^n$ can be created by a Minkowski circuit with t sumset $(+)$ gates. Our goal is to show that then there exist t sumsets $X + Y \subseteq A$ with the following property:

$(*)$ for every norm measure $\mu : \mathbb{N}^n \to \mathbb{R}_+$, and for every vector $a \in A$ of norm $\mu(a) \geq m$, at least one of these sumsets $X + Y$ contains vectors $x \in X$ and $y \in Y$ such that $x + y = a$ and $\epsilon m/2 \leq \mu(x) \leq \epsilon m$.

Recall that a function $\mu : \mathbb{N}^n \to \mathbb{R}_+$ is a *norm measure* if $\mu(e_i) \leq 1$ holds for each of the n unit 0-1 vectors e_i, and $\mu(x + y) \leq \mu(x) + \mu(y)$ holds for all vectors $x, y \in \mathbb{N}^n$.

Fix a Minkowski circuit of size t producing A. This circuit gives us t sumsets $X_v + Y_v \subseteq A$, the contents of sumset gates v. So, it is enough to show that these sumsets already have the desired property $(*)$.

To show this, fix some norm measure $\mu : \mathbb{N}^n \to \mathbb{R}_+$, and a vector $a \in A$ of norm $\mu(a) \geq m$. By the trace lemma, there is a trace T_a in the circuit such that vector a belongs to the contents $X_v + Y_v$ of all nodes v in this trace. That is, for every node v in T_a, there is a vector $x \in X_v$ such that $a = x + y$ for some vector $y \in Y_v$. The decomposition $a = x + y$ may not be unique. So, define the *weight* of the gate v to be the maximum norm

$$l_v = \max\{\mu(x) \colon x \in X_v \text{ and } x + y = a \text{ for some } y \in Y_v\}$$

of a vector $x \in X_v$ in such a decomposition of a. We give zero weight $l_w = 0$ to all nodes w in the circuit whose contents do not contain vector a.

Recall that a trace T_a for a given vector a is constructed by using the content propagation lemma: start at the output gate of the circuit, and construct T_a backwards by the following rule. If a sumset gate v is included in T_a, then both its inputs are included T_a. If a union gate v is included in T_a, then include in T_a that of its inputs whose content contains vector a. This latter rule does not specify what to do when vector a belongs to the contents of *both* inputs of a union gate. For the current proof, we will assume that, in such an ambiguous situation, the choice is "greedy": that of the two inputs is included whose weight is larger; if both inputs have the same weight, then any of them can be included in the trace.

Under this proviso, the weight measure of gates has the following properties.

CLAIM 7. *Let v be a gate with inputs u and w in the circuit. Suppose that gate v belongs to the trace T_a.*

- *Subadditivity: if v is a sumset gate, then $l_v \leq l_u + l_w$.*
- *Monotonicity: if v is a union gate and gate u is in the trace, then $l_v \leq l_u$.*

Proof. First let v be a sumset gate. Then both inputs u and w belong to the trace. Let $x \in X_v$ be a vector of norm $\mu(x) = l_v$ such that $x + y = a$ holds for some $y \in Y_v$. Since x belongs to $X_v = X_u + X_w$, there are vectors $x_u \in X_u$ and $x_w \in X_w$ such that $x = x_u + x_w$ and, hence, also $x_u + x_w + y = a$. By the inclusions (2), vector $x_w + y$ belongs to Y_u, and vector $x_u + y$ belongs to Y_w, implying that $\mu(x_u) \leq l_u$ and $\mu(x_w) \leq l_w$. Hence, $l_v = \mu(x) = \mu(x_u + x_w) \leq \mu(x_u) + \mu(x_w) \leq l_u + l_w$.

Now let v be a union gate, and u be its input belonging to the trace. If vector a does not belong to the content of the second input w of v, then $l_w = 0$. Otherwise, we still have $l_u \leq l_w$, due to our greedy choice when constructing the trace T_a. Now

let $x \in X_v$ be a vector of norm $\mu(x) = l_v$ such that $x + y_x = a$ holds for some $y_x \in Y_v$. Since in this (union gate) case we have $Y_v = Y_u \cap Y_w$, vector y_x belongs to both Y_u and Y_w. Vector x must belong to X_u or to X_w. If $x \in X_u$ then $l_u \geq \mu(x) = l_v$. If $x \in X_w$ then $l_w \geq \mu(x) = l_v$ and, hence, also $l_u \geq l_v$, as desired. ∎

CLAIM 8. *There is a sumset gate in T_a whose weight lies between $\epsilon m/2$ and ϵm.*

Proof. The weight of the output gate is the norm $\mu(a) \geq m$ of the vector a itself. So, the output gate has weight strictly larger than ϵm (because $\epsilon < 1$). On the other hand, the weight of the ith source node is $\mu(e_i) \leq 1$, which is at most ϵm (because $\epsilon \geq 1/m$). So, we can start at the output gate and traverse the trace backwards until a (sumset) gate v of weight $> \epsilon m$ is found such that the weights of both its inputs are at most ϵm; this must be a sumset gate, due to the monotonicity of the weight at union gates. Since the weight of gate v is $> \epsilon m$, the subadditivity of weight implies that the weight of at least one of these inputs must be $> \epsilon m/2$, as desired.

This completes the proof of Claim 8 and, hence, also the proof Theorem D. ∎

Remark 10. This "$\frac{1}{3}$–$\frac{2}{3}$ trick" (in the case when $\epsilon = 2/3$) is standard, and was already used in the previous proofs of Theorem 4 and its variants [16, 34, 17, 28, 15]. The novelty of our argument is that we applied this trick to *traces* of individual vectors $a \in A$ in the circuit, not just to the entire circuit. This allowed us to put individual vectors into their balanced sumsets.

8. Proof of Theorem E. Recall that the *length* $|a|$ of a vector a is the number of its nonzero entries. Let $A \subseteq \{0,1\}^n$ be an antichain and $B = \{a \in A\colon |a| = m\}$ be the set of all vectors in A of length exactly m; note that the length $|a|$ of a 0-1 vector is just the number of ones in it.

Let $m \geq 2$ and $1/m \leq \epsilon < 1$. Recall that a sumset $R = X + Y$ is *orthogonal* if $\langle x, y\rangle = 0$ holds for all vectors $x \in X$ and $y \in Y$. A sumset is *strongly (m,ϵ)-balanced* if it is orthogonal, and all vectors in X have the same length lying between $\epsilon m/2$ and ϵm. Our goal is to prove the following two claims.

(i) If no vector of A with fewer than m ones is contained in any vector of the sumset $B + B$, then B is a union of at most $\mathrm{Min}(A)$ strongly (m,ϵ)-balanced sumsets.

(ii) If no vector of A with more than m ones shares $\epsilon m/2$ or more ones with any vector of B, then B is a union of at most $\mathrm{Max}(A)$ strongly (m,ϵ)-balanced sumsets.

To do this, suppose that an optimization problem (minimization or maximization) on A can be solved by a tropical circuit of size t. Then, by the structural lemma, there exists a set $F \subset \mathbb{N}^n$ of vectors such that $A \subseteq F$, F can be created by a Minkowski circuit of size at most t, and F either lies above A if we work with $(\min, +)$ circuits, or lies below A if we work with $(\max, +)$ circuits.

By Corollary 5 (when we take the length of vectors as their norm measure), there exist at most t (m,ϵ)-balanced sumsets $X + Y \subseteq F$ such that every vector in F of length at least m and, hence, also for every vector $a \in B$ (since $B \subseteq F$), belongs to at least one of these sumsets. So fix such an (m,ϵ)-balanced sumset $X + Y \subseteq F$. The theorem will follow from the following claim.

CLAIM 9. *Let $a, x \in X$ and $b, y \in Y$ be such that both vectors $a + b$ and $x + y$ belong to B. Then* (i) $|a| = |x|$, (ii) $\langle a, y\rangle = 0$, *and* (iii) $a + y \in B$.

Having these three properties, we can obtain the desired strongly (m,ϵ)-balanced sumset $X' + Y'$ by letting $X' \subseteq X$ and $Y' \subseteq Y$ be the projections of the set B onto the parts X and Y. That is, X' consists of all vectors $x \in X$ such that $x + y \in B$

holds for at least one vector $y \in Y$, and similarly for Y'. By (i), the sumset $X' + Y'$ is strongly (m, ϵ)-balanced and, by (ii), it is orthogonal. Moreover, we have the inclusions $B \cap (X + Y) \subseteq X' + Y' \subseteq B$, where the first inclusion is trivial, and the second follows from property (iii). Since the t original sumsets $X + Y$ contained all vectors of B, the inclusions imply that B is actually a union of the reduced sumsets $X' + Y'$, as desired.

So, it remains to prove Claim 9.

Take any vectors $a, x \in X$ and $b, y \in Y$ such that both vectors $a + b$ and $x + y$ belong to B. Since both $a+b$ and $x+y$ must be 0-1 vectors (they belong to B) we have $\langle a, b\rangle = \langle x, y\rangle = 0$ (vectors a and b, as well as vectors x and y must be orthogonal). Hence, we also have $|a + b| = |a| + |b| = m$ and $|x + y| = |x| + |y| = m$. Call a vector *short* if its length is $< m$, and *long* if its length is $> m$.

Proof for $(\min, +)$ *circuits.* Since the set F in this case must lie *above* A, and since the "mixed" vector $a + y$ belongs to F, there must be a vector $c \in A$ such that $a + y \geq c$. Our assumption in this case is that no vector of $B + B$ can contain any vector of A shorter than m. In particular, this implies that

$$|a + y| \geq |c| \geq m\,. \tag{4}$$

Indeed, were vector $a + y$ short (shorter than m) then vector c would also be short. But then the sum $(a + b) + (x + y)$ of two vectors in B would contain a short vector c of A, contradicting our assumption.

To prove claim (i), assume contrariwise that $|a| < |x|$. Then $|a + y| \leq |a| + |y| = |a| + (m - |x|) < m$, a contradiction with (4). So, $|a| = |x|$. Hold now $\langle a, y\rangle \neq 0$, then vectors a and y would share a common nonzero position, implying that $|a+y| \leq |a| + |y| - 1 = |x| + |y| - 1 = m - 1$, a contradiction with (4). Hence, claim (ii) also holds.

By claims (i) and (ii), we have $|a + y| = |a| + |y| = |x| + |y| = m$. Together with $a + y \geq c$ and $|c| \geq m$, we have that $a + y = c$ and $|c| = m$. But since B contains *all* vectors in A of length m (including vector c) this means that $a + y$ must belong to B, as desired. ∎

Proof for $(\max, +)$ *circuits.* Since the set F in this case must lie *below* A, and the mixed vector $a + y$ belongs to F, there must be a vector $c \in A$ such that $a + y \leq c$. So, since a is a 0-1 vector, claim (ii) trivially holds in this case. In particular, we have $|a + y| = |a| + |y|$. To prove claims (i) and (iii), we will use our assumption that the scalar product of a vector in B with any long vector in A (if there are any) must be smaller than $\epsilon m/2$. Since the sumset $X + Y$ was (m, ϵ)-balanced, and $a \in X$, we also know that $|a| \geq \epsilon m/2$.

To prove claim (i), assume contrariwise that $|a| > |x|$. Then $|a + y| = |a| + |y| = |a| + (m - |x|) > m$, meaning that $a + y$ is a long vector. Since $a + y \leq c$, vector $c \in A$ is then also a long vector. But then the scalar product $\langle a + b, c\rangle \geq \langle a + b, a + y\rangle \geq |a| \geq \epsilon m/2$ of a vector $a + b$ in B with a long vector c in A is not smaller than $\epsilon m/2$, a contradiction.

To prove claim (iii), assume contrariwise that $a + y \notin B$. By claims (i) and (ii), the vector $a + y$ is a 0-1 vector of length $|a| + |y| = |x| + |y| = m$. Since B includes *all* vectors of A of length m, and since $a + y \notin B$, the vector $c \geq a + y$ of A must be a long vector. But then, again, we have that the scalar product $\langle a + b, c\rangle$ of a vector $a + b$ in B with a long vector c in A is not smaller than $\epsilon m/2$, a contradiction. ∎

Appendix A. Proof of the structural lemma. We will prove this lemma only in the case of (min, +) circuits; the proof for (max, +) circuits is similar. Our

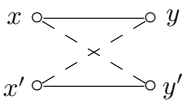

FIG. 3. *Illustration for Lemma* 10. *Dashed lines are edges of the graph G_A of A, nondashed are those of its subgraph G_B.*

goal is to show that for any antichain $A \subset \{0,1\}^n$, the tropical complexity $\mathrm{Min}(A)$ of A coincides with the minimum of the Minkowski complexity $\mathrm{L}(F)$ of a set $F \subset \mathbb{N}^n$ such that $A \subseteq F$ and F lies above A. Since the inequality $\mathrm{Min}(A) \leq \mathrm{L}(F)$ is trivial (take $F = A$), it is enough to show that there exists a set F with these two properties for which $\mathrm{Min}(A) \geq \mathrm{L}(F)$.

Proof. Take a $(\min, +)$ circuit of size $\mathrm{Min}(A)$ solving the minimization problem $\min_A(x) = \min\{\langle a, x\rangle : a \in A\}$ on A. Let $F \subset \mathbb{N}^n$ be the set of vectors created by the Minkowski version of this circuit. Hence, $\mathrm{L}(F) \leq \mathrm{Min}(A)$. We known that $\min_F(x) = \min_A(x)$ must then hold for all input weightings $x \in \mathbb{N}^n$.

To show that F must lie above A, assume contrariwise that there is a vector $b \in F$ such that $b \not\geq a$ for all vectors $a \in A$. Take an assignment x such that $x_i = 0$ for all $i \in S_b$, and $x_i = 1$ for all $i \notin S_b$; recall that $S_b = \{i\colon b_i \neq 0\}$ is the support of vector b. On this weighting x, we have $\min_F(x) \leq \langle b, x\rangle = 0$. But by our assumption, $S_a \setminus S_b \neq \emptyset$ holds for all vectors $a \in A$, and hence, $\langle a, x\rangle \geq 1 > \min_F(x)$ holds for all these vectors, a contradiction with $\min_F(x) = \min_A(x)$.

To show that $A \subseteq F$, suppose contrariwise that there is a vector $a \in A \setminus F$. Consider an input weighting x with $x_i = 1$ for $i \in S_a$, and $x_i = n+1$ for $i \notin S_a$. Then $\min_A(x) \leq \langle a, x\rangle = \langle a, a\rangle \leq n$. To get a desired contradiction, it is enough to show that $\langle b, x\rangle > \langle a, a\rangle$ holds for all $b \in F$.

Case 1: $S_b \not\subseteq S_a$. Then $\langle b, x\rangle \geq n + 1 > \langle a, a\rangle$.

Case 2: $S_b = S_a$. Since a is a 0-1 vector and $b \neq a$, there must be a position $i \in S_b$ where $b_i \geq 2$. Hence, $\langle b, x\rangle \geq \langle a, a\rangle + 1 > \langle a, a\rangle$.

Case 3: $S_b \subset S_a$ (proper inclusion). We show that this case is impossible. Since F must lie above A, there must be a vector $a' \in A$ such that $a' \leq b$ and, hence, also $S_{a'} \subseteq S_b \subset S_a$. But since both a and a' are 0-1 vectors, this contradicts with A being an antichain. ∎

Appendix B. Sidon sets in terms of graphs. Recall that every set $A \subset \mathbb{N}^n$ of vectors has its associated bipartite graph G_A whose nodes are vectors in $\mathbb{N}^n$, and two nodes x and y are adjacent precisely when $x + y \in A$.

LEMMA 10. *A set $B \subseteq A$ is a Sidon set inside A if and only if no copy of $K_{2,2}$ in G_A contains two disjoint edges of G_B.*

Proof. ($\Rightarrow$): Suppose G_A has a copy of $K_{2,2}$ with two disjoint edges in G_B (see Figure 3). So, there are vectors $a = x + y$ and $b = x' + y'$ in B for some vectors $x \neq x'$ and $y \neq y'$ in $\mathbb{N}^n$ such that the vectors $c = x + y'$ and $d = x' + y$ belong to A. Then clearly $a + b = c + d$ but neither $c = a$ nor $c = b$ can hold, meaning that B is not a Sidon set inside A.

($\Leftarrow$): Assume that B is not a Sidon set inside A. Then there are vectors $a, b \in B$ and $c, d \in A$ such that $a + b = c + d$ but $c \notin \{a, b\}$. To show that then G_A must contain a copy of $K_{2,2}$ with two disjoint edges in G_B, it is enough to show that there exist vectors $x \neq x'$ and $y \neq y'$ in $\mathbb{N}^n$ such that $x + y = a$, $x' + y' = b$ and $\{x, x'\} + \{y, y'\} \subseteq A$. We define the desired vectors componentwise.

If $a_i < c_i$ then take $x_i = 0$, $x'_i = c_i - a_i$, $y_i = a_i$, and $y'_i = d_i$ $(= a_i - c_i + b_i)$. If $a_i \geq c_i$ then take $x_i = a_i - c_i$, $x'_i = 0$, $y_i = c_i$, and $y'_i = b_i$ $(= c_i - a_i + d_i)$. Then all four vectors x, x', y, y' belong to $\mathbb{N}^n$, vectors $x + y = a$ and $x' + y' = b$ belong to B, and the "cross-vectors," vectors $x + y' = d$ and $x' + y = c$, belong to A. Moreover, $c \neq a$ implies $x \neq x'$ and $c \neq b$ implies $y \neq y'$. ∎

Appendix C. Cover-free sets of size $2^{\epsilon n}$. Using ideas of Justesen codes, Friedman [10] has shown how to construct, for every large (but fixed) constant q and any sufficiently large m, an explicit code $C \subseteq \{1, \ldots, q\}^m$ of $|C| = q^{\Omega(m)}$ vectors (code words) with the minimal Hamming distance $d > m/2$ between any two distinct code words. As observed by Alon [1], each triple (x, y, z) of code words must then have a position j such that $x_j \notin \{y_j, z_j\}$. Indeed, otherwise the sum of distances of x from y and from z would not exceed m. But this sum must be at least $2d > m$, a contradiction.

Now replace in each code word each occurrence of the ith symbol by the 0-1 vector of length q with exactly one 1 in the ith position. This gives us an explicit set $A \subseteq \{0,1\}^n$ of $|C| = q^{\Omega(m)}$ 0-1 vectors of length $n = qm$. Since each vector of A has exactly m ones, the set is homogeneous.

To see that A is cover free, take any three distinct vectors a, b, c in A. By the property $d > m/2$ of the original code C, the corresponding triple (x, y, z) of code words must have a position $1 \leq j \leq m$ such that $x_j \notin \{y_j, z_j\}$. Since distinct symbols were replaced by distinct unit vectors, there will be a position i in which $a_i = 1$ and $b_i = c_i = 0$. Hence, the set A is cover free.

Appendix D. Sidon sets of size $2^{n/2}$. Lindström [24] and Cilleruelo [4] have shown that the maximum size of a Sidon set in $\{0,1\}^n$ is asymptotically equal to $2^{n/2}$. On the other hand, an *explicit* Sidon set $A \subset \{0,1\}^n$ of size $|A| = 2^{n/2}$ was earlier constructed by Lindström [23] as follows.

Let $n = 2m$, and $A \subset \{0,1\}^n$ be the set of all vectors (x, x^3) with $x \in \{0,1\}^m$, where we view vectors x as elements of $\mathrm{GF}(2^m)$ when raising them to a power. The set A is clearly not cover free because it is even not an antichain: say, the all-0 vector belongs to it. But it is not difficult to verify that A is a Sidon set.

To show this, fix any two vectors $c, d \in \{0,1\}^m$, and consider the equation $(x, x^3) + (y, y^3) = (c, c^3) + (d, d^3)$. It is enough to show that this equation has at most one unordered pair $\{x, y\}$ of 0-1 solutions over the semigroup $(\mathbb{N}^{2m}, +)$. If $c = d$ then there is only one solution $x = y = c$. So, assume that $c \neq d$. It is enough to show that then the equation cannot have more than one solution $\{x, y\}$ even over the field $\mathrm{GF}(2^{2m})$.

The equation is equivalent to the system of two equations $x+y = a$ and $x^3+y^3 = b$ with $a = c+d \neq 0$ and $b = c^3+d^3$. Since we are working over a field of characteristic 2, the identity $(x+y)^3 = x^3+y^3+3xy(x+y)$ turns into $axy = (x+y)^3+(x^3+y^3) = a^3+b$. Thus, x and y must satisfy $x + y = a \neq 0$ and $xy = a^2 + b/a$. By Vieta's formulas, x and y are then the roots of the polynomial $aX^2 + a^2X + (a^3 + b)$, and there can be only one pair of them.

Appendix E. Sets of tropical complexity $2^{n/2-o(n)}$. Let q be a prime power, $t \geq 2$ an integer, and consider the field $\mathbb{F} = \mathrm{GF}(q^t)$ with q^t elements. The *norm* is a mapping $N : \mathrm{GF}(q^t) \to \mathrm{GF}(q)$ given by $N(a) = a \cdot a^q \cdots a^{q^{t-1}} = a^{(q^t-1)/(q-1)}$. Consider the set $A = \{a \in \mathbb{F} \colon N(a) = 1\}$ of all elements of unit norm. It is known (see, e.g., [22]) that $|A| = (q^t - 1)/(q - 1)$.

Kollár, Rónyai, and Szabó [21, Theorem 3.3] proved that, for every t distinct elements $a_1, \ldots, a_t$ of $\mathbb{F}$, the system of equations

$$N(a_1 + x) = 1, N(a_2 + x) = 1, \ldots, N(a_t + x) = 1$$

has at most $t!$ solutions $x \in \mathbb{F}$. Hence, the set A is $(t, t!)$-sparse over the group $(\mathbb{F}, +)$. Now let $q = 2^r$ and $m = rt$. By viewing elements of $\mathrm{GF}(2^m)$ as vectors in $\{0,1\}^m$, we obtain a set $A \subseteq \{0,1\}^m$ of $|A| = (2^{rt}-1)/(2^r-1) \geq 2^{r(t-1)} = 2^{m-m/t}$ vectors which is $(t, t!)$-sparse over $(\mathbb{F}, +)$ and, hence, also over the semigroup $(\mathbb{N}^n, +)$.

The constructed set $A \subset \{0,1\}^m$ is not homogeneous, but it can be extended to a homogeneous set by using the following simple trick (which emerged during the discussions with Igor Sergeev). Namely, we can extend every set $A \subseteq \{0,1\}^m$ to a homogeneous set $A^* = \{(a, \overline{a})\colon a \in A\} \subseteq \{0,1\}^{2m}$, where $\overline{a}$ denotes the componentwise negation of a 0-1 vector a. For example, $\overline{(1,1,0,1,0)} = (0,0,1,0,1)$. Note that the set A^* is already homogeneous because every vector has exactly m ones.

CLAIM 11. *If $A \subseteq \{0,1\}^m$ is (k,l)-sparse then $A^* \subseteq \{0,1\}^{2m}$ is also (k,l)-sparse.*

Proof. Assume contrariwise that A^* is not (k,l)-sparse. Then $X + Y \subseteq A^*$ holds for some sets $X, Y \subseteq \{0,1\}^{2m}$ with $|X| > k$ and $|Y| > l$. Let X' and Y' be the projections of X and Y onto the first m positions. The sumset $X' + Y'$ must lie in the set A. So, it remains to show that $|X'| \geq |X|$ and $|Y'| \geq |Y|$, because then A is also not (k,l)-sparse, and we have the desired contradiction.

To show $|X'| \geq |X|$, assume contrariwise that some two distinct vectors $(x,y) \neq (x',y')$ in X have the same projection on the first m positions, i.e., that $x = x'$. Take any vector (u,v) in Y. Since the sumset $X+Y$ lies in A^*, there must be vectors $a, b \in A$ such that $(x,y)+(u,v) = (a,\overline{a})$ and $(x,y')+(u,v) = (b,\overline{b})$. But then $x+u = a$ and $x+u = b$ imply $a = b$. From $y+v = \overline{a}$ and $y'+v = \overline{b} = \overline{a}$, we then have $y = y'$. Together with $x = x'$, this contradicts $(x,y) \neq (x',y')$. The proof of $|Y'| \geq |Y|$ is the same. ∎

Now let $B = A^* \subseteq \{0,1\}^n$ with $n = 2m$ be the homogeneous extension of the norm set $A \subseteq \{0,1\}^m$ constructed above. By Claim 11, B is $(t, t!)$-sparse and has $|A| \geq 2^{m-m/t}$ vectors. Assuming that $\sqrt{m}$ is an integer, take $t = \sqrt{m}$. Then Theorem 3 and the reduction lemma imply that both $\mathrm{Min}(B)$ and $\mathrm{Max}(B)$ are at least

$$\frac{|A|}{(t!)^2} \geq \frac{|A|}{t^{2t}} \geq 2^{m-m/t-2t\log t} = 2^{m-\sqrt{m}\log(2m)} \geq 2^{n/2-\sqrt{n}\log n}\,.$$

To our best knowledge, no larger explicit lower bound on tropical complexity is known.

Acknowledgments. I thank Georg Schnitger and Igor Sergeev for many inspiring discussions. I am also thankful to referees for useful suggestions and corrections.

REFERENCES

[1] N. Alon, *Explicit construction of exponential sized families of k-independent sets*, Discrete Math., 58 (1986), pp. 191–193.
[2] R. Bellman, *On a routing problem*, Quart. Appl. Math., 16 (1958), pp. 87–90.
[3] X. Chen, N. Kayal, and A. Wigderson, *Partial derivatives in arithmetic complexity and beyond*, Found. Trends Theor. Comput. Sci., 6 (2011), pp. 1–138.
[4] J. Cilleruelo, *Sidon sets in N^d*, J. Combin. Theory Ser. A, 117 (2010), pp. 857–871.
[5] P. Erdős, P. Frankl, and Z. Füredi, *Families of finite sets in which no set is covered by the union of two others*, J. Combin. Theory Ser. A, 33 (1982), pp. 158–166.
[6] P. Erdős and E. Harzheim, *Congruent subsets of infinite sets of natural numbers*, J. Reine Angew. Math., 367 (1986), pp. 207–214.
[7] P. Erdős and P. Turán, *On a problem of Sidon in additive number theory, and on some related problems*, J. Lond. Math. Soc. (2), 16 (1941), pp. 212–215.
[8] R. Floyd, *Algorithm 97, shortest path*, Comm. ACM, 5 (1962), p. 345.
[9] L. Ford, *Network Flow Theory*, Technical report P-923, The Rand Corp., 1956.
[10] J. Friedman, *Constructing $O(n \log n)$ size monotone formulae for the kth elementary symmetric polynomial of n boolean variables*, in 25th Annual Symposium on Foundations of Computer Science, IEEE Computer Society, Los Angeles, 1984, pp. 506–515.

[11] S. Gashkov, *On one method of obtaining lower bounds on the monotone complexity of polynomials*, Vestnik Moskov. Univ., Ser. 1 Math., Mekh, 5 (1987), pp. 7–13.
[12] S. Gashkov and I. Sergeev, *A method for deriving lower bounds for the complexity of monotone arithmetic circuits computing real polynomials*, Sb. Math., 203 (2012), pp. 1411–1147.
[13] D. Grigoriev and G. Koshevoy, *Complexity of tropical Schur polynomials*, J. Symbolic Comput., 74 (2016), pp. 46–54.
[14] M. Held and R. M. Karp, *A dynamic programming approach to sequencing problems*, SIAM J. on Appl. Math., 10 (1962), pp. 196–210.
[15] P. Hrubes and A. Yehudayoff, *Homogeneous formulas and symmetric polynomials*, Comput. Complexity, 20 (2011), pp. 559–578.
[16] L. Hyafil, *On the parallel evaluation of multivariate polynomials*, SIAM J. Comput., 8 (1979), pp. 120–123.
[17] M. Jerrum and M. Snir, *Some exact complexity results for straight-line computations over semirings*, J. ACM, 29 (1982), pp. 874–897.
[18] S. Jukna, *Boolean Function Complexity: Advances and Frontiers*, Springer, Berlin, 2012.
[19] S. Jukna, *Lower bounds for tropical circuits and dynamic programs*, Theory Comput. Syst., 57 (2015), pp. 160–194.
[20] W. Kautz and R. Singleton, *Nonrandom binary superimposed codes*, IEEE Trans. Inform. Theory, 10 (1964), pp. 363–377.
[21] J. Kollár, L. Rónyai, and T. Szabó, *Norm-graphs and bipartite Turán numbers*, Combinatorica, 16 (1996), pp. 399–406.
[22] R. Lidl and H. Niederreiter, *Introduction to Finite Fields and their Applications.*, Cambridge University Press, Cambridge, 1986.
[23] B. Lindström, *Determination of two vectors from the sum*, J. Combin. Theory Ser. B, 6 (1969), pp. 402–407.
[24] B. Lindström, *On B_2-sequences of vectors*, J. Number Theory, 4 (1972), pp. 261–265.
[25] E. Moore, *The shortest path through a maze*, in Proceedings of an Internatational Symposium on Switching Theory, Vol. II, Harvard Univ. Press, Cambridge, 1959, pp. 285–292.
[26] K. O'Bryant, *A complete annotated bibliography of work related to Sidon sequences*, Electr. J. Comb., 11 (2004), DS11.
[27] N. Pippenger, *On another Boolean matrix*, Theoret. Comput. Sci., 11 (1980), pp. 49–56.
[28] R. Raz and A. Yehudayoff, *Multilinear formulas, maximal-partition discrepancy and mixed-sources extractors*, J. Comput. System Sci., 77 (2011), pp. 167–190.
[29] C. Schnorr, *A lower bound on the number of additions in monotone computations*, Theoret. Comput. Sci., 2 (1976), pp. 305–315.
[30] R. Sengupta and H. Venkateswaran, *A lower bound for monotone arithmetic circuits computing* 0–1 *permanent*, Theoret. Comput. Sci., 209 (1998), pp. 389–398.
[31] E. Shamir and M. Snir, *On the depth complexity of formulas*, Math. System Theory, 13 (1980), pp. 301–322.
[32] A. Shpilka and A. Yehudayoff, *Arithmetic circuits: A survey of recent results and open questions*, Found. Trends Theor. Comput. Sci., 5 (2010), pp. 207–388.
[33] P. Tiwari and M. Tompa, *A direct version of Shamir and Snir's lower bounds on monotone circuit depth*, Inform. Process. Lett., 49 (1994), pp. 243–248.
[34] L. Valiant, *Negation can be exponentially powerful*, Theoret. Comput. Sci., 12 (1980), pp. 303–314.
[35] S. Warshall, *A theorem on boolean matrices*, J. ACM, 9 (1962), pp. 11–12.