

Threesomes, Degenerates, and Love Triangles

Allan Grønlund
MADALGO, Aarhus University
Aarhus, Denmark

Seth Pettie
University of Michigan
Ann Arbor, MI

Abstract—The 3SUM problem is to decide, given a set of n real numbers, whether any three sum to zero. It is widely conjectured that a trivial $O(n^2)$ -time algorithm is optimal and over the years the consequences of this conjecture have been revealed. This 3SUM conjecture implies $\Omega(n^2)$ lower bounds on numerous problems in computational geometry and a variant of the conjecture implies strong lower bounds on triangle enumeration, dynamic graph algorithms, and string matching data structures.

In this paper we refute the 3SUM conjecture. We prove that the decision tree complexity of 3SUM is $O(n^{3/2}\sqrt{\log n})$ and give two subquadratic 3SUM algorithms, a deterministic one running in $O(n^2/(\log n/\log \log n)^{2/3})$ time and a randomized one running in $O(n^2(\log \log n)^2/\log n)$ time with high probability. Our results lead directly to improved bounds for k -variate linear degeneracy testing for all odd $k \geq 3$. The problem is to decide, given a linear function $f(x_1, \dots, x_k) = \alpha_0 + \sum_{1 \leq i \leq k} \alpha_i x_i$ and a set $A \subset \mathbb{R}$, whether $0 \in f(A^k)$. We show the decision tree complexity of this problem is $O(n^{k/2}\sqrt{\log n})$.

Finally, we give a subcubic algorithm for a generalization of the $(\min, +)$ -product over real-valued matrices and apply it to the problem of finding zero-weight triangles in weighted graphs. We give a depth- $O(n^{3/2}\sqrt{\log n})$ decision tree for this problem, as well as an algorithm running in time $O(n^3(\log \log n)^2/\log n)$.

I. INTRODUCTION

The time hierarchy theorem [24] implies that there exist problems in \mathbf{P} with complexity $\Omega(n^k)$ for every fixed k . However, it is consistent with current knowledge that all problems of practical interest can be solved in $\tilde{O}(n)$ time in a reasonable model of computation. Efforts to build a useful complexity theory inside \mathbf{P} have been based on the conjectured hardness of certain archetypal problems, such as 3SUM, $(\min, +)$ -matrix product, and CNF-SAT. See, for example, the conditional lower bounds in [22], [30], [31], [25], [1], [2], [32], [15], [36].

In this paper we study the complexity of 3SUM and related problems such as linear degeneracy testing (LDT) and finding zero-weight triangles. Let us define the problems formally.

3SUM: Given a set $A \subset \mathbb{R}$, determine if there exists $a, b, c \in A$ such that $a + b + c = 0$.

INTEGER3SUM: Given a set $A \subseteq \{-U, \dots, U\} \subset \mathbb{Z}$, determine if there exists $a, b, c \in A$ such that $a + b + c = 0$.

k -LDT and k -SUM: Fix a k -variate linear function $\phi(x_1, \dots, x_k) = \alpha_0 + \sum_{i=1}^k \alpha_i x_i$, where $\alpha_0, \dots, \alpha_k \in \mathbb{R}$. Given a set $A \subset \mathbb{R}$, determine if $\phi(\mathbf{x}) = 0$ for any $\mathbf{x} \in A^k$. When ϕ is $\sum_{i=1}^k x_i$ the problem is called k -SUM.

ZEROTRIANGLE: Given a weighted undirected graph $G = (V, E, w)$, where $w : E \rightarrow \mathbb{R}$, determine if there

exists a triangle $(a, b, c) \in V^3$ for which $w(a, b) + w(b, c) + w(c, a) = 0$. (From the definition of *love* : a score of zero, one could also call this the LOVETRIANGLE problem.)

These problems are often defined with further constraints that do not change the problem in any substantive way [22]. For example, the input to 3SUM can be three sets $A, B, C \subset \mathbb{R}$ and the problem is to determine if there exists $a \in A, b \in B, c \in C$ such that $a + b + c = 0$. Even if there is only one set, there is sometimes an additional constraint that a, b , and c be *distinct* elements.

As a problem in its own right, 3SUM has no compelling practical applications. However, lower bounds on 3SUM imply lower bounds on dozens of other problems that *are* of practical interest. Before reviewing existing 3SUM algorithms we give a brief survey of conditional lower bounds that depend on the hardness of 3SUM.

A. Implications of the 3SUM Conjectures

It is often conjectured that 3SUM requires $\Omega(n^2)$ time and that INTEGER3SUM requires $\Omega(n^{2-o(1)})$ time [30], [2]. These conjectures have been shown to imply strong lower bounds on numerous problems in computational geometry [22], [4], [8], [33] dynamic graph algorithms [30], [2], and pattern matching [3], [6], [13], [16]. For example, the 3SUM conjecture implies that the following problems require at least $\Omega(n^2)$ time.

- Given an n -point set in \mathbb{R}^2 , determine whether it contains three collinear points [22].
- Given two n -edge convex polygons, determine whether one can be placed inside the other via rotation and translation [8].
- Given n triangles in \mathbb{R}^2 , determine whether their union contains a hole, or determine the area of their union [22].

Through a series of reductions, Pătraşcu [30] proved that the INTEGER3SUM conjecture implies lower bounds on triangle enumeration and various problems in dynamic data structures, even when all updates and queries are presented in advance. Some lower bounds implied by the INTEGER3SUM conjecture include the following.

- Given an undirected m -edge graph, enumerating up to m triangles (3-cycles) requires at least $\Omega(m^{4/3-o(1)})$ time [30].¹

¹Bjorklund et al. [9] recently proved that the exponent $4/3$ is optimal if the matrix multiplication exponent ω is 2.

- Given a sequence of m updates to a directed graph (edge insertions and edge deletions) and two specified vertices s, t , determining whether t is reachable from s after each update requires at least $\Omega(m^{4/3-o(1)})$ time [2].
- Given an edge-weighted undirected graph, deciding whether there exists a zero-weight triangle requires at least $\Omega(n^{3-o(1)})$ time [37].

In recent years conditional lower bounds have been obtained from other plausible conjectures: that computing the $(\min, +)$ -product of two $n \times n$ matrices requires $\Omega(n^{3-o(1)})$ time, that CNF-SAT requires $2^{\Omega(n)}$ time (Exponential Time Hypothesis, or ETH), and that CNF-SAT requires $2^{(1-o(1))n}$ time (Strong ETH). We now know that if the ETH holds, k -SUM requires $n^{\Omega(k)}$ time [31], and if the Strong ETH holds, $(3/2-\epsilon)$ -approximating the diameter of an m -edge unweighted graph requires $\Omega(m^{2-o(1)})$ time [32], [15]. Williams and Williams [36] proved that numerous problems are equivalent to $(\min, +)$ -matrix multiplication, inasmuch as a truly subcubic ($O(n^{3-\Omega(1)})$) algorithm for one would imply truly subcubic algorithms for all the others.

B. Algorithms, Lower Bounds, and Reductions

The evidence in favor of the 3SUM and INTEGER3SUM conjectures is rather thin. Erickson [19] and Ailon and Chazelle [5] proved that any k -linear decision tree for solving k -LDT must have depth $\Omega(n^{k/2})$ when k is even and $\Omega(n^{(k+1)/2})$ when k is odd. In particular, any 3-linear decision tree for 3SUM has depth $\Omega(n^2)$. (An s -linear decision tree is one where each internal node asks for the sign of a linear expression of s elements.) The INTEGER3SUM problem is obviously not harder than 3SUM, but no other relationship between these two problems is known. Indeed, the assumption that elements are integers opens the door to a variety of algorithmic techniques that cannot be modeled as decision trees. Using the fast Fourier transform it is possible to solve INTEGER3SUM in $O(U \log U)$ time, which is subquadratic even for a rather large universe size U .² Baran, Demaine, and Pătraşcu [7] showed that INTEGER3SUM can be solved in $O(n^2/(\log n/\log \log n)^2)$ time (with high probability) on the word RAM, where $U = 2^w$ and $w > \log n$ is the machine word size. The algorithm uses a mixture of randomized universe reduction (via hashing), word packing, and table lookups.

It is straightforward to reduce k -LDT to a 2SUM problem or unbalanced 3SUM problem, depending on whether k is even or odd. When k is odd one forms certain sets A, B, C where $|A| = |B| = n^{(k-1)/2}$ and $|C| = n$, then sorts them in $O(n^{(k-1)/2} \log n)$ time. The standard three-set 3SUM algorithm on A, B, C takes $O(|C|(|A| + |B|)) = O(n^{(k+1)/2})$ time. When $k \geq 4$ is even there is no set C . Using Lambert’s algorithm [26] (cf. [34]), A and B can be sorted in $O(n^{k/2} \log n)$ time while performing only $O(n^{k/2})$ comparisons. These algorithms can be modeled as k -linear decision trees, and are therefore optimal in this model by the

²Erickson [19] credits R. Seidel with this 3SUM algorithm.

lower bounds of [19], [5]. However, it was known that all k -LDT problems can be solved by n -linear decision trees with depth $O(n^5 \log n)$ [27], or with depth $O(n^4 \log(nK))$ if the coefficients of the linear function are integers with absolute value at most K [28]. Unfortunately these decision trees are not efficiently constructible. The time required to determine which comparisons to make is exponential in n .

The ZEROTRIANGLE problem was highlighted in a recent article by Williams and Williams [37]. They did not give any subcubic algorithm, but did show that a subcubic ZEROTRIANGLE algorithm would have implications for INTEGER3SUM via an intermediate problem called CONVOLUTION3SUM.

- CONVOLUTION3SUM: Given a vector $A \in \mathbb{R}^n$, determine if there exist i, j for which $A(i) + A(j) = A(i + j)$.
- INTEGERCONV3SUM: The same as CONVOLUTION3SUM, except that $A \in \{0, \dots, U - 1\}^n$ and $U \leq 2^w$, where $w = \Omega(\log n)$ is the machine word size.

Pătraşcu [30] defined the CONVOLUTION3SUM problem and gave a randomized reduction from INTEGER3SUM to INTEGERCONV3SUM. Williams and Williams [37] gave a reduction from CONVOLUTION3SUM to ZEROTRIANGLE. Neither of these reductions is frictionless. Define $T_{I3S}, T_{IC3S}, T_{C3S}$ and T_{ZT} to be the complexities of the various problems on inputs of length n , or graphs with n vertices. Clearly $T_{IC3S}(n) \leq T_{C3S}(n)$. The reductions show that for any k , $T_{I3S}(n) = O(n^2/k + k^3 \cdot T_{IC3S}(n/k))$ and $T_{C3S}(n) = O(\sqrt{n} \cdot T_{ZT}(\sqrt{n}))$. Note that even if ZEROTRIANGLE had an $O(n^2)$ -time algorithm (optimal on dense graphs), this would only give an $O(n^{9/5})$ bound for INTEGER3SUM.

C. New Results

We give the first subquadratic bounds on both the decision tree complexity of 3SUM and the algorithmic complexity of 3SUM, which also gives the first deterministic subquadratic algorithm for INTEGER3SUM.³ Our method leads to similar improvements to the decision tree complexity of k -LDT when $k \geq 3$ is odd, but *not* to the algorithmic complexity of k -LDT. Refer to Table I for a summary of prior work and our results.

Theorem 1. *There is a 4-linear decision tree for 3SUM with depth $O(n^{3/2} \sqrt{\log n})$. Furthermore, 3SUM can be solved deterministically in $O(n^2/(\log n/\log \log n)^{2/3})$ time and, using randomization, in $O(n^2(\log \log n)^2/\log n)$ time w.h.p.*

Theorem 2. *When $k \geq 3$ is odd, there is a $(2k - 2)$ -linear decision tree for k -LDT with depth $O(n^{k/2} \sqrt{\log n})$.*

Theorem 1 refutes the 3SUM conjecture and casts serious doubts on the optimality of many $O(n^2)$ algorithms in computational geometry. Theorem 1 also answers a question of Erickson [19] and Ailon and Chazelle [5] about whether

³We assume a simplified Real RAM model. Real numbers are subject to only two unit-time operations: addition and comparison. In all other respects the machine behaves like a $w = O(\log n)$ -bit word RAM with the standard repertoire of unit-time AC^0 operations: bitwise Boolean operations, left and right shifts, addition, and comparison.

3SUM		
trivial	n^2	
new	$n^{3/2}\sqrt{\log n}$	DEC. TREE
	$n^2 / \left(\frac{\log n}{\log \log n}\right)^{2/3}$	
	$n^2 / \frac{\log n}{(\log \log n)^2}$	RAND.
INTEGER3SUM		
trivial	n^2	
Seidel (1997)	$U \log U$	
Baran, Demaine, and Pătraşcu (2005)	$n^2 / \left(\frac{\log n}{\log \log n}\right)^2$	RAND.
	$n^2 / \frac{w}{\log^2 w}$	RAND.
new	$n^2 / \left(\frac{\log n}{\log \log n}\right)^{2/3}$	
CONVOLUTION3SUM		
trivial	n^2	
new	$n^{3/2}\sqrt{\log n}$	DEC. TREE
	$n^{3/2}$	RAND., DEC. TREE
	$n^2 / \frac{\log n}{(\log \log n)^2}$	
	$n^2 / \frac{\log n}{\log \log n}$	RAND.
ZEROTRIANGLE		
trivial	n^3	
new	$n^{5/2}\sqrt{\log n}$	DEC. TREE
	$n^{5/2}$	RAND., DEC. TREE
	$n^3 / \frac{\log n}{(\log \log n)^2}$	
	$n^3 / \frac{\log n}{\log \log n}$	RAND.
	$m^{5/4}\sqrt{\log m}$	DEC. TREE
	$m^{5/4}$	RAND., DEC. TREE
	$m^{3/2} / \left(\frac{\log m}{(\log \log m)^2}\right)^{1/4}$	
	$m^{3/2} / \left(\frac{\log m}{\log \log m}\right)^{1/4}$	RAND.

TABLE I

RESULTS IN THE DECISION TREE MODEL ARE INDICATED BY DEC. TREE, RESULTS USING RANDOMIZATION ARE INDICATED BY RAND. IN THE ZEROTRIANGLE PROBLEM n AND m ARE THE NUMBER OF VERTICES AND EDGES WHEREAS IN ALL OTHER PROBLEMS n IS THE LENGTH OF THE INPUT. IN THE INTEGER3SUM PROBLEM $w = \Omega(\log n)$ IS THE MACHINE WORD SIZE AND $U \leq 2^w$ THE SIZE OF THE UNIVERSE.

$(k + 1)$ -linear decision trees are more powerful than k -linear decision trees in solving k -LDT problems. In the case of $k = 3$, they are.

We define a new product of three real-valued matrices called *target-min-plus*, which is trivially computable in $O(n^3)$ time. We observe that ZEROTRIANGLE is reducible to a target-min-plus product, then give subcubic bounds on the decision tree and algorithmic complexity of target-min-plus. Theorem 3 is an immediate consequence.

Theorem 3. *The decision tree complexity of ZEROTRIANGLE is $O(n^{5/2}\sqrt{\log n})$ on n -vertex graphs and its randomized decision tree complexity is $O(n^{5/2})$ w.h.p. There is a deterministic ZEROTRIANGLE algorithm running in $O(n^3(\log \log n)^2/\log n)$ time and a randomized algorithm running in $O(n^3 \log \log n/\log n)$ time w.h.p.*

Any m -edge graph contains $O(m^{3/2})$ triangles which can be enumerated in $O(m^{3/2})$ time, so ZEROTRIANGLE can clearly be solved in $O(m^{3/2})$ time as well. We improve this bound for all m .

Theorem 4. *The decision tree complexity of ZEROTRIANGLE on m -edge graphs is $O(m^{5/4}\sqrt{\log m})$ and, using randomization, $O(m^{5/4})$ w.h.p. The ZEROTRIANGLE problem can be solved in $O(m^{3/2}/(\log m/(\log \log m)^2)^{1/4})$ time deterministically or $O(m^{3/2}/(\log m/\log \log m)^{1/4})$ w.h.p.*

By invoking the Williams-Williams reduction [37], our ZEROTRIANGLE algorithms give subquadratic bounds on the complexity of CONVOLUTION3SUM. By designing CONVOLUTION3SUM algorithms from scratch we can obtain speedups comparable to those of Theorem 3.

Theorem 5. *The decision tree complexity of CONVOLUTION3SUM is $O(n^{3/2}\sqrt{\log n})$ and its randomized decision tree complexity is $O(n^{3/2})$ w.h.p. The CONVOLUTION3SUM problem can be solved in $O(n^2(\log \log n)^2/\log n)$ time deterministically, or in $O(n^2 \log \log n/\log n)$ time w.h.p.*

Due to the antiquated page limits imposed on us, complete proofs of Theorems 2, 3, 4, and 5 do not appear in this extended abstract. Refer to [23].

D. An Overview

All of our algorithms borrow liberally from Fredman's 1976 articles on the decision tree complexity of $(\min, +)$ -matrix multiplication [21] and the complexity of sorting $X + Y$ [20]. Throughout the paper we shall refer to the ingenious observation that $a + b < c + d$ iff $a - c < d - b$ as *Fredman's trick*. In order to shave off $\text{poly}(\log n)$ factors in runtime we apply the bichromatic domination technique invented by Chan [14] and developed further by Bremner, Chan, Demaine, Erickson, Hurtado, Iacono, Langerman, Pătraşcu, and Taslakian [11].

In Section II we review a number of useful lemmas due to Fredman [20], Buck [12], and Chan [14] about sorting with partial information, the complexity of hyperplane arrangements, and the complexity of dominance reporting in \mathbb{R}^d . In Section III we review a standard $O(n^2)$ -time 3SUM algorithm

and in Section IV we present an $\tilde{O}(n^{3/2})$ -depth decision tree for 3SUM. Subquadratic algorithms for 3SUM are presented in Section V. Section VI presents new bounds on the decision tree and algorithmic complexity of ZEROTRIANGLE. Section VII concludes with some open problems.

II. USEFUL LEMMAS

Fredman [20] considered the problem of sorting a list of n numbers known to be arranged in one of $\Pi \leq n!$ permutations. When Π is sufficiently small the list can be sorted using a linear number of comparisons.

Lemma 6. (Fredman 1976 [20]) *A list of n numbers whose sorted order is one of Π permutations can be sorted with $2n + \log \Pi$ pairwise comparisons.*

Throughout the paper $[N]$ denotes the first $[N]$ natural numbers $\{0, \dots, [N] - 1\}$, where N may or may not be an integer. We apply Lemma 6 to the problem of sorting Cartesian sums. Given lists $A = (a_i)_{i \in [n]}$ and $B = (b_i)_{i \in [n]}$ of distinct numbers, define $A + B = \{a_i + b_j \mid i, j \in [n]\}$. We often regard $A + B$ as an $|A| \times |B|$ matrix (which may contain multiple copies of the same number) or as a point in the $2n$ -dimensional space \mathbb{R}^{2n} , whose coordinates are named $x_1, \dots, x_n, y_1, \dots, y_n$. The points in \mathbb{R}^{2n} that agree with a fixed permutation of $A + B$ form a convex cone bounded by the $\binom{n}{2}$ hyperplanes $H = \{x_i + y_j - x_k - y_l \mid i, j, k, l \in [n]\}$. The sorted order of $A + B$ is encoded as a sign vector $\{-1, 0, 1\}^{\binom{n}{2}}$ depending on whether (A, B) lies on, above, or below a particular hyperplane in H . Therefore the number of possible sorted orders of $A + B$ is exactly the number of regions (of all dimensions) defined by the arrangement H . (Regions of dimension less than $2n$ correspond to instances in which some numbers appear multiple times.)

Lemma 7. (Buck 1943 [12]) *Consider the partition of space defined by an arrangement of m hyperplanes in \mathbb{R}^d . The number of regions of dimension $k \leq d$ is at most*

$$\binom{m}{d-k} \left(\binom{m-d+k}{0} + \binom{m-d+k}{1} + \dots + \binom{m-d+k}{k} \right)$$

and the number of regions of all dimensions is $O(m^d)$.

In one of our algorithms we will construct the hyperplane arrangement explicitly. Edelsbrunner, O'Rourke, and Seidel [18] proved that the natural incremental algorithm takes $O(m^d)$ time (linear in the size of the arrangement), but any trivial $m^{O(d)}$ -time algorithm suffices in our application. The hyperplane arrangements we use correspond to fragments of the Cartesian sum $A + B$. Lemma 8 is a direct consequence of Lemmas 6 and 7.

Lemma 8. *Let $A = (a_i)_{i \in [n]}$ and $B = (b_i)_{i \in [n]}$ be two lists of numbers and let $F \subseteq [n]^2$ be a set of positions in the $n \times n$ grid. The number of realizable orders of $(A + B)|_F \stackrel{\text{def}}{=} \{a_i + b_j \mid (i, j) \in F\}$ is $O\left(\binom{|F|}{2}^{2n}\right)$ and therefore $(A + B)|_F$ can be sorted with at most $2|F| + 2n \log |F| + O(1)$ comparisons.*

It is sometimes convenient to assume that the elements of a Cartesian sum are distinct, even though numbers may appear multiple times. Lemma 9 illustrates one way to break ties consistently. The proof is straightforward.

Lemma 9. *Let $A = (a_i)$ and $B = (b_i)$ be two lists of numbers. Define $a'_i = (a_i, i, 0)$ and $b'_j = (b_j, 0, j)$. The Cartesian sum $A' + B'$ is totally ordered, and is a linear extension of the partially ordered $A + B$. (Addition over tuples is pointwise addition; tuples are ordered lexicographically. The tuple (u, v, w) can be regarded as a representation of a real number $u + \epsilon_1 v + \epsilon_2 w$ where $\epsilon_1 \gg \epsilon_2$ are sufficiently small so as not to invert strictly ordered elements of $A + B$.)*

Given a set P of red and blue points in \mathbb{R}^d , the bichromatic dominating pairs problem is to enumerate every pair $(p, q) \in P^2$ such that p is red, q is blue, and p is greater than q at each of the d coordinates. A natural divide and conquer algorithm [29, p. 366] runs in time linear in the output size and $O(n \log^d n)$. Chan [14] provided an improved analysis when d is logarithmic in n . For the sake of completeness we give a short proof of Lemma 10 in Appendix A.

Lemma 10. (Chan [14]) *Given a set $P \subseteq \mathbb{R}^d$ of red and blue points, it is possible to return all bichromatic dominating pairs $(p, q) \in P^2$ in time linear in the output size and $c_\epsilon^d |P|^{1+\epsilon}$. Here $\epsilon \in (0, 1)$ is arbitrary and $c_\epsilon = 2^\epsilon / (2^\epsilon - 1)$.*

We typically invoke Lemma 10 with $\epsilon = 1/2, c_\epsilon \approx 3.42$, and $d = \delta \log n$, where $\delta > 0$ is sufficiently small to make the running time subquadratic, excluding the time allotted to reporting the output.

III. THE QUADRATIC 3SUM ALGORITHM

We shall review a standard $O(n^2)$ algorithm for the three-set version of 3SUM and introduce some terminology used in Sections IV and V. We are given sets $A, B, C \subset \mathbb{R}$ and must determine if there exists $a \in A, b \in B, c \in C$ such that $a + b + c = 0$. For each $c \in C$ the algorithm searches for $-c$ in the Cartesian sum $A + B$. Each search takes $O(|A| + |B|)$ time, for a total of $O(|C|(|A| + |B|))$. We view $A + B$ as being a matrix whose rows correspond to A and columns correspond to B , both listed in increasing order.

1. Sort A and B in increasing order as $A(0) \dots A(|A| - 1)$ and $B(0) \dots B(|B| - 1)$.
2. For each $c \in C$,
 - 2.1. Initialize $lo \leftarrow 0$ and $hi \leftarrow |B| - 1$.
 - 2.2. Repeat:
 - 2.2.1. If $-c = A(lo) + B(hi)$, report witness “ $(A(lo), B(hi), c)$ ” and continue.
 - 2.2.2. If $-c < A(lo) + B(hi)$ then decrement hi , otherwise increment lo .
 - 2.3. Until $lo = |A|$ or $hi = -1$.
3. If no witnesses were found report “no witnesses.”

Note that when a witness is discovered in Step 2.2.1 the algorithm continues to search for more witnesses involving c . Since the elements in each row and each column of the $A + B$ matrix are distinct, it does not matter whether we increment

lo or decrement hi after finding a witness. We choose to increment lo in such situations; this choice is reflected in Lemma 11 and its applications in Sections V-C and V-E.

Define the *contour* of x , $\text{CONTOUR}(x, A + B)$, to be the sequence of positions (lo, hi) encountered while searching for x in $A + B$. When $A + B$ is understood from context we will write it as $\text{CONTOUR}(x)$. If $A + B$ is viewed as a topographic map, with the lowest point in the NW corner and highest point in the SE corner, $\text{CONTOUR}(x)$ represents the path taken by an agent attempting to stay as close to altitude x as possible, starting in the NE corner (at position $(0, |B| - 1)$) and ending when it falls off the western or southern side of the map. Lemma 11 is straightforward.

Lemma 11. *Every occurrence of x in $A + B$ lies on $\text{CONTOUR}(x)$. Let $y = (A + B)(i, j)$ be any element of $A + B$. Then $y > x$ iff either (i, j) lies strictly below $\text{CONTOUR}(x)$ or both (i, j) and $(i, j - 1)$ lie on $\text{CONTOUR}(x)$. Similarly, $y \leq x$ iff either (i, j) lies strictly above $\text{CONTOUR}(x)$ or both (i, j) and $(i + 1, j)$ lie on $\text{CONTOUR}(x)$.*

IV. A SUBQUADRATIC 3SUM DECISION TREE

Recall that we are given a set $A \subset \mathbb{R}$ of reals and must determine if there exist $a, b, c \in A$ summing to zero. We first state the algorithm, then establish its correctness and efficiency.

1. Sort A in increasing order as $A(0) \dots A(n-1)$. Partition A into $\lceil n/g \rceil$ groups $A_0, \dots, A_{\lceil n/g \rceil - 1}$ of size at most g , where $A_i \stackrel{\text{def}}{=} \{A(ig), \dots, A((i+1)g - 1)\}$ and $A_{\lceil n/g \rceil - 1}$ may be smaller. The first and last elements of A_i are $\min(A_i) = A(ig)$ and $\max(A_i) = A((i+1)g - 1)$.
2. Sort $D \stackrel{\text{def}}{=} \bigcup_{i \in [n/g]} (A_i - A_i) = \{a - a' \mid a, a' \in A_i \text{ for some } i\}$.
3. For all $i, j \in [n/g]$, sort $A_{i,j} \stackrel{\text{def}}{=} A_i + A_j = \{a + b \mid a \in A_i \text{ and } b \in A_j\}$ using the sorted order on D .
4. For k from 1 to n ,
 - 4.1. Initialize $\text{lo} \leftarrow 0$ and $\text{hi} \leftarrow \lfloor k/g \rfloor$ to be the group index of $A(k)$.
 - 4.2. Repeat:
 - 4.2.1. If $-A(k) \in A_{\text{lo}, \text{hi}}$, report “witness found” and halt.
 - 4.2.2. If $\max(A_{\text{lo}}) + \min(A_{\text{hi}}) > -A(k)$ then decrement hi , otherwise increment lo .
 - 4.3. Until $\text{hi} < \text{lo}$.
5. Report “no witnesses” and halt.

With appropriate modifications this algorithm also solves the three-set version of 3SUM, where the input is $A, B, C \subset \mathbb{R}$.

A. Efficiency of the Algorithm

Step 1 requires $n \log n$ comparisons. One can clearly execute Step 2 using $O(|D| \log |D|)$ comparisons. By Lemmas 6 and 7, Step 2 actually only requires $O(n \log n + |D|) = O(n \log n + gn)$ comparisons to sort D . Using Fredman’s trick, Step 3 requires no comparisons at all, given the sorted order

on D . (If $a, a' \in A_i$ and $b, b' \in A_j$, $a + b < a' + b'$ holds iff $a - a' < b' - b$.) For each iteration of the outer loop (Step 4) there are at most $\lceil n/g \rceil$ iterations of the inner loop (Step 4.2) since each iteration ends by either incrementing lo or decrementing hi. In Step 4.2.1 we can determine whether $-A(k)$ is in $A_{\text{lo}, \text{hi}}$ with a binary search, in $\log |A_{\text{lo}, \text{hi}}| = \log(g^2)$ comparisons. In total the number of comparisons is on the order of $n \log n + gn + (n^2 \log g)/g$, which is $O(n^{3/2} \sqrt{\log n})$ when $g = \sqrt{n \log n}$.

B. Correctness of the Algorithm

The purpose of the outer loop (Step 4) is to find $a, b \in A$, for which $a, b \leq A(k)$ and $a + b + A(k) = 0$. This is tantamount to finding indices lo, hi for which $a \in A_{\text{lo}}, b \in A_{\text{hi}}$, and $-A(k) \in A_{\text{lo}, \text{hi}}$. We maintain the loop invariant that if there exist a, b for which $a + b + A(k) = 0$, then both a and b lie in $A_{\text{lo}}, A_{\text{lo}+1}, \dots, A_{\text{hi}}$. Suppose the algorithm has not halted in Step 4.2.1, that is, there are no solutions with $a \in A_{\text{lo}}, b \in A_{\text{hi}}$. If $\max(A_{\text{lo}}) + \min(A_{\text{hi}}) > -A(k)$ then there can clearly be no solutions with $b \in A_{\text{hi}}$ since $b \geq \min(A_{\text{hi}})$, so decrementing hi preserves the invariant. Similarly, if $\max(A_{\text{lo}}) + \min(A_{\text{hi}}) < -A(k)$ then there can be no solutions with $a \in A_{\text{lo}}$ since $a \leq \max(A_{\text{lo}})$, so incrementing lo preserves the invariant. If it is ever the case that $\text{hi} < \text{lo}$ then, by the invariant, no solutions exist.

C. Algorithmic Implementation

This 3SUM algorithm can be implemented to run in $O(n^2 \log n)$ time while performing only $O(n^{3/2} \log n)$ comparisons. Using any optimal sorting algorithm Steps 1–3 can be executed in $O(gn \log(gn) + (n/g)^2 \cdot g^2 \log g) = O(n^2 \log n)$ time while using $O(gn \log(gn))$ comparisons. Now the boxes $\{A_{i,j}\}$ have been explicitly sorted, so the binary searches in Step 4.2.1 can be executed in $O(\log g)$ time per search. The total running time is $O(n^2 \log n)$ and the number of comparisons is now minimized when $g = \sqrt{n}$, for a total of $O(n^{3/2} \log n)$ comparisons. We do not know of any polynomial time 3SUM algorithm that performs $O(n^{3/2} \sqrt{\log n})$ comparisons.

V. SOME SUBQUADRATIC 3SUM ALGORITHMS

In our 3SUM decision tree, sorting D (Step 2) is a comparison-efficient way to accomplish Step 3, but it only lets us *deduce* the sorted order of the boxes $\{A_{i,j}\}$. It does not give us a useful representation of these sorted orders, namely one that lets us implement each comparison of the binary search in Step 4.2.1 in $O(1)$ time. In this section we present several methods for sorting the boxes based on bichromatic dominating pairs, as in Chan [14] and Bremner et al. [11]; see Lemma 10. The total time spent performing binary searches in Step 4.2.1 will be $O(n^2 \log g/g)$, so our goal is to make g as large as possible, provided that the cost of sorting the boxes is of a lesser order.

As a warmup we give, in Section V-A, a relatively simple subquadratic 3SUM algorithm running in $O(n^2 (\log \log n)^{3/2} / (\log n)^{1/2})$ time. In Section V-B we

present a more sophisticated algorithm, some of whose parameters can be selected either deterministically or randomly. Sections V-C and V-E give two parameterizations of the algorithm, which lead to an $O(n^2(\log \log n/\log n)^{2/3})$ time deterministic 3SUM algorithm and an $O(n^2(\log \log n)^2/\log n)$ -time randomized 3SUM algorithm.

A. A Simple Subquadratic 3SUM Algorithm

Choose the group size to be $g = \Theta(\sqrt{\log n/\log \log n})$. The algorithm enumerates every permutation $\pi : [g^2] \rightarrow [g^2]$, where $\pi = (\pi_r, \pi_c)$ is decomposed into row and column functions $\pi_r, \pi_c : [g^2] \rightarrow [g]$. By definition π is the correct sorting permutation iff $A_{i,j}(\pi(t)) < A_{i,j}(\pi(t+1))$ for all $t \in [g^2-1]$.⁴ Since $A_{i,j} = A_i + A_j$ this inequality can also be written $A_i(\pi_r(t)) + A_j(\pi_c(t)) < A_i(\pi_r(t+1)) + A_j(\pi_c(t+1))$. By Fredman's trick this is equivalent to saying that the (red) point p_j dominates the (blue) point q_i , where

$$\begin{aligned} p_j &= (\dots, A_j(\pi_c(t+1)) - A_j(\pi_c(t)), \dots) \\ q_i &= (\dots, A_i(\pi_r(t)) - A_i(\pi_r(t+1)), \dots). \end{aligned}$$

We find all such dominating pairs. By Lemma 10 the time to report red/blue dominating pairs, over all $(g^2)!$ invocations of the procedure, is $O\left((g^2)!c^{g^2-1}(2n/g)^{1+\epsilon} + (n/g)^2\right)$, the last term being the total size of the outputs. For $\epsilon = 1/2$ and $g = \frac{1}{2}\sqrt{\log n/\log \log n}$ the first term is negligible. The total running time is therefore $O((n/g)^2)$ for dominance reporting and $O(n^2 \log g/g) = O(n^2(\log \log n)^{3/2}/(\log n)^{1/2})$ for the binary searches in Step 4.2.1.

Since there are at most g^{8g} realizable permutations of $A_{i,j}$, not $(g^2)!$ (see Lemma 8 and Fredman [20]), we could possibly shave off another $\sqrt{\log \log n}$ factor by setting $g = \Theta(\sqrt{\log n})$. However, with a bit more work it is possible to save poly($\log n$) factors, as we now show.

B. A Faster 3SUM Algorithm

To improve the running time of the simple algorithm we must sort larger boxes. Our approach is to partition the blocks into layers and sort each layer separately. So long as each layer has size $\Theta(\log n)$, the cost of red/blue dominance reporting will be negligible. The main difficulty is that the natural boundaries between layers are unknown and different for each of the blocks in $\{A_{i,j}\}$.

Let $P \subset [g]^2$ be a set of p positions in the $g \times g$ grid that includes the corner positions $(0,0)$ and $(g-1, g-1)$. How we select the remaining $p-2$ positions in P will be addressed later. For each $(l,m) \in P$, consider $\text{CONTOUR}(A_{i,j}(l,m), A_{i,j})$, that is, the path in $[g]^2$ taken by the standard 3SUM algorithm of Section III when searching for $A_{i,j}(l,m)$ inside $A_{i,j}$. Clearly $\text{CONTOUR}(A_{i,j}(l,m))$ goes through position (l,m) . For any two $(l,m), (l',m') \in P$, $\text{CONTOUR}(A_{i,j}(l,m))$ and $\text{CONTOUR}(A_{i,j}(l',m'))$ may intersect in several places (see Figure 1) though they never cross. According to Lemma 11,

⁴Without loss of generality we can assume $A_{i,j}$ is totally ordered. See Lemma 9.

250	272	362	368	372	385	416	546	549	606
289	311	401	407	411	424	455	585	588	645
299	321	411	417	421	434	465	595	598	655
311	333	423	429	433	446	477	607	610	667
325	347	437	443	447	460	491	621	624	681
331	353	443	449	453	466	497	627	630	687
363	385	475	481	485	498	529	659	662	719
384	406	496	502	506	519	550	680	683	740
412	434	524	530	534	547	578	708	711	768
415	437	527	533	537	550	581	711	714	771

Fig. 1. A tripartition of a block defined by two contours (search paths). The contours of the elements at positions (3,5) and (8,6) are blue and yellow, respectively, and their intersection is green.

the two contours define a *tripartition* (R, S, T) of the positions of $[g]^2$ into three regions, where

$$\begin{aligned} A_{i,j}(R) &\subset (-\infty, A_{i,j}(l, m)] \\ A_{i,j}(S) &\subset (A_{i,j}(l, m), A_{i,j}(l', m')) \\ A_{i,j}(T) &\subset [A_{i,j}(l', m'), \infty). \end{aligned}$$

Here $A_{i,j}(X) = \{A_{i,j}(x) \mid x \in X\}$ for a subset $X \subseteq [g]^2$. Note that (R, S, T) is fully determined by the shapes of the contours, not the specific contents of $A_{i,j}$.⁵ See Figure 2(a) for an illustration.

A contour is defined by at most $2g-1$ comparisons between the search element and elements of the block. Suppose that $\tau = (\tau_r, \tau_c)$ is purported to be $\text{CONTOUR}(A_{i,j}(l, m))$, that is, $\tau(0) = (\tau_r(0), \tau_c(0)) = (0, g-1)$ is the starting position of (lo, hi) and $\tau(t+1) \in \tau(t) + \{(1,0), (0,-1)\}$ depending on whether lo is incremented or hi is decremented after the $(t+1)$ th comparison. The contour ends at the first t^* for which $\tau(t^*) = (g, \cdot)$ or $(\cdot, -1)$ depending on whether the search for $A_{i,j}(l, m)$ falls off the southern or western boundary of $A_{i,j}$. Clearly τ is the correct contour if and only if

$$\begin{aligned} A_{i,j}(l, m) &< A_{i,j}(\tau(t)) && \text{when } \tau(t+1) = \tau(t) + (0, -1) \\ A_{i,j}(l, m) &> A_{i,j}(\tau(t)) && \text{when } \tau(t+1) = \tau(t) + (1, 0) \end{aligned}$$

for every $t \in [t^*]$, excluding the t for which $\tau(t) = (l, m)$ since in this case we have equality: $A_{i,j}(l, m) = A_{i,j}(l, m)$. Restating this, τ is the correct contour if the (red) point p_j dominates the (blue) point q_i , defined as

$$\begin{aligned} p_j &= (\dots, \sigma(t)(A_j(\tau_c(t)) - A_j(m)), \dots) \\ q_i &= (\dots, \sigma(t)(A_i(l) - A_i(\tau_r(t))), \dots), \end{aligned}$$

where $\sigma(t) \in \{1, -1\}$ is the proper sign:

$$\sigma(t) = \begin{cases} 1 & \text{when } \tau(t+1) = \tau(t) + (0, -1), \text{ and} \\ -1 & \text{when } \tau(t+1) = \tau(t) + (1, 0). \end{cases}$$

The coordinate t for which $\tau(t) = (l, m)$ is, of course, omitted from p_j and q_i , so both vectors have length at most $2g-2$.

Call a pair (τ, τ') of contours *legal* if

- (i) Whenever τ and τ' do not intersect, τ is above τ' .

⁵We continue to assume that ties are broken to make $A_{i,j}$ totally ordered.

- (ii) There are two $(l, m), (l', m') \in P$ such that τ contains (l, m) and τ' contains (l', m') .
- (iii) Let (R, S, T) be the tripartition of $[g]^2$ defined by (τ, τ') , where S are those positions lying strictly between (l, m) and (l', m') . Then $P \cap S = \emptyset$ and $|S| \leq s$, where s is a parameter to be determined.

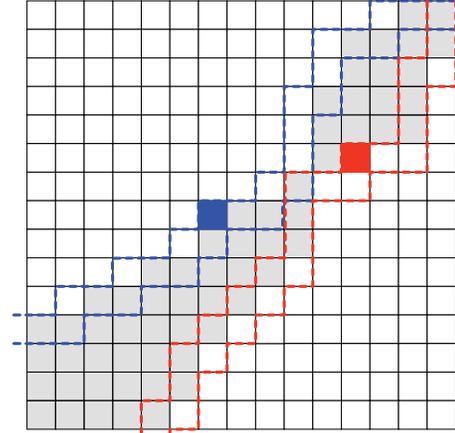
Let us clarify criterion (iii). It states that if $A_{i,j}$ is any *specific* box for which τ and τ' are correct contours of $A_{i,j}(l, m)$ and $A_{i,j}(l', m')$, the number of positions $(l'', m'') \in [g]^2$ for which $A_{i,j}(l'', m'') \in (A_{i,j}(l, m), A_{i,j}(l', m'))$ is at most s , and no such position appears in P .

Our algorithm enumerates every legal pair (τ, τ') of contours, at most 2^{4g} in total. Let $(l, m), (l', m') \in P$ be the points lying on τ, τ' and (R, S, T) be the tripartition of $[g]^2$ defined by (τ, τ') . For each (τ, τ') the algorithm enumerates every realizable permutation $\pi : [|S|] \rightarrow S$ of the elements at positions in S . By Lemma 8 there are $O\left(\binom{s}{2}^{2g}\right) < 2^{4g \log s}$ such permutations, which can be enumerated in $O(2^{4g \log s})$ time. For each (τ, τ', π) we create red points $\{p_j\}_{j \in [n/g]}$ and blue points $\{q_i\}_{i \in [n/g]}$ in \mathbb{R}^{4g+s-5} such that p_j dominates q_i iff $\tau = \text{CONTOUR}(A_{i,j}(l, m))$ and $\tau' = \text{CONTOUR}(A_{i,j}(l', m'))$ are the correct contours (w.r.t. $A_{i,j}$) and π is the correct sorting permutation of $A_{i,j}(S)$. The first $4g-4$ coordinates encode the correctness of τ and τ' and the last $s-1$ coordinates encode the correctness of π .

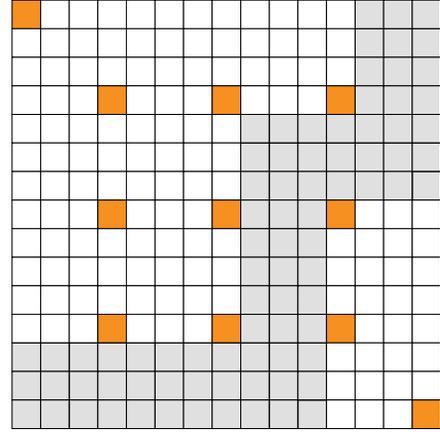
According to Lemma 10, the time to report all dominating pairs is $O(p(n/g)^2 + 2^{4g} \cdot 2^{4g \log s} \cdot (c_\epsilon)^{4g+s-5} (2n/g)^{1+\epsilon})$. The first term is the output size, since by criterion (iii) of the definition of *legal*, at most $p-1$ pairs are reported for each of the $([n/g])^2$ boxes. There are $2^{4g} 2^{4g \log s}$ choices for (τ, τ', π) and the dimension of the point set is at most $4g+s-5$, but could be smaller if the contours happen to be short or $|S| < s$. Fixing $\epsilon = 1/2$, if $g \log s$ and $g+s$ are both $O(\log n)$ (with a sufficiently small leading constant) the running time of the algorithm will be dominated by the time spent reporting the output.

Call a box $A_{i,j}$ *bad* if the output of the dominating pairs algorithm fails to determine its sorted order. The only way a box can be bad is if an otherwise legal (τ, τ') with tripartition (R, S, T) were correct for $A_{i,j}$ but failed to be legal because $|S| > s$, leaving the sorted order of $A_{i,j}(S)$ unknown.

If all boxes are not bad we can search for $x \in \mathbb{R}$ in $A_{i,j}$ in $O(\log g)$ time using binary search, as follows. Each box $A_{i,j}$ was associated with a list of $p-1$ triples of the form (τ, τ', π) returned by the dominating pairs algorithm, one for each pair of successive elements in $A_{i,j}(P)$. The first step is to find the predecessor of x in $A_{i,j}(P)$, that is, to find the consecutive $(l, m), (l', m') \in P$ for which $A_{i,j}(l, m) \leq x < A_{i,j}(l', m')$. Let (τ, τ', π) be the triple associated with $(l, m), (l', m')$ and (R, S, T) be the tripartition of (τ, τ') . Each legal, realizable (τ, τ', π) is encoded as a bit string with length $4g(1 + \log s)$, which must fit comfortably in one machine word. Before executing the algorithm proper we build, in $o(n)$ time, a lookup table indexed by tuples (τ, τ', π, r) that contains the location in $[g]^2$ of the element



(a)



(b)

Fig. 2. Illustrations of tripartitions defined by two contours in a $[15] \times [15]$ grid. Top: the blue and red locations are in P . Two possible contours are indicated by blue and red paths. They define a tripartition (R, S, T) with S marked in gray. Bottom: P is chosen to include two corner locations and an evenly spaced $q \times q$ grid. Any tripartition (R, S, T) defined by a legal pair of contours has $P \cap S = \emptyset$, implying that $|S| \leq 2g^2/(q+1)$. An example of an S nearly achieving that size is marked in gray.

with rank r in S , sorted according to π . Using this lookup table it is straightforward to perform a binary search for x in $A_{i,j}(S)$, in $O(\log |S|) = O(\log g)$ time.

C. A Randomized Parameterization of the Algorithm

Throughout let $\delta > 0$ be a sufficiently small constant. In the randomized implementation of our algorithm we choose $g = s = \delta \ln n / \ln \ln n$ and $p = 2 + 3\delta \ln n$. The points $(0, 0)$ and $(g-1, g-1)$ must be in P and the remaining $3\delta \ln n$ points are chosen uniformly at random. With these parameters

the probability of a box being *bad* is sufficiently low to keep the expected cost per search $O(\log g)$.

Lemma 12. *The probability a particular box is bad is at most $1/g$.*

Proof: Let π be the sorted order for some box $A_{i,j}$. The probability $A_{i,j}$ is bad is precisely the probability that there are $s+1$ consecutive elements (according to π) that are not included in P . The probability that this occurs for a particular set of $s+1$ elements is less than $(1 - (p-2)/g^2)^{s+1} < e^{-(p-2)(s+1)/g^2} < e^{-3 \ln \ln n} < 1/g^3$. By a union bound over all $g^2 - (s+1)$ sets of $s+1$ consecutive elements, the probability $A_{i,j}$ is bad is at most $1/g$. ■

The expected time per search is therefore $O(\log g) + 1/g \cdot O(g) = O(\log g)$. By linearity of expectation the expected total running time is $O(p(n/g)^2 + n^2(\log g)/g) = O(n^2(\log \log n)^2 / \log n)$.

Remark 1. We could have set the parameters differently and achieved the same running time. For example, setting $g = p = \Theta(\log n / \log \log n)$ and $s = \Theta(\log n)$ would also work. The advantage of keeping $s = O(\log n / \log \log n)$ is simplicity: we can afford to enumerate all $s!$ permutations of $S \subset [g]^2$ rather than explicitly construct a hyperplane arrangement in order to enumerate only those *realizable* permutations of S .

D. High Probability Bounds

The running time of the algorithm may deviate from its expectation with non-negligible probability since the *badness* events for the boxes $\{A_{i,j}\}$ can be strongly positively correlated. The easiest way to obtain high probability bounds is simply to choose $L = c \log n$ random point sets $\{P_l\}_{l \in [L]}$, estimate the cost of the algorithm under each point set, then execute the algorithm under the point set with the best estimated cost. The first step is to run a truncated version of the algorithm in order to determine which queries will be asked in Step 4.2.1. Rather than answer the query $-A(k) \in A_{i,j}$ we simply record the triple (k, i, j) in a list \mathcal{Q} to be answered later. The running time of the algorithm under P_l is $O(n^2(\log g)/g)$ plus g times the number of bad triples in \mathcal{Q} , that is, those (k, i, j) for which $A_{i,j}$ is bad according to P_l .

Let ϵ_l be the true fraction of bad triples in \mathcal{Q} according to P_l and $\hat{\epsilon}_l$ be the estimate of ϵ_l obtained by the following procedure. Sample $M = cg^2 \ln n$ elements of \mathcal{Q} uniformly at random and for each, test whether the given block is bad according to P_l by sorting its elements, in $O(g^2 \log g)$ time. If X is the number of blocks discovered to be bad, report the estimate $\hat{\epsilon}_l = X/M$. By a standard version of the Chernoff bound⁶ we have

$$\begin{aligned} \Pr(|\hat{\epsilon}_l - \epsilon_l| > 1/g) \\ = \Pr(|X - \mathbb{E}(X)| > M/g) < 2e^{-2(M/g)^2/M} = 2n^{-2c}. \end{aligned}$$

⁶If X is the number of successes in n independent Bernoulli trials, $\Pr(X > \mathbb{E}(X) + t)$ and $\Pr(X < \mathbb{E}(X) - t)$ are both upper bounded by $e^{-2t^2/n}$. See [17, Thm. 1.1].

By Lemma 12, $\mathbb{E}(\epsilon_l) \leq 1/g$ for each l , so by Markov's inequality $\Pr(\min_{l \in [L]} \{\epsilon_l\} \leq 2/g) \geq 1 - 2^{-L} = 1 - n^{-c}$. With high probability, each $\hat{\epsilon}_l$ deviates from ϵ_l by at most $1/g$, so the running time of the algorithm will be within a constant factor of its expectation with probability $1 - O(n^{-c})$. The time to pick the best point set P_{l^*} , l^* being $\operatorname{argmin}_{l \in [L]} \{\hat{\epsilon}_l\}$, is $O(LMg^2 \log g) = o(\log^6 n)$. We could set L and M as high as $n/\operatorname{polylog}(n)$, making the probability that the algorithm deviates from its expectation exponentially small, $\exp(-n/\operatorname{polylog}(n))$.

E. A Deterministic Parameterization of the Algorithm

We achieve a subquadratic worst-case 3SUM algorithm by choosing g, s, p , and P such that no block can be bad. Fix $g = (\delta \log n)^{2/3} (\log \log n)^{1/3}$ and $p = 2 + q^2 < (\delta \log n)^{2/3} (\log \log n)^{4/3}$ for an integer q to be determined. Aside from the two obligatory points, P contains an evenly spaced $q \times q$ grid in $[g] \times [g]$. Setting $\Delta = \lfloor \frac{g+1}{q+1} \rfloor$, P is defined as follows. See Figure 2(b).

$$\begin{aligned} P = \{ & (0, 0), (g-1, g-1) \} \\ & \cup \{ (k\Delta - 1, l\Delta - 1) \mid \text{where } 1 \leq k, l \leq q \}. \end{aligned}$$

We now argue that no box can be bad if $s = 2g(\Delta - 1)$. For any legal pair of contours (τ, τ') with tripartition (R, S, T) , no element of P can be contained in S . That is, in any row (or column) containing elements of P , the width (or height) of the band S is at most $\Delta - 1$. Since both τ and τ' are monotone paths in $[g]^2$ (non-decreasing by row, non-increasing by column), we always have $|S| < 2g(\Delta - 1) < 2g^2/(q+1) < 2\delta \log n$. See Figure 2(b) for a worst-case example. For δ sufficiently small the overhead for reporting dominating pairs will be negligible. The overall running time is therefore $O(n^2(\log g)/g) = O(n^2/(\log n / \log \log n)^{2/3})$.

VI. ZERO TRIANGLES

We consider a matrix product called *target-min-plus* that subsumes the $(\min, +)$ -product (aka distance product) and the ZEROTRIANGLE problem of [37]. Given real matrices $A \in (\mathbb{R} \cup \{\infty\})^{r \times s}$, $B \in (\mathbb{R} \cup \{\infty\})^{s \times t}$, and a target matrix $T \in (\mathbb{R} \cup \{-\infty, \infty\})^{r \times t}$, the goal is to compute $C = \odot(A, B, T)$, where

$$\begin{aligned} C(i, j) = \min \{ & A(i, k) + B(k, j) \mid k \in [s] \\ & \text{and } A(i, k) + B(k, j) \geq T(i, j) \} \end{aligned}$$

as well as the matrix of witnesses, that is, the k (if any) for which $C(i, j) = A(i, k) + B(k, j)$. This operation reverts to the $(\min, +)$ -product when $T(i, j) = -\infty$. It can also solve ZEROTRIANGLE on a weighted graph $G = (V, E, w)$ by setting A, B , and T as follows. Let $A(i, j) = B(i, j) = w(i, j)$, where $w(i, j) \stackrel{\text{def}}{=} \infty$ if $(i, j) \notin E$, and let $T(i, j) = -w(i, j)$ if $(i, j) \in E$ and ∞ otherwise. If $C(i, j) = T(i, j)$ then there is a zero weight triangle containing (i, j) and the witness matrix gives the third corner of the triangle.

The trivial target-min-plus algorithm runs in $O(rst)$ time and performs the same number of comparisons. We can

compute the target-min-plus product using fewer comparisons using Fredman’s trick.

Theorem 13. *The decision-tree complexity of the target-min-plus product of three $n \times n$ matrices is $O(n^{5/2}\sqrt{\log n})$. This product can be computed in $O(n^3(\log \log n)^2/\log n)$ time.*

Proof: We first show that the target-min-plus product $\odot(A, B, T)$ can be determined with $O((r+t)s^2 + rt \log s)$ comparisons, where A, B , and T are $r \times s, s \times t$, and $r \times t$ matrices, respectively. Begin by sorting the set

$$D = \{A(i, k) - A(i, k'), B(k', j) - B(k, j) \mid i \in [r], j \in [t], \text{ and } k, k' \in [s]\}.$$

By Lemma 8 the number of comparisons required to sort D is $O(|D| + (r+t)s \log(rst)) = O((r+t)s^2 + (r+t)s \log(rst))$. We can now deduce the sorted order on

$$S(i, j) = \{A(i, k) + B(k, j) \mid k \in [s]\},$$

for any pair $(i, j) \in [r] \times [t]$, and can therefore find $C(i, j) = \min(S(i, j) \cap [T(i, j), \infty))$ with a binary search over $S(i, j)$ using $\log s$ additional comparisons. The total number of comparisons is $O((r+t)s^2 + rt \log s)$. Note that this provides no improvement when A and B are square, that is, when $r = s = t = n$. Following Fredman [21] we partition A and B into rectangular matrices and compute their target-min-plus products separately.

Choose a parameter g and partition A into $A_0, \dots, A_{\lfloor n/g \rfloor - 1}$ and B into $B_0, \dots, B_{\lfloor n/g \rfloor - 1}$ where A_ℓ contains columns $\ell g, \dots, (\ell + 1)g - 1$ of A and B_ℓ contains the corresponding rows of B . For each $\ell \in [n/g]$, compute the target-min-plus product $C_\ell = \odot(A_\ell, B_\ell, T)$ and set $C(i, j) = \min_{\ell \in [n/g]}(C_\ell(i, j))$. This algorithm performs $O((n/g) \cdot (ng^2 + n^2 \log n))$ comparisons to compute $\{C_\ell\}_{\ell \in [n/g]}$ and $n^2(n/g)$ comparisons to compute C . When $g = \sqrt{n \log n}$ the number of comparisons is $O(n^{5/2}\sqrt{\log n})$.

To compute the product efficiently we use the geometric dominance approach of Chan [14] and Bremner et al. [11]. Choose a parameter $g = \Theta(\log n / \log \log n)$ and partition A into $n \times g$ matrices $\{A_\ell\}$ and B into $g \times n$ matrices $\{B_\ell\}$. For each $\ell \in [n/g]$ and permutation $\pi : [g] \rightarrow [g]$ we will find those pairs $(i, j) \in [n]^2$ for which π is the sorted order on $\{A_\ell(i, k) + B_\ell(k, j) \mid k \in [g]\}$. Such a triple satisfies the inequality $A_\ell(i, \pi(k)) + B_\ell(\pi(k), j) < A_\ell(i, \pi(k+1)) + B_\ell(\pi(k+1), j)$, for all $k \in [g-1]$. By Fredman’s trick this is equivalent to saying that the (red) point

$$(\dots, A_\ell(i, \pi(k+1)) - A_\ell(i, \pi(k)), \dots)$$

dominates the (blue) point

$$(\dots, B_\ell(\pi(k), j) - B_\ell(\pi(k+1), j), \dots)$$

in each of the $g-1$ coordinates. By Lemma 10 the total time for all $[n/g] \cdot g!$ invocations of the dominance algorithm is $O((n/g) \cdot g! \cdot c_\epsilon^{g-1} (2n/g)^{1+\epsilon})$ plus the output size, which is precisely $n^2 \lfloor n/g \rfloor$. For $\epsilon = 1/2$ and $g = \Theta(\log n / \log \log n)$

the running time is $O(n^3/g)$. We can now compute the target-min-plus product $C_\ell = \odot(A_\ell, B_\ell, T)$ in $O(n^2 \log g)$ time by iterating over all $(i, j) \in [n]^2$ and performing a binary search to find the minimum element in $\{A_\ell(i, k) + B_\ell(k, j) \mid k \in [g]\} \cap [T(i, j), \infty)$. Since $C = \odot(A, B, T)$ contains the pointwise minima of $\{C_\ell\}$, the total time to compute the target-min-plus product is $O(n^3(\log g)/g) = O(n^3(\log \log n)^2/\log n)$. ■

The $\sqrt{\log n}$ factor in the decision tree complexity of target-min-plus comes from the binary searches, n/g searches per pair $(i, j) \in [n]^2$. If the searches were sufficiently correlated (either for fixed (i, j) or fixed ℓ) then there would be some hope that we could evade the information theoretic lower bound of $\Omega(\log g)$ per search. Using random sampling we form a hierarchy of rectangular target-min-plus products such that the solutions at one level gives a hint for the solutions at the next lower level. The cost of finding the solution, given the hint from the previous level, is $O(1)$ in expectation. The same approach lets us shave off another $\log \log n$ factor off the algorithmic complexity of target-min-plus. Refer to [23] for proof of Theorem 14.

Theorem 14. *The randomized decision tree complexity of the target-min-plus product of three $n \times n$ matrices is $O(n^{5/2})$. It can be computed in $O(n^3 \log \log n / \log n)$ time w.h.p.*

VII. CONCLUSIONS AND OPEN PROBLEMS

The original applications of the 3SUM conjecture were to prove *conditional* $\Omega(n^2)$ lower bounds on both the decision tree complexity and algorithmic complexity of geometric problems. Our results refute the conjecture (very strongly in the case of decision tree complexity) and therefore reopen several fundamental questions in computational geometry. For example, what is the (algebraic) decision tree complexity of detecting three colinear points in a subset of \mathbb{R}^2 ? Is there a subquadratic algorithm for this problem?

The most intriguing open question is whether 3SUM has a truly subquadratic algorithm, or even one running in time $O(n^2/2^{(\log n)^\delta})$, for some $\delta > 0$. The geometric dominance approach we use seems incapable of obtaining more than a $\text{poly}(\log n)$ -speedup, but there is hope that techniques employed in Williams’s (min, +)-product algorithm [35] can be repurposed to solve 3SUM and its cousins, such as the *target-min-plus* product defined in Section VI.

ACKNOWLEDGMENT

This research was performed while S. Pettie was on sabbatical at the Center for Massive Data Algorithmics (MADALGO) at Aarhus University, which is supported by the Danish National Research Foundation grant DNR84. A. Grønland is supported in part by a grant from InnovationsFonden. S. Pettie is supported by NSF grants CCF-1217338 and CNS-1318294 and a grant from the US-Israel Binational Science Foundation.

REFERENCES

- [1] A. Abboud and K. Lewi. Exact weight subgraphs and the k -sum conjecture. In *Proceedings of the 40th Int’l Colloquium on Automata, Languages, and Programming (ICALP)*, pages 1–12, 2013.

- [2] A. Abboud and V. Vassilevska Williams. Popular conjectures imply strong lower bounds for dynamic problems. In *Proceedings 55th IEEE Symposium on Foundations of Computer Science (FOCS)*, 2014.
- [3] A. Abboud, V. Vassilevska Williams, and O. Weimann. Consequences of faster sequence alignment. In *Proceedings 41st Int'l Colloquium on Automata, Languages, and Programming (ICALP)*, pages 39–51, 2014.
- [4] O. Aichholzer, F. Aurenhammer, E. D. Demaine, F. Hurtado, P. Ramos, and J. Urrutia. On k -convex polygons. *Comput. Geom.*, 45(3):73–87, 2012.
- [5] N. Ailon and B. Chazelle. Lower bounds for linear degeneracy testing. *J. ACM*, 52(2):157–171, 2005.
- [6] A. Amir, T. M. Chan, M. Lewenstein, and N. Lewenstein. Consequences of faster sequence alignment. In *Proceedings 41st Int'l Colloquium on Automata, Languages, and Programming (ICALP)*, pages 114–125, 2014.
- [7] I. Baran, E. D. Demaine, and M. Pătrașcu. Subquadratic algorithms for 3SUM. *Algorithmica*, 50(4):584–596, 2008.
- [8] G. Barequet and S. Har-Peled. Polygon containment and translational min-Hausdorff-distance between segment sets are 3SUM-hard. *Int. J. Comput. Geometry Appl.*, 11(4):465–474, 2001.
- [9] A. Björklund, R. Pagh, V. Vassilevska Williams, and U. Zwick. Listing triangles. In *Proceedings 41st Int'l Colloquium on Automata, Languages, and Programming (ICALP)*, pages 223–234, 2014.
- [10] M. Blum, R. W. Floyd, V. Pratt, R. L. Rivest, and R. E. Tarjan. Time bounds for selection. *J. Comput. Syst. Sci.*, 7(4):448–461, 1973.
- [11] D. Bremner, T. M. Chan, E. D. Demaine, J. Erickson, F. Hurtado, J. Iacono, S. Langerman, M. Pătrașcu, and P. Taslakian. Necklaces, convolutions, and $X + Y$. *Algorithmica*, 69:294–314, 2014.
- [12] R. C. Buck. Partition of space. *Amer. Math. Monthly*, 50:541–544, 1943.
- [13] A. Butman, P. Clifford, R. Clifford, M. Jalsenius, N. Lewenstein, B. Porat, E. Porat, and B. Sach. Pattern matching under polynomial transformation. *SIAM J. Comput.*, 42(2):611–633, 2013.
- [14] T. M. Chan. All-pairs shortest paths with real weights in $o(n^3/\log n)$ time. *Algorithmica*, 50(2):236–243, 2008.
- [15] S. Chechik, D. Larkin, L. Roditty, G. Schoenebeck, R. E. Tarjan, and V. Vassilevska Williams. Better approximation algorithms for the graph diameter. In *Proceedings of the 25th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1041–1052, 2014.
- [16] K.-Y. Chen, P.-H. Hsu, and K.-M. Chao. Approximate matching for run-length encoded strings is 3SUM-hard. In *Combinatorial Pattern Matching*, volume 5577 of *Lecture Notes in Computer Science*, pages 168–179. Springer, 2009.
- [17] D. P. Dubhashi and A. Panconesi. *Concentration of Measure for the Analysis of Randomized Algorithms*. Cambridge University Press, 2009.
- [18] H. Edelsbrunner, J. O'Rourke, and R. Seidel. Constructing arrangements of lines and hyperplanes with applications. *SIAM J. Comput.*, 15(2):341–363, 1986.
- [19] J. Erickson. Bounds for linear satisfiability problems. *Chicago J. Theor. Comput. Sci.*, 1999.
- [20] M. L. Fredman. How good is the information theory bound in sorting? *Theoretical Computer Science*, 1(4):355–361, 1976.
- [21] M. L. Fredman. New bounds on the complexity of the shortest path problem. *SIAM J. Comput.*, 5(1):83–89, 1976.
- [22] A. Gajentaan and M. H. Overmars. On a class of $O(n^2)$ problems in computational geometry. *Comput. Geom.*, 5:165–185, 1995.
- [23] A. Grønlund and S. Pettie. Threesomes, degenerates, and love triangles. *CoRR*, abs/1404.0799, 2014.
- [24] J. Hartmanis and R. E. Stearns. On the computational complexity of algorithms. *Trans. Amer. Math. Soc.*, 117:285–306, 1965.
- [25] Z. Jafarholi and E. Viola. 3SUM, 3XOR, triangles. *CoRR*, abs/1305.3827, 2013.
- [26] J.-L. Lambert. Sorting the sums $(x_i + y_j)$ in $O(n^2)$ comparisons. *Theor. Comput. Sci.*, 103(1):137–141, 1992.
- [27] S. Meiser. Point location in arrangements of hyperplanes. *Information and Computation*, 106(2):286–303, 1993.
- [28] F. Meyer auf der Heide. A polynomial linear search algorithm for the n -dimensional knapsack problem. *J. ACM*, 31(3):668–676, 1984.
- [29] F. P. Preparata and M. I. Shamos. *Computational Geometry*. Springer, New York, NY, 1985.
- [30] M. Pătrașcu. Towards polynomial lower bounds for dynamic problems. In *Proceedings 42nd ACM Symposium on Theory of Computing (STOC)*, pages 603–610, 2010.
- [31] M. Pătrașcu and R. Williams. On the possibility of faster SAT algorithms. In *Proceedings of the 21st Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1065–1075, 2010.
- [32] L. Roditty and V. Vassilevska Williams. Fast approximation algorithms for the diameter and radius of sparse graphs. In *Proceedings 45th ACM Symposium on Theory of Computing (STOC)*, pages 515–524, 2013.
- [33] M. A. Soss, J. Erickson, and M. H. Overmars. Preprocessing chains for fast dihedral rotations is hard or even impossible. *Comput. Geom.*, 26(3):235–246, 2003.
- [34] W. L. Steiger and I. Streinu. A pseudo-algorithmic separation of lines from pseudo-lines. *Info. Proc. Lett.*, 53(5):295–299, 1995.
- [35] R. Williams. Faster all-pairs shortest paths via circuit complexity. In *Proceedings of the 46th Annual ACM Symposium on Theory of Computing (STOC)*, pages 664–673, 2014.
- [36] V. Vassilevska Williams and R. Williams. Subcubic equivalences between path, matrix and triangle problems. In *Proc. 51th IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 645–654, 2010.
- [37] V. Vassilevska Williams and R. Williams. Finding, minimizing, and counting weighted subgraphs. *SIAM J. Comput.*, 42(3):831–854, 2013.

APPENDIX

We shall review a standard divide and conquer dominating pairs algorithm [29, p. 366] and give a short proof of Lemma 10 due to Chan [14].

We are given n red and blue points in $P \subset \mathbb{R}^d$, at least one of each color, and wish to report all pairs (p, q) where $p = (p_i)_{i \in [d]}$ is red, $q = (q_i)_{i \in [d]}$ is blue and $p_i \geq q_i$ for each $i \in [d]$. When $d = 0$ the algorithm simply reports every pair of points, so assume $d \geq 1$. Find the median h on the last coordinate in $O(n)$ time [10] and partition P into disjoint sets P_L, P_R of size at most $\lceil n/2 \rceil$, where

$$P_L \subset \{p \in P \mid p_{d-1} \leq h\}$$

$$P_R \subset \{p \in P \mid p_{d-1} \geq h\}.$$

Furthermore, there cannot be a red $p \in P_L$ and blue $q \in P_R$ such that $p_{d-1} = q_{d-1} = h$. At this point all dominating pairs are in P_L , or P_R , or have one point in each, in which case the blue point is necessarily in P_L and the red in P_R . We make three recursive calls to find dominating pairs of each variety. The first two calls are on $\lceil n/2 \rceil$ points in \mathbb{R}^d . The third recursive call is on all blue points in P_L and all red points in P_R ; after stripping their last coordinate they lie in \mathbb{R}^{d-1} .

Excluding the cost of reporting the output, the running time of this algorithm is bounded by $T_d(n)$, defined inductively as

$$T_0(n) = T_d(1) = 0$$

$$T_d(n) = 2T_d(n/2) + T_{d-1}(n) + n.$$

We prove by induction that $T_d(n) \leq c_\epsilon^d n^{1+\epsilon} - n$, a bound which holds in all base cases. Assuming the claim holds for all smaller values of d and n ,

$$\begin{aligned} T_d(n) &\leq 2(c_\epsilon^d (n/2)^{1+\epsilon} - n/2) + (c_\epsilon^{d-1} n^{1+\epsilon} - n) + n \\ &= (c_\epsilon^d / 2^\epsilon + c_\epsilon^{d-1}) n^{1+\epsilon} - n \\ &= (1/2^\epsilon + 1/c_\epsilon) \cdot c_\epsilon^d n^{1+\epsilon} - n \\ &= c_\epsilon^d n^{1+\epsilon} - n \quad \{\text{By defn. of } c_\epsilon = 2^\epsilon / (2^\epsilon - 1).\} \end{aligned}$$