Monochromatic Triangles, Intermediate Matrix Products, and Convolutions

Andrea Lincoln
CSAIL, MIT, Cambridge MA, USA
andreali@mit.edu

Adam Polak
Institute of Theoretical Computer Science, Faculty of Mathematics and Computer Science, Jagiellonian University, Krakow, Poland
polak@tcs.uj.edu.pl

Virginia Vassilevska Williams
CSAIL, MIT, Cambridge MA, USA
virgi@mit.edu

Abstract

The most studied linear algebraic operation, matrix multiplication, has surprisingly fast $O(n^\omega)$ time algorithms for $\omega < 2.373$. On the other hand, the $(\min, +)$ matrix product which is at the heart of many fundamental graph problems such as All-Pairs Shortest Paths, has received only minor $n^{o(1)}$ improvements over its brute-force cubic running time and is widely conjectured to require $n^{3-o(1)}$ time. There is a plethora of matrix products and graph problems whose complexity seems to lie in the middle of these two problems. For instance, the Min-Max matrix product, the Minimum Witness matrix product, All-Pairs Shortest Paths in directed unweighted graphs and determining whether an edge-colored graph contains a monochromatic triangle, can all be solved in $\widetilde{O}(n^{(3+\omega)/2})$ time. While slight improvements are sometimes possible using rectangular matrix multiplication, if $\omega = 2$, the best runtimes for these "intermediate" problems are all $\widetilde{O}(n^{2.5})$.

A similar phenomenon occurs for convolution problems. Here, using the FFT, the usual $(+, \times)$-convolution of two n-length sequences can be solved in $O(n \log n)$ time, while the $(\min, +)$ convolution is conjectured to require $n^{2-o(1)}$ time, the brute force running time for convolution problems. There are analogous intermediate problems that can be solved in $O(n^{1.5})$ time, but seemingly not much faster: Min-Max convolution, Minimum Witness convolution, etc.

Can one improve upon the running times for these intermediate problems, in either the matrix product or the convolution world? Or, alternatively, can one relate these problems to each other and to other key problems in a meaningful way?

This paper makes progress on these questions by providing a network of fine-grained reductions. We show for instance that APSP in directed unweighted graphs and Minimum Witness product can be reduced to both the Min-Max product and a variant of the monochromatic triangle problem, so that a significant improvement over $n^{(3+\omega)/2}$ time for any of the latter problems would result in a similar improvement for both of the former problems. We also show that a natural convolution variant of monochromatic triangle is fine-grained equivalent to the famous 3SUM problem. As this variant is solvable in $O(n^{1.5})$ time and 3SUM is in $O(n^2)$ time (and is conjectured to require $n^{2-o(1)}$ time), our result gives the first fine-grained equivalence between natural problems of different running times. We also relate 3SUM to monochromatic triangle, and a coin change problem to monochromatic convolution, and thus to 3SUM.

2012 ACM Subject Classification Theory of computation → Problems, reductions and completeness

Keywords and phrases 3SUM, fine-grained complexity, matrix multiplication, monochromatic triangle

Digital Object Identifier 10.4230/LIPIcs.ITCS.2020.53

Funding *Andrea Lincoln*: Partially supported by NSF Grant CCF-1909429.
Adam Polak: Partially supported by the National Science Center, Poland under grants 2017/27/N/ST6/01334 and 2018/28/T/ST6/00305.

Virginia Vassilevska Williams: Supported by an NSF CAREER Award, NSF Grants CCF-1528078, CCF-1514339 and CCF-1909429, a BSF Grant BSF:2012338, a Google Research Fellowship and a Sloan Research Fellowship.

Acknowledgements We would like to thank Amir Abboud for fruitful discussions at an early stage of our research. Part of the research was done when the second author was visiting MIT.

1 Introduction

Matrix multiplication is arguably the most fundamental linear algebraic operation. It is an important primitive for an enormous variety of applications. Within algorithmic research it has a very special role since it is one of the few problems for which we have surprisingly fast and completely counter-intuitive algorithms. Starting with Strassen's breakthrough [36] in 1969, a long line of research culminated in the current bound $\omega < 2.373$ [42, 30], where ω is the smallest real number so that $n \times n$ matrix multiplication can be performed in $\mathcal{O}(n^{\omega+\varepsilon})$ time for all $\varepsilon > 0$.

In many applications, one needs to compute matrix products that are a bit different (often called funny [2]) from the usual definition of matrix multiplication over a ring such as the integers ($C_{ij} = \sum_k A_{ik} \cdot B_{kj}$). Such examples include matrix products over semirings such as the $(\min, +)$-product (often called distance product) which is over the tropical $((\min, +))$ semiring and the Max-Min product which is over the $(\max, \min)$-semiring. Both these products are equivalent to certain types of path optimization problems in graphs. The distance product of $n \times n$ matrices is equivalent to the All-Pairs Shortest Paths (APSP) problem in n-node graphs, so that a $T(n)$ time algorithm for one problem would imply an $\mathcal{O}(T(n))$ time algorithm for the other [21]. Similarly, the Max-Min product is equivalent to the so called All-Pairs Bottleneck Paths (APBP) in graphs (e.g. [35]).

There seems to be a distinct complexity difference between APSP and APBP (and hence the corresponding matrix products), however. The fastest algorithms for APSP and the distance product run in $n^3/\exp(\sqrt{\log n})$ time [46], which is only better by an $n^{o(1)}$ factor than the trivial cubic time algorithm for the distance product. Meanwhile, as was first shown by [38, 39], APBP and the Max-Min product admit a much faster than cubic time algorithm via a reduction to (normal) matrix multiplication; the fastest running time is $\mathcal{O}(n^{(3+\omega)/2})$ [19].

APSP is in fact conjectured to not admit any truly subcubic, $\mathcal{O}(n^{3-\varepsilon})$ time algorithms for $\varepsilon > 0$. Fine-grained complexity has strengthened this hypothesis by providing a large class of problems that are equivalent to APSP and the distance product, via fine-grained subcubic reductions. Thus the reason why distance product is seemingly so difficult is because there are many problems that are equivalent to it and researchers from different communities have all failed to solve these problems faster.

The best known running time for the $n \times n$ Max-Min product, $\widetilde{\mathcal{O}}(n^{(3+\omega)/2})$, while nontrivially subcubic, seems difficult to improve upon. In fact, $\widetilde{\mathcal{O}}(n^{(3+\omega)/2})$ is the best known running time for many other matrix and graph problems besides the Max-Min product: the Dominance product [32] and Equality product [47, 29], All-Pairs Nondecreasing Paths (APNP) and the $(\min, \leqslant)$-product [37, 41, 18]. For some of these problems [49, 24] one can obtain slightly improved running times using rectangular matrix multiplication [23]. However, the closer ω is to 2, the smaller the improvement, and when $\omega = 2$, the $\widetilde{\mathcal{O}}(n^{(3+\omega)/2}) = \widetilde{\mathcal{O}}(n^{2.5})$ running time is the best known for all of these problems. Since their running time exponent is essentially the average of the brute-force exponent 3 and the fast matrix multiplication exponent ω, we will call these problems "intermediate".

There are two additional problems that are intermediate if $\omega = 2$: the Minimum Witness product, which is equivalent to the problem of computing All-Pairs Least Common Ancestors in a DAG, and All-Pairs Shortest Paths (APSP) in unweighted directed graphs. For both problems we know algorithms running in $\widetilde{\mathcal{O}}(n^{(3+\omega)/2}) \leqslant \widetilde{\mathcal{O}}(n^{2.687})$ time [6, 2], and both algorithms can be improved upon, by using rectangular matrix multiplication [15, 50]. The improvement is already seen in a naive implementation, i.e. cutting rectangular matrices into square blocks, which gives an $\widetilde{\mathcal{O}}(n^{2+1/(4-\omega)}) \leqslant \widetilde{\mathcal{O}}(n^{2.615})$ time. Employing a specialized rectangular matrix multiplication algorithm [23], brings the runtime down to $\widetilde{\mathcal{O}}(n^{2.529})$. When $\omega = 2$, however, all the improvements vanish and those running times become $\widetilde{\mathcal{O}}(n^{2.5})$.

Is the 2.5 *running time exponent (for* $\omega = 2$*) for all of these problems a coincidence, or can we relate all of them via fine-grained reductions, and use plausible hypotheses to explain it?*

This is a question that many have asked, but unfortunately there are only two partial answers: First, it is known that Equality product and Dominance product are equivalent ([47, 29], also follows from Proposition 3.4 in [45]), and that they are equivalent to All-Pairs ℓ_{2p+1} Distances [29]. The second result is that the Max-Min product is equivalent to approximate APSP in weighted graphs without scaling [10]. The main question above remains *wide open*.

Parallel to the world of matrix products, there is a very similar landscape of *convolution* problems. While it is well-known that the $(+,\times)$-convolution[1] of two n-length vectors can be computed in $\mathcal{O}(n \log n)$ time using the Fast Fourier Transform (FFT), these techniques no longer work for the $(\min,+)$-convolution, and this problem is conjectured to require $n^{2-o(1)}$ time (see e.g. [14]). Similar to the "intermediate" matrix product problems, there are analogous "intermediate" convolution problems, all in $\widetilde{\mathcal{O}}(n^{3/2})$ time[2]: Max-Min convolution, Dominance convolution, Minimum Witness convolution, etc.

The convolution landscape is even somewhat cleaner than the matrix product one. As the normal convolution $((+,\times))$ is already in (near-)linear time, there are no analogues of rectangular matrix multiplication speedups, and all intermediate problems happen to have exactly the same running time (up to polylogarithmic factors). Still, there is no real formal explanation of why they have the same running time. The only reductions between these convolutions are analogous to the matrix product ones: Dominance convolution is equivalent to Equality convolution [29], and approximate $(\min,+)$-convolution is equivalent to exact Max-Min convolution [10].

1.1 Our contributions

In this paper we provide new fine-grained reductions between several intermediate matrix product and all-pairs graph problems, and between intermediate convolution problems, also relating these to other key problems from fine-grained complexity such as 3SUM. See Figure 1 for a pictorial representation of our results.

Reductions for Graph Problems and Matrix Products

Several of our reductions concern the All-Edges Monochromatic Triangle (AE-MonoΔ) problem: Given an n-node graph in which each edge has a color from 1 to n^2, decide for each

[1] The $(+,\times)$-convolution of two vectors a and b is the vector c such that $c_i = \sum_j a_j b_{i-j}$.

[2] The exponent $(3+\omega)/2$ for intermediate matrix products is the average of the fast matrix multiplication exponent and the brute force matrix product exponent, and the exponent $3/2$ for intermediate convolution problems is the average of the fast convolution exponent 1 and the brute force exponent 2.

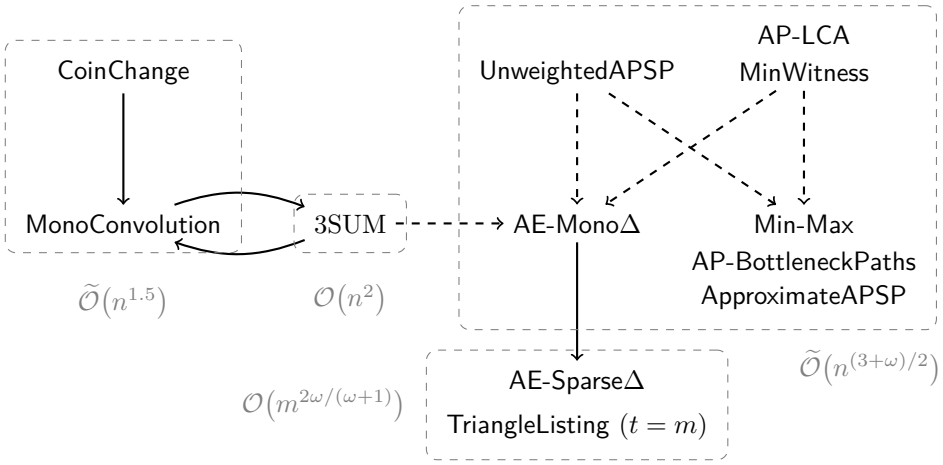

Figure 1 Our results. An arrow pointing from problem A to problem B means that problem A reduces to problem B in the fine-grained sense. Dashed arrows denote reductions which become tight when $\omega = 2$.

edge whether it belongs to a *monochromatic* triangle, a triangle whose all three edges have the same color. Vassilevska, Williams and Yuster [40] studied the decision variant of AE-MonoΔ in which one asks whether the given graph contains a monochromatic triangle. They provided an $\mathcal{O}(n^{(3+\omega)/2})$ time algorithm for the decision problem, but that algorithm is in fact strong enough to also solve the all-edges variant AE-MonoΔ, making AE-MonoΔ one of the "intermediate" problems of interest.

To obtain their $\mathcal{O}(n^{(3+\omega)/2})$ time algorithm, Vassilevska, Williams and Yuster [40] implicitly reduce AE-MonoΔ (in a black-box way) to the AE-SparseΔ problem of deciding for every edge e in an m-edge graph whether e is in a triangle. The fastest known algorithm for AE-SparseΔ is by Alon, Yuster and Zwick [3], running in $\mathcal{O}(m^{2\omega/(\omega+1)})$ time, and the problem is known to be runtime equivalent to the problem of *listing* up to m triangles in an m-edge graph [20]. The black-box reduction of [40] from AE-MonoΔ to AE-SparseΔ implies that a significant improvement over the $\mathcal{O}(m^{2\omega/(\omega+1)})$ time for AE-SparseΔ would translate to an improvement over $\mathcal{O}(n^{(3+\omega)/2})$ for AE-MonoΔ.

▶ **Theorem 1** (implicit in [40]). *If* AE-SparseΔ *is in* $\mathcal{O}(m^{2\omega/(\omega+1)-\varepsilon})$ *time, for some* $\varepsilon > 0$, *then* AE-MonoΔ *is in* $\mathcal{O}(n^{(3+\omega)/2-\delta})$ *time, for some* $\delta > 0$.

Our first set of results shows that AE-MonoΔ is powerful enough to capture two well-studied intermediate problems: the Minimum Witness product of two Boolean matrices and the All-Pairs Shortest Paths problem in directed unweighted graphs.

The Minimum Witness product (MinWitness) C of two Boolean matrices A and B is defined as $C_{ij} = \min\{k \mid A_{ik} = B_{kj} = 1\}$ (where the minimum is defined to be ∞ if there is no witness k). It is well-known [15] that MinWitness is equivalent to determining for every pair u, v of vertices in a DAG, the least common ancestor of u and v, i.e. solving the All-Pairs Least Common Ancestors (AP-LCA) problem. The fastest known algorithm for MinWitness runs in $\mathcal{O}(n^{2.529})$ time using rectangular matrix multiplication, and in $\mathcal{O}(n^{2+1/(4-\omega)})$ time just using square matrix multiplication [15].

The All-Pairs Shortest Paths (APSP) problem in unweighted graphs is very well-studied. While in undirected graphs, the problem is known to be solvable in $\widetilde{\mathcal{O}}(n^\omega)$ time [34], the problem in directed graphs is one of our intermediate problems. Its fastest algorithm

(similarly to MinWitness) runs in $\mathcal{O}(n^{2.529})$ time using rectangular matrix multiplication, and in $\widetilde{\mathcal{O}}(n^{2+1/(4-\omega)})$ time just using square matrix multiplication [50]. We will refer to the APSP problem in directed unweighted graphs as UnweightedAPSP.

We present reductions from MinWitness and UnweightedAPSP to AE-MonoΔ with only polylogarithmic overhead.

▶ **Theorem 2.** *If* AE-MonoΔ *is in* $T(n)$ *time, then* MinWitness *is in* $\mathcal{O}(T(n)\log n)$ *time.*

▶ **Theorem 3.** *If* AE-MonoΔ *is in* $T(n)$ *time, then* UnweightedAPSP *is in* $\mathcal{O}(T(n)\log^2 n)$ *time.*

The above reductions tightly relate MinWitness and UnweightedAPSP to AE-MonoΔ if $\omega = 2$, showing that any improvement over the 2.5 exponent for AE-MonoΔ, gives the same improvement for MinWitness and UnweightedAPSP. Due to the tight reduction (Theorem 1) from AE-MonoΔ to AE-SparseΔ, we also obtain that an $\mathcal{O}(m^{4/3-\varepsilon})$ time algorithm, with $\varepsilon > 0$, for AE-SparseΔ would give $\mathcal{O}(n^{2.5-\delta})$ time algorithms, for $\delta > 0$, for MinWitness and UnweightedAPSP, presenting another tight relationship for the case when $\omega = 2$.

Our next result is that improving over the exponent 2.5 for AE-MonoΔ is at least as hard as obtaining a truly subquadratic time algorithm for the 3SUM problem.

▶ **Theorem 4.** *If* AE-MonoΔ *is in* $\mathcal{O}(n^{5/2-\varepsilon})$ *time, then* 3*SUM is in (randomized)* $\widetilde{\mathcal{O}}\left(n^{2-\frac{4}{5}\varepsilon}\right)$ *time.*

In 3SUM one is given n integers and is asked whether three of them sum to 0. The problem is easy to solve in $\mathcal{O}(n^2)$ time, and slightly subquadratic time algorithms exist [4, 11]. 3SUM is a central problem in fine-grained complexity [43]. It is hypothesized to require $n^{2-o(1)}$ time (on a word-RAM with $\mathcal{O}(\log n)$ bit words), and many fine-grained hardness results are conditioned on this hypothesis (see [22, 43]). Our reduction shows that, under the 3SUM Hypothesis, the exponent 2.5 for AE-MonoΔ cannot be beaten, and this is tight if $\omega = 2$. We note that before our work no intermediate matrix, graph, or convolution problem was known to be 3SUM-hard.

Next, we consider the Min-Max product (Min-Max) of two matrices A and B, defined as $C_{ij} = \min_k \max(A_{ik}, B_{kj})$. The Min-Max product is equivalent to the aforementioned Max-Min product (just negate the matrix entries) and the All-Pairs Bottleneck Paths problem, and is thus solvable in $\mathcal{O}(n^{(3+\omega)/2})$ time [19].

A very simple folklore reduction shows that Min-Max on $n \times n$ integer matrices is at least as hard as MinWitness on $n \times n$ Boolean matrices, giving a tight relationship when $\omega = 2$.

▶ **Theorem 5** (folklore)**.** *If* Min-Max *is in* $T(n)$ *time, then* MinWitness *is in* $\mathcal{O}(T(n))$ *time.*

Our next result states that All-Pairs Shortest Paths in directed unweighted graphs (UnweightedAPSP) is also tightly reducible to Min-Max. This gives a second intermediate problem that is at least as hard as both MinWitness and UnweightedAPSP.

▶ **Theorem 6.** *If* Min-Max *is in* $T(n)$ *time, then* UnweightedAPSP *is in* $\mathcal{O}(T(n)\log n)$ *time.*

The above theorem also follows from a recent independent result by Barr, Kopelowitz, Porat and Roditty [5]. In particular, they reduce All-Pairs Shortest Paths in directed graphs with edge weights from $\{-1, 0, 1\}$ to Min-Max. Interestingly, they use a substantially different approach than ours. While their argument can be seen as inspired by Seidel's algorithm for unweighted APSP in undirected graphs [34], ours resembles Zwick's algorithm for directed graphs [50].

Reductions for Convolution Problems

Our main result for convolution problems regards the convolution version of AE-MonoΔ, which we call MonoConvolution: Given three integer sequences a, b, c, decide for each index i if there exists j such that $a_j = b_{i-j} = c_i$. We show that MonoConvolution is actually fine-grained equivalent to 3SUM.

▶ **Theorem 7.** *If* MonoConvolution *is in* $\mathcal{O}\left(n^{3/2-\varepsilon}\right)$ *time, then* 3SUM *is in (randomized)* $\widetilde{\mathcal{O}}\left(n^{2-\frac{4}{3}\varepsilon}\right)$ *time.*

▶ **Theorem 8.** *If* 3SUM *is in* $\mathcal{O}\left(n^{2-\varepsilon}\right)$ *time, then* MonoConvolution *is in* $\widetilde{\mathcal{O}}\left(n^{3/2-\varepsilon/4}\right)$ *time.*

This equivalence is arguably the first fine-grained equivalence between natural problems with *different* running time complexities: MonoConvolution is a problem in $\mathcal{O}\left(n^{3/2}\right)$ time, whereas 3SUM is in $\mathcal{O}\left(n^2\right)$ time, and a polynomial improvement on one of these running times would result in a polynomial improvement over the other. All previous fine-grained equivalences were between problems with the same running time exponent: the problems equivalent to APSP [44, 1] are all solvable in $\mathcal{O}\left(N^{1.5}\right)$ time where N is the size of their input, the problems equivalent to Orthogonal Vectors [12] or to $(\min, +)$-convolution [14] are all in quadratic time, the problems equivalent to CNF-SAT [13] are all in $\mathcal{O}(2^n)$ time, etc. While tight fine-grained reductions between problems with different running times are well-known, there was no such equivalence until our result, largely since it often seems difficult to reduce a problem with a smaller asymptotic running time to one with a larger running time, something our Theorem 8 overcomes. Note that the same apparent difficulty is overcome by the reduction from AE-MonoΔ to AE-SparseΔ in Theorem 1, as well as by the reductions from MinWitness and UnweightedAPSP to AE-SparseΔ, which follow from combining Theorems 2 and 3 with Theorem 1.

Theorem 8 together with Theorem 4 give a reduction from MonoConvolution to AE-MonoΔ. Previously reductions from a convolution to the corresponding graph/matrix problem were known only for problems with best known algorithms running in brute-force time, i.e. quadratic time for convolution and cubic time for product, e.g. $(\min, +)$-convolution and $(\min, +)$-product [7].

Finally, we relate MonoConvolution to an unweighted variant of a coin change problem [48, 28] that is related to the minimum word break problem [8]. Given a set of coin values from $\{1, 2, \ldots, n\}$, the CoinChange problem asks to determine for each integer value up to n what is the minimum number of coins (allowing repetitions) that sum to that value. We reduce CoinChange to MonoConvolution with only a polylogarithmic overhead. CoinChange is solvable in $\widetilde{\mathcal{O}}\left(n^{3/2}\right)$ time [9], and our reduction implies that any improvement over the known running times of MonoConvolution or 3SUM would also improve the running time for CoinChange.

▶ **Theorem 9.** *If* MonoConvolution *is in* $T(n)$ *time, then* CoinChange *is in* $\mathcal{O}\left(T(n)\log^2 n\right)$ *time.*

2 Preliminaries

In this section we recall formal definitions of all the problems involved in the reductions presented in the paper. We split these problems by their time complexity.

2.1 Problems in $\widetilde{\mathcal{O}}\left(n^{(3+\omega)/2}\right)$ time

▶ **Definition 10** (All-Edges Monochromatic Triangle, AE-MonoΔ). *Given an n-node graph G in which each edge has a color from 1 to n^2, decide for each edge whether it belongs to a* monochromatic *triangle, a triangle where all three edges have the same color.*

▶ **Definition 11** (Min-Max matrix product, Min-Max). *Given two $n \times n$ matrices A and B, compute matrix C such that*

$$C_{ij} = \min_k \max(A_{ik}, B_{kj}).$$

▶ **Definition 12** (Minimum Witness matrix product, MinWitness). *Given two $n \times n$ Boolean matrices A and B, compute matrix C such that*

$$C_{ij} = \min(\{k \mid A_{ik} = B_{kj} = 1\} \cup \{\infty\}).$$

▶ **Definition 13** (All-Pairs Shortest Paths in directed unweighted graphs, UnweightedAPSP). *Given an n-node unweighted directed graph $G = (V, E)$, compute for each pair of vertices $u, v \in V$ the length of a shortest path from u to v. Note that all path lengths will be in $\{0, 1, \ldots, n-1\} \cup \{\infty\}$.*

2.2 Problems in $\mathcal{O}\left(m^{2\omega/(\omega+1)}\right)$ time

▶ **Definition 14** (All-Edges Sparse Triangle, AE-SparseΔ). *Given an m-edge graph G decide for each edge whether it belongs to a triangle.*

2.3 Problems in $\mathcal{O}(n^2)$ time

▶ **Definition 15** (3SUM). *Given three lists, A, B and C, of n integers, determine if there exist $a \in A$, $b \in B$, and $c \in C$ such that $a + b = c$.*

Let us note that the 3SUM problem is defined in several different ways in literature. They differ as to whether the input is split into three list or all the numbers are in a single list, and whether one looks for $a + b = c$ or $a + b + c = 0$. All these variants are equivalent by simple folklore reductions.

2.4 Problems in $\widetilde{\mathcal{O}}(n^{1.5})$ time

▶ **Definition 16** (MonoConvolution). *Given three sequences a, b, c, all of length n, compute the sequence d such that*

$$d_i = \begin{cases} 1 & \text{if } \exists_j\ a_j = b_{i-j} = c_i, \\ 0 & \text{otherwise.} \end{cases}$$

▶ **Definition 17** (CoinChange). *Given a set of coin values $C \subseteq \{1, 2, \ldots, n\}$, assume you have for each $c \in C$ an infinite supply of coins of value c, and determine for each $v \in \{1, 2, \ldots, n\}$ the minimum number of coins that sums up to v.*

CoinChange can be solved in $\widetilde{\mathcal{O}}(n^{1.5})$ time [9]. The algorithm splits the coins into heavy coins, with weight at least $\sqrt{n}$, and light coins, with weight less than $\sqrt{n}$. The minimum sum for a value can use at most $\sqrt{n}$ heavy coins. By running FFT $\sqrt{n}$ times the algorithm produces a vector with the minimum number of heavy coins needed to sum to every value. That takes $\mathcal{O}(n^{1.5} \log n)$ time in total. Then a classical dynamic programming algorithm is run for the $\sqrt{n}$ light coins and n values, in $\mathcal{O}(n^{1.5})$ time.

2.5 Self-reducibility of 3SUM

In our proofs of Theorems 4 and 7 we use the following fact about 3SUM.

▶ **Lemma 18.** *For any $\alpha \in [0,1]$, a single instance of 3SUM of size n can be reduced to $\mathcal{O}(n^{2\alpha})$ instances of 3SUM of size $\mathcal{O}(n^{1-\alpha})$ each. The reduction runs in time linear in the total size of produced instances, and the original instance is a yes-instance if and only if at least one of the produced instances is a yes-instance.*

This fact appears in many 3SUM-related papers, e.g. [4, 26, 27, 31, 33]. In [4] it was proved using a randomized almost linear hashing scheme [17]. An alternative proof – using a domination argument to provide a deterministic reduction – appeared e.g. in [31, 25], and is based on ideas of [16].

3 Reductions for Graph and Matrix Problems

First, let us recall the algorithm of Vassilevska, Williams and Yuster [40] for AE-MonoΔ. We rephrase the argument so that it not only shows how to solve AE-MonoΔ in $\mathcal{O}(n^{(3+\omega)/2})$ time, but also proves that any polynomial improvement over the $\mathcal{O}(m^{2\omega/(\omega+1)})$ time algorithm of Alon, Yuster and Zwick [3] for AE-SparseΔ translates to a polynomial improvement for AE-MonoΔ.

▶ **Theorem 1** (implicit in [40])**.** *If* AE-SparseΔ *is in $\mathcal{O}(m^{2\omega/(\omega+1)-\varepsilon})$ time, for some $\varepsilon > 0$, then* AE-MonoΔ *is in $\mathcal{O}(n^{(3+\omega)/2-\delta})$ time, for some $\delta > 0$.*

Proof. Assume AE-SparseΔ is in $\mathcal{O}(m^\alpha)$ time. Take an AE-MonoΔ instance. For each color consider the subgraph composed of all the edges of that color. Each such subgraph constitutes an independent instance of AE-SparseΔ. However, simply using the $\mathcal{O}(m^\alpha)$ time algorithm on all of these instances is not efficient enough. Intuitively, some of the instances might be too dense.

Instead, for a parameter t to be determined later, take the t largest subgraphs (in terms of the number of edges). For each of them solve the problem by using fast matrix multiplication to compute the square of the adjacency matrix. This takes $\mathcal{O}(tn^\omega)$ time in total. Let m_i denote the number of edges in the i-th of the remaining subgraphs. Clearly, $\forall_i\ m_i \leqslant n^2/t$, and $\sum_i m_i \leqslant n^2$. On each of those subgraphs use the $\mathcal{O}(m^\alpha)$ time AE-SparseΔ algorithm. This takes an order of

$$\sum_i m_i^\alpha = \sum_i m_i \cdot m_i^{\alpha-1} \leqslant \sum_i m_i \cdot (n^2/t)^{\alpha-1} \leqslant n^2 \cdot (n^2/t)^{\alpha-1}$$

time. The total runtime is thus $\mathcal{O}(tn^\omega + n^{2\alpha}/t^{\alpha-1})$. Optimize by setting $t = n^{(2\alpha-\omega)/\alpha}$, and get an $\mathcal{O}(n^{\omega+2-(\omega/\alpha)})$ time.

Observe that for $\alpha = 2\omega/(\omega+1)$ the runtime is $\mathcal{O}(n^{(3+\omega)/2})$. Moreover, for $\alpha < 2\omega/(\omega+1)$ the exponent in the runtime becomes strictly smaller. ◀

Now, we proceed to show how to use AE-MonoΔ to solve two popular graph problems. We start with the problem of finding All-Pairs Least Common Ancestors in directed acyclic graphs, which is runtime-equivalent to MinWitness [15]. We reduce a single instance of MinWitness to $\log n$ instances of AE-MonoΔ.

▶ **Theorem 2.** *If* AE-MonoΔ *is in $T(n)$ time, then* MinWitness *is in $\mathcal{O}(T(n)\log n)$ time.*

Proof. The main idea is to use a parallel binary search. For each entry of the output matrix C we will keep an interval which that entry is guaranteed to lie in. With a single call to AE-MonoΔ we will be able to halve all the intervals.

W.l.o.g. assume the last column of A and last row of B are all ones, so that the output is always finite. For $\ell \in [\log n]$, let $C^{(\ell)}$ denote the matrix pointing to 2^ℓ-length intervals in which entries of C lie, that is $C^{(\ell)}_{ij}$ is the unique integer such that $C_{ij} \in \left[2^\ell \cdot C^{(\ell)}_{ij}, 2^\ell \cdot (C^{(\ell)}_{ij}+1)\right)$.

We will compute $C^{(\ell)}$ for $\ell = \lceil \log n \rceil, \ldots, 1, 0$. Observe that $C^{(\lceil \log n \rceil)}$ is the zero matrix. Knowing $C^{(\ell+1)}$, we compute $C^{(\ell)}$ as follows. We create a tripartite graph $G = (I \cup J \cup K, E)$, with each of I, J, K containing n vertices. We add edges between I and K according to the matrix A. Edges from the k-th column get the label $\lfloor k/2^\ell \rfloor$. We add edges between K and J according to the matrix B. Edges from the k-th row get the label $\lfloor k/2^\ell \rfloor$. Finally, we add the full bipartite clique between I and J. The edge between the i-th vertex of I and the j-th vertex of J gets the label $2 \cdot C^{(\ell+1)}$. That edge forms a monochromatic triangle if and only if $C_{ij} \in \left[2^\ell \cdot 2 \cdot C^{(\ell+1)}_{ij}, 2^\ell \cdot (2 \cdot C^{(\ell+1)}_{ij} + 1)\right)$, i.e. $C^{(\ell)}_{ij} = 2 \cdot C^{(\ell+1)}_{ij}$. Otherwise, it must be that $C_{ij} \in \left[2^\ell \cdot (2 \cdot C^{(\ell+1)}_{ij} + 1), 2^\ell \cdot (2 \cdot C^{(\ell+1)}_{ij} + 2)\right)$, i.e. $C^{(\ell)}_{ij} = 2 \cdot C^{(\ell+1)}_{ij} + 1$. Therefore, solving AE-MonoΔ on G suffices to compute $C^{(\ell)}$. Finally, observe that $C = C^{(0)}$. ◀

With a slightly more involved argument we show how to solve UnweightedAPSP with $\mathcal{O}\left(\log^2 n\right)$ calls to AE-MonoΔ.

▶ **Theorem 3.** *If* AE-MonoΔ *is in* $T(n)$ *time, then* UnweightedAPSP *is in* $\mathcal{O}\left(T(n) \log^2 n\right)$ *time.*

Proof. We solve UnweightedAPSP in $\log n$ rounds, in the i-th round we compute matrix $D^{\leqslant 2^i}$ of lengths of shortest paths of length up to 2^i (other entries equal to ∞). Each round will consist of a parallel binary search, similar to the one we use in our reduction from MinWitness to AE-MonoΔ (Theorem 2). The algorithm is based on the fact that in unweighted graphs every path can be split roughly in half, i.e. if the distance from u to v equals to k, then there must exist a vertex w such that the distances from u to w and from w to v equal to $\lfloor k/2 \rfloor + \{0, 1\}$.

To start, note that $D^{\leqslant 2^0}$ is a $\{0, 1, \infty\}$-matrix that can be easily obtained from the adjacency matrix of the input graph. Now, assume we already computed $D^{\leqslant 2^i}$ and let us proceed to compute $D^{\leqslant 2^{i+1}}$. To avoid excessive indexing, let A denote $D^{\leqslant 2^i}$, and B denote $D^{\leqslant 2^{i+1}}$. For each entry of the output matrix B we will keep an interval which that entry is guaranteed to lie in. With a single call to AE-MonoΔ we will be able to halve all the intervals.

For $\ell \in \{0, 1, \ldots, i+2\}$, let $B^{(\ell)}$ denote the matrix pointing to 2^ℓ-length intervals in which entries of B lie, that is $B^{(\ell)}_{uv}$ equals to the unique integer such that $B_{uv} \in \left[2^\ell \cdot B^{(\ell)}_{uv}, 2^\ell \cdot (B^{(\ell)}_{uv} + 1) - 1\right]$, or to infinity in case B_{uv} is infinite.

We will iterate over ℓ from $i+2$ down to 0. First, we need to compute $B^{(i+2)}$, whose entries are either zeros or infinities. Recall that we already know the matrix $A = D^{\leqslant 2^i}$. Consider a pair of nodes u and v that are at distance at most 2^{i+1}. There must exist a node w such that $A_{uw} \leqslant 2^i$ and $A_{wv} \leqslant 2^i$, that is, equivalently both A_{uw} and A_{wv} are finite. We obtain the matrix $B^{(i+2)}$ by squaring the $(0,1)$ matrix obtained from A by putting ones at the finite entries and zeros elsewhere. That single matrix multiplication can be easily simulated by a single call to AE-MonoΔ, using just two colors.

Once we have the matrix $B^{(\ell+1)}$ we want to compute $B^{(\ell)}$. For this we first note that if $B^{(\ell+1)}_{uv} = j$ then $B^{(\ell)}_{uv}$ is either $2j$ or $2j+1$. If $B^{(\ell)}_{uv} = 2j$, then there must exist a vertex w

such that

$$A_{uw} \in \big[2^{\ell-1}\cdot(2j), 2^{\ell-1}\cdot(2j+1)\big), \quad \text{and} \quad A_{wv} \in \big[2^{\ell-1}\cdot(2j), 2^{\ell-1}\cdot(2j+1)\big]. \tag{1}$$

Furthermore, if $B^{(\ell)}_{uv} > 2j$, then there is no w such that the above condition holds. This will allow us to distinguish between the $2j$ and $2j+1$ cases by coloring the matrix A based on which range the entries fall in. Note that the ranges in Condition (1) do not overlap with corresponding ranges for different integer values $j' \neq j$. Thus we will be able to use a single call to AE-MonoΔ to check in parallel for all values of $B^{(\ell)}_{uv}$ if they are the smaller even value $2\cdot B^{(\ell+1)}_{uv}$ or the larger odd value $2\cdot B^{(\ell+1)}_{uv}+1$.

We construct an AE-MonoΔ instance with a tripartite graph with the vertex set $U \sqcup V \sqcup W$ where U, V and W are disjoint copies of the original vertex set. The edges between U and V correspond to our desired output. If $B^{(\ell+1)}_{uv} = j$ then we color the edge $(u,v) \in U \times V$ with j. The edges between U and W correspond to the first part of Condition (1), i.e. if $A_{uw} \in \big[2^{\ell-1}\cdot(2j), 2^{\ell-1}\cdot(2j+1)\big)$, then we add the edge (u,w) to $U \times W$ with color j. The edges between W and V correspond to the second part of Condition (1), i.e. if $A_{wv} \in \big[2^{\ell-1}\cdot(2j), 2^{\ell-1}\cdot(2j+1)\big]$, then we add the edge (w,v) to $W \times V$ with color j. Any edge (u,v) in $U \times V$ that is in a monochromatic triangle implies $B^{(\ell)}_{uv} = 2\cdot B^{(\ell+1)}_{uv}$. Conversely, any edge (u,v) that is not a part of any monochromatic triangle implies $B^{(\ell)}_{uv} = 2\cdot B^{(\ell+1)}_{uv}+1$.

We iterate down until $B^{(0)}$, and observe that $B^{(0)} = B$. Thus, with $\mathcal{O}(\log n)$ calls we can compute $B = D^{\leqslant 2^{i+1}}$ from $A = D^{\leqslant 2^i}$. To solve UnweightedAPSP the total number of calls we need to make to AE-MonoΔ is $\mathcal{O}\big(\log^2(n)\big)$. Therefore, if AE-Mono$\Delta$ can be solved in $T(n)$ time then UnweightedAPSP can be solved in $\mathcal{O}\big(T(n)\log^2(n)\big)$ time. ◀

Now we show that AE-MonoΔ is 3SUM-hard. In our proof we use as a black-box the following reduction from 3SUM to AE-SparseΔ.

▶ **Lemma 19** (Kopelowitz, Pettie, Porat [27]). *A single instance of 3SUM of size n can be reduced to a single instance of* AE-SparseΔ *with $\Theta(n\log n)$ vertices and $\Theta\big(n^{3/2}\log n\big)$ edges.*

▶ **Theorem 4.** *If* AE-MonoΔ *is in $\mathcal{O}\big(n^{5/2-\varepsilon}\big)$ time, then 3SUM is in (randomized) $\widetilde{\mathcal{O}}\Big(n^{2-\frac{4}{5}\varepsilon}\Big)$ time.*

Proof. Given an instance of 3SUM of size N, we use the self-reduction (Lemma 18), and reduce it to $\mathcal{O}\big(N^{2/5}\big)$ instances of size $\mathcal{O}\big(N^{4/5}\big)$ each. Then, we reduce each of these instances to an AE-SparseΔ instance with $n = \Theta\big(N^{4/5}\log N\big)$ vertices and $m = \Theta\big(N^{6/5}\log N\big)$ edges, using Lemma 19. Now we will show how to combine these $\mathcal{O}\big(N^{2/5}\big)$ AE-SparseΔ instances to form polylogarithmically many AE-MonoΔ instances, each with $\mathcal{O}\big(N^{4/5}\log N\big)$ vertices, which will finish the proof.

Assume w.l.o.g. that all the created graphs are over the same vertex set $[n]$. If we were lucky enough and the edge sets of the created AE-SparseΔ instances were disjoint, the reduction would be essentially done. Indeed, we could simply union the edge sets to create a single graph, and use colors to track from which graph every edge originates. Solving that one AE-MonoΔ instance would provide answers to all AE-SparseΔ instances. Sadly, the chances of such a favorable collision-free scenario are very slim. The remaining part of the proof shows how to deal with multiple AE-SparseΔ instances containing the same edge.

We randomly permute the vertex sets, for each graph independently. For a fixed $(u,v) \in [n]^2$, the probability that a fixed graph contains the edge (u,v) equals to $p = m/\binom{n}{2} = \mathcal{O}\big((N^{2/5}\log N)^{-1}\big)$. The expected number of (u,v) edges across all graphs is $\mathcal{O}\big(N^{2/5}\cdot p\big) = \mathcal{O}(1/\log N)$. By Chernoff bound, the probability that the number of (u,v) edges exceeds

$c \log n$ is less than $e^{\Theta(c \log n)}$. We take c large enough so that, by union bound, with probability at least $1/2$ no edge appears more than $c \log n$ times across all graphs. For each $(u, v) \in [n]^2$ we arbitrarily number all (u, v) edges with consecutive positive integers from 1 up to at most $c \log n$. We iterate over all triples $(i, j, k) \in [c \log n]^3$. For every triple we create a tripartite graph with the vertex set $V_1 \sqcup V_2 \sqcup V_3$, for $V_1 = V_2 = V_3 = [n]$. We create an edge (u, v) between V_1 and V_2 if there exists an edge (u, v) with number i assigned to it in any of the AE-SparseΔ instances. Note that there can be at most one such instance. We set the color of the newly created edge to the identifier of the instance it originates from. Similarly, we create edges between V_2 and V_3 using edges with number j assigned, and between V_3 and V_1 using number k. That gives us $(c \log n)^3$ instances of AE-MonoΔ. Note that every triangle present in any of the AE-SparseΔ instance corresponds to a single monochromatic in one of the AE-MonoΔ instances, and vice versa. We solve all AE-MonoΔ instances and combine the outputs in order to get the output for all AE-SparseΔ instances, and eventually for the 3SUM instance. ◀

The next two theorems use techniques similar to Theorems 2 and 3 to give reductions to Min-Max.

▶ **Theorem 5** (folklore). *If* Min-Max *is in* $T(n)$ *time, then* MinWitness *is in* $\mathcal{O}(T(n))$ *time.*

Proof. Given two $(0, 1)$ matrices A and B, we construct matrices A' and B' such that

$$A'_{ik} = \begin{cases} k & \text{if } A_{ik} = 1, \\ \infty & \text{if } A_{ik} = 0, \end{cases} \quad \text{and} \quad B'_{kj} = \begin{cases} k & \text{if } B_{kj} = 1, \\ \infty & \text{if } B_{kj} = 0. \end{cases}$$

Observe that the $(\min, \max)$-product of A' and B' equals to the minimum witness product of A and B. ◀

▶ **Theorem 6.** *If* Min-Max *is in* $T(n)$ *time, then* UnweightedAPSP *is in* $\mathcal{O}(T(n) \log n)$ *time.*

Proof. The reduction is similar to the reduction from UnweightedAPSP to AE-MonoΔ (Theorem 3) in that we also have $\log n$ rounds, and in the i-th round we compute matrix $D^{\leqslant 2^i}$ of lengths of shortest paths of length up to 2^i (other entries equal to ∞). The key difference is that, in each round, instead of performing a binary search and issuing $\log n$ calls to AE-MonoΔ, we issue just two calls to Min-Max.

As before, first note that $D^{\leqslant 2^0}$ is a $\{0, 1, \infty\}$-matrix that can be easily obtained from the adjacency matrix of the input graph. Now, assume we already computed $D^{\leqslant 2^i}$ and let us proceed to compute $D^{\leqslant 2^{i+1}}$. Let $\ell = 2^i$. Naturally, $D^{\leqslant 2\ell}$ is the $(\min, +)$-product of $D^{\leqslant \ell}$ with itself, but this sole observation is not enough for our purposes. We will exploit the fact that $D^{\leqslant \ell}$ is not an arbitrary matrix – but a (truncated) matrix of shortest paths in an unweighted graph – in order to compute that specific $(\min, +)$-product using a Min-Max algorithm. Let $A \mathbin{⍉} B$ denote the $(\min, \max)$-product of matrices A and B.

First, we handle even-length paths. We compute $E = 2 \cdot (D^{\leqslant \ell} \mathbin{⍉} D^{\leqslant \ell})$. Note that $D^{\leqslant 2\ell}_{uv} \leqslant E_{uv}$ for all $u, v \in V$, because for any two integers a, b we have $a + b \leqslant 2 \cdot \max(a, b)$. Moreover, if $D^{\leqslant 2\ell}_{uv} = 2k$, then there must exist $w \in V$ such that $D^{\leqslant \ell}_{uw} = D^{\leqslant \ell}_{wv} = k$, and thus $D^{\leqslant \ell}_{uw} + D^{\leqslant \ell}_{wv} = 2 \cdot \max(D^{\leqslant \ell}_{uw}, D^{\leqslant \ell}_{wv})$ and $D^{\leqslant 2\ell}_{uv} = E_{uv}$.

For odd-length paths we proceed in a similar manner, just the formulas become slightly more obscure. We compute $O = 2 \cdot (D^{\leqslant \ell} \mathbin{⍉} (D^{\leqslant \ell} - 1)) + 1$. Note that $D^{\leqslant 2\ell}_{uv} \leqslant O_{uv}$ for all $u, v \in V$, because for any two integers a, b we have $a + b \leqslant 2 \cdot \max(a, b - 1) + 1$. Moreover, if $D^{\leqslant 2\ell}_{uv} = 2k + 1$, then there must exist $w \in V$ such that $D^{\leqslant \ell}_{uw} = k$ and $D^{\leqslant \ell}_{wv} = k + 1$, and thus $D^{\leqslant \ell}_{uw} + D^{\leqslant \ell}_{wv} = 2 \cdot \max(D^{\leqslant \ell}_{uw}, D^{\leqslant \ell}_{wv} - 1) + 1$ and $D^{\leqslant 2\ell}_{uv} = O_{uv}$.

Consequently, we compute $D^{\leqslant 2\ell}_{uv} = \min(E_{uv}, O_{uv})$, for all $u, v \in V$. ◀

4 Reductions for Convolution Problems

In this section we provide two reductions which together show that MonoConvolution is fine-grained equivalent to 3SUM. Recall that the best known algorithms for MonoConvolution require time $n^{3/2-o(1)}$, and the best algorithms for 3SUM require time $n^{2-o(1)}$, so this is an equivalence between problems of different time complexity. At the end of the section we reduce CoinChange to MonoConvolution.

First, let us recall the All-Integers variant of 3SUM, which parallels the All-Edges variants of our graph problems. That variant is easier to work with than the original 3SUM problem for our purposes. Luckily if either variant has a subquadratic algorithm then they both do [44].

▶ **Definition 20** (All-Integers 3SUM). *Given three lists A, B, C of n integers each, output the list of all integers $c \in C$ such that there exist $a \in A$ and $b \in B$ such that $a + b = c$.*

▶ **Lemma 21** (Vassilevska Williams, Williams [44]). *If 3SUM is in $\mathcal{O}(n^{2-\varepsilon})$ time, then All-Integers 3SUM is in $\mathcal{O}(n^{2-\varepsilon/2})$ time.*

An important ingredient of our reduction from 3SUM to AE-MonoΔ (Theorem 7) is the following range reduction for 3SUM.

▶ **Lemma 22** (Baran, Demaine, Pătraşcu, rephrased, see Section 2.1 of [4]). *For every positive integer output size s, there exists a family of hash functions H such that:*

1. *Every hash function $h \in H$ hashes to the range $\{0, 1, \ldots, R-1\}$ for $R = 2^s$.*
2. *For all integers $a, b, c \in \mathbb{Z}$ and all hash functions $h \in H$, if $a + b = c$, then*

$$h(a) + h(b) \equiv h(c) + \{-1, 0, 1\} \mod R.$$

3. *Given an integer c and two lists of n integers A and B such that there are no $a \in A, b \in B$ with $a + b = c$, the probability, over hash functions h drawn uniformly at random from H, that there exist $a \in A, b \in B$ such that $h(a) + h(b) \equiv h(c) + \{-1, 0, 1\} \mod R$ is at most $\mathcal{O}(n^2/R)$.*

We are now ready to show that 3SUM can be solved efficiently with a MonoConvolution algorithm. Our reduction uses the fact that we can re-write a 3SUM instance with n integers in $[-R, R]$ as a convolution of $\mathcal{O}(R)$-length $(0, 1)$ vectors, where a one in the i-th position corresponds to the number i in the original 3SUM instance. We will combine several such instances into one MonoConvolution instance by giving each instance its own number. A one in position i in a convolution instance labelled j will result in the MonoConvolution instance having j in position i.

▶ **Theorem 7.** *If MonoConvolution is in $\mathcal{O}(n^{3/2-\varepsilon})$ time, then 3SUM is in (randomized) $\widetilde{\mathcal{O}}\left(n^{2-\frac{4}{3}\varepsilon}\right)$ time.*

Proof. Given an instance of 3SUM of size n, we reduce it to $\mathcal{O}(n^{2/3})$ instances of size $\mathcal{O}(n^{2/3})$ each, using the self-reduction (Lemma 18). Although for the self-reduction itself it would be sufficient just to solve 3SUM on each of these instances – i.e. decide if there exist a, b, c with $a + b = c$ – we are going to solve the All-Integers 3SUM variant – i.e. decide for each c if there exist a and b with $a + b = c$.

To each created instance we apply a hashing scheme of Lemma 22 in order to reduce the universe size down to $R = n^{4/3}$. This introduces false positives for each element with probability $\mathcal{O}((n^{2/3})^2)/R = \mathcal{O}(1)$. Note that the hashing has one-sided error, i.e. if for some

element c there are no a and b such that $h(a)+h(b) \equiv h(c)+\{-1,0,1\} \mod R$, then with certainty there are no a and b such that $a+b=c$. To mitigate the effect of false positives we create $\mathcal{O}(\log n)$ copies of each instance, each copy with an independently drawn hash function. Note that for every fixed element c, if there are no a, b with $a+b=c$, then the probability that in each of the independent $\mathcal{O}(\log n)$ copies we detect that $h(a)+h(b) \equiv h(c)+\{-1,0,1\} \mod R$ for some $h(a), h(b)$ is $1/\text{poly}(n)$, and we can make the degree of the polynomial arbitrarily large by choosing an appropriate multiplicative constant for the number of copies. Therefore we can use the union bound to argue that with at least $2/3$ probability there are no false positives across all instances and all elements.

Imagine that for some sub-instance A, B, C of size $n^{2/3}$ we obtain the All-Integers 3-SUM results for all $\mathcal{O}(() \lg(n))$ hashed instance copies. That is, we know for every value $h(c)$ in each of the instances if there exists a $h(a)+h(b) \equiv h(c)+\{-1,0,1\}$. Then, we can check for every original $c \in C$ if the answer for $h(c)$ was yes for all $\mathcal{O}(() \lg(n))$ hash functions. If the answer was YES for all hash functions we can conclude with high probability that there is a 3-SUM involving c in A, B, C. If the answer is NO for some hash function then we can conclude there is not a three sum involving c in A, B, C.

Here an important point is that we need to solve *all* of the $\mathcal{O}\big(n^{2/3} \log n\big)$ instances above of All-Integers 3SUM on $n^{2/3}$ integers each over a range $[\mathcal{O}\big(n^{4/3}\big)]$. We will embed solving all instances simultaneously into solving a small number of MonoConvolution instances.

Each of the above $\mathcal{O}\big(n^{2/3} \log n\big)$ instances of All-Integers 3SUM easily reduces to an (OR, AND)-convolution of $(0,1)$ vectors of length $\mathcal{O}\big(n^{4/3}\big)$, each with only $\mathcal{O}\big(n^{2/3}\big)$ nonzero entries, and with only $\mathcal{O}\big(n^{2/3}\big)$ relevant output coordinates one needs to compute. If only we had no collisions – i.e. two instances with the same nonzero input coordinate or the same relevant output coordinate – we could easily combine all the convolution instances into a single instance of MonoConvolution, with $\mathcal{O}\big(n^{2/3} \log n\big)$ different colors/values. However, the collisions are unavoidable. In order to circumvent these collisions, we will add small random shifts, and use a similar analysis as in the 3SUM to AE-MonoΔ reduction of Theorem 4.

Specifically, for each instance we chose a shift s uniformly at random from a range of size $\mathcal{O}\big(n^{4/3}\big)$, we add s to all elements in A, add s too all elements in B, and add $-2s$ to all elements in C. These shifts do not change whether for a given triplet a, b, c the condition $a+b=c$ holds or not. For a fixed value $v \in [R]$ the expected number of instances containing v is $\mathcal{O}(\log n)$: for each particular instance, the probability that one of its numbers is lands at v after the shift is $n^{2/3}/R = 1/n^{2/3}$; then union bounding over all the instances gives an expectation of $\mathcal{O}(\log n)$.

Since the shifts are independent, we can use a Chernoff bound to bound the probability that the number of instances containing v exceeds $c \log n$ by $\leqslant e^{\Theta(c \log n)}$. We take c large enough so that, by union bound, the probability that no value is contained in more than $c \log n$ instances is at least $2/3$.

Then, once again following the example of Theorem 4, we reduce the problem to $(c \log n)^3$ instances of MonoConvolution as follows.

For each value $r \in [R]$, let the instances that contain r in their A sets be $in_A(r)[1], ..., in_A(r)[c \log n]$. Define $in_B(r)[1], ..., in_B(r)[c \log n]$ and $in_C(r)[1], ..., in_C(r)[c \log n]$ analogously.

We now create an instance of MonoConvolution for each choice of $(x, y, z) \in [c \log n]^3$. In instance (x, y, z) we create vectors a, b, c, where for each $r \in [R]$, we set $a[r] = in_A(r)[x]$, $b[r] = in_B(r)[y]$ and $c[r] = in_C(r)[z]$.

Then for any instance i that contains r in A, s in B and t in C, we would have $in_A(r)[x] = in_B(s)[y] = in_C(t)[z] = i$ for some x, y, z and so we will place i in $a[r], b[s]$, and $c[t]$ for that choice of x, y, z. Thus all value collisions will be handled.

◀

This next reduction finishes the equivalence between MonoConvolution and 3SUM. It uses a high-frequency/low-frequency split. For elements that appear at a high frequency we use FFT. For elements of low frequency we make calls to All-Integers 3SUM. Recall that a subquadratic algorithm for 3SUM implies a subquadratic algorithm for All-Integers 3SUM.

▶ **Theorem 8.** *If 3SUM is in $\mathcal{O}(n^{2-\varepsilon})$ time, then* MonoConvolution *is in $\widetilde{\mathcal{O}}(n^{3/2-\varepsilon/4})$ time.*

Proof. For a parameter t to be determined later, consider the t most frequent values. For each of these values use FFT to calculate the standard $(+,\times)$-convolution of two $(0,1)$ vectors formed from vectors a and b by putting ones everywhere that value appears, and zeros everywhere else. Examine, where the outputs of these convolutions are nonzero, in order to determine the part of output to MonoConvolution corresponding to occurrences of the frequent values in vector c. This takes $\widetilde{\mathcal{O}}(tn)$ time in total.

Let n_i denote the number of occurrences of the i-th of the remaining values in all three sequences. Clearly, $\forall_i\ n_i \leqslant 3n/t$, and $\sum_i n_i \leqslant 3n$. For each value v out of those remaining values construct sets of indices at which it appears in vectors a, b, c, i.e. $A = \{j : a_j = v\}$, $B = \{j : b_j = v\}$, $C = \{j : c_j = v\}$, and solve All-Integers 3SUM on these sets. For each element j reported by the All-Integers 3SUM algorithm assign the corresponding output of MonoConvolution $d_j = 1$. By Lemma 21, solving these All-Integers 3SUM instances takes an order of

$$\sum_i n_i^{2-\varepsilon/2} = \sum_i n_i \cdot n_i^{1-\varepsilon/2} \leqslant \sum_i n_i \cdot (3n/t)^{1-\varepsilon/2} \leqslant 3n \cdot (3n/t)^{1-\varepsilon/2}$$

time. The total time is thus $\widetilde{\mathcal{O}}\big(tn + n \cdot (n/t)^{1-\varepsilon/2}\big)$. Optimize by setting $t = n^{1/2-\varepsilon/4}$, and get the desired runtime. ◀

The following theorem connects the CoinChange problem to our network of reduction. The proof uses the same structure and techniques as the reduction from UnweightedAPSP to AE-MonoΔ in Theorem 3.

▶ **Theorem 9.** *If* MonoConvolution *is in $T(n)$ time, then* CoinChange *is in $\mathcal{O}\big(T(n)\log^2 n\big)$ time.*

Proof. Let S denote the array of output values, i.e. $S[v]$ equals to the minimum number of coins that sum to v. Parallel to the proof of Theorem 3, let $S^{\leqslant 2^i}[v]$ be infinity if $S[v] > 2^i$, and otherwise equal to $S[v]$. We solve CoinChange in $\log n$ rounds, in the i-th round we compute $S^{\leqslant 2^i}$.

Note that $S^{\leqslant 2^0}[0] = 0$, and, for $v \geqslant 1$, $S^{\leqslant 2^0}[v] = 1$ if $v \in C$ and $S^{\leqslant 2^0}[v] = \infty$ otherwise. Further note that $S = S^{\leqslant 2^{\log n}}$. We will show how to compute $S^{\leqslant 2^{i+1}}$ given $S^{\leqslant 2^i}$. We will then iterate i from 0 up to $\log n$.

Following the style of Theorem 3, we avoid overparameterizing by setting $A = S^{\leqslant 2^i}$ and $B = S^{\leqslant 2^{i+1}}$. Let $B^{(\ell)}$ be an array pointing to 2^ℓ-length intervals in which entries of B lie, i.e. $B^{(\ell)}[v] = j$ if $B[v] \in [2^\ell j, 2^\ell(j+1)-1]$. We will iterate ℓ from $i+2$ down to 0 to compute B from A.

First we show how to compute $B^{(i+2)}$ from A. If there is a way to sum to v with at most 2^{i+1} coins, then there must be a $u \in [0,n]$ such that both $A[u]$ and $A[v-u]$ are at most 2^i. Conversely, if there is no way to sum to v with at most 2^i coins, then there will be no u that meets the above criteria. Therefore we create a $(0,1)$ vector a with $a_v = 1$ if and only if $A[v]$ is finite. Then, we compute the $(+,\times)$-convolution of a with itself, in near-linear time using

FFT. We set $B[v] = 0$ where the convolution output is non-zero and $B[v] = \infty$ everywhere else.

Now we show how to compute $B^{(\ell)}$ from A and $B^{(\ell+1)}$. Note that if $B^{(\ell+1)}[v] = j$ then $B^{(\ell)}[v] \in \{2j, 2j+1\}$. Next, note that if $B^{(\ell)}[v] = 2j$, then there must exist an integer $u \in [0, n]$ such that

$$A[u] \in \left[2^{\ell-1} \cdot (2j), 2^{\ell-1} \cdot (2j+1)\right), \quad \text{and} \quad A[v-u] \in \left[2^{\ell-1} \cdot (2j), 2^{\ell-1} \cdot (2j+1)\right]. \quad (2)$$

Furthermore, if $B^{(\ell)}[v] > 2j$, then there is no u such that the above condition holds. This will allow us to distinguish between the $2j$ and $2j+1$ cases. Note that the ranges in Condition (2) do not overlap with corresponding ranges for different integer values $j' \neq j$. Thus, we will be able to use a single call to MonoConvolution to check in parallel for all values v if $B^{(\ell)}[v]$ is the smaller even value $2B^{(\ell+1)}[v]$ or the larger odd value $2B^{(\ell+1)}[v] + 1$.

We construct a MonoConvolution instance with three input vectors a, b, c. The first input vector corresponds to the first part of Condition (2), i.e. if $A[v] \in \left[2^{\ell-1} \cdot (2j), 2^{\ell-1} \cdot (2j+1)\right)$, then $a_v = j$. Any entries a_v unset by this condition are given the special value $a_v = -1$. The second vector corresponds to the second part of Condition (2), i.e. if $A[v] \in \left[2^{\ell-1} \cdot (2j), 2^{\ell-1} \cdot (2j+1)\right]$, then $b_v = j$. Similarly, any entries b_v unset by this condition are given the special value $b_v = -1$. The last vector corresponds to our desired output, i.e. $c_v = B^{(\ell+1)}[v]$. Let d denote the vector output by this MonoConvolution call. Now, if $d_v = 1$ then $B^{(\ell)}[v] = 2B^{(\ell+1)}[v]$, else $B^{(\ell)}[v] = 2B^{(\ell+1)}[v] + 1$.

We iterate down until $B^{(0)}$, and observe that $B^{(0)} = B$. Thus, with $\mathcal{O}(\log n)$ calls to MonoConvolution we can compute $B = S^{\leqslant 2^{i+1}}$ from $A = S^{\leqslant 2^i}$. To solve CoinChange the total number of MonoConvolution calls is $\mathcal{O}\left(\log^2 n\right)$. Therefore if MonoConvolution can be solved in $T(n)$ time, then CoinChange can be solved in $\mathcal{O}\left(T(n) \log^2 n\right)$ time. ◀

References

1. Amir Abboud, Fabrizio Grandoni, and Virginia Vassilevska Williams. Subcubic equivalences between graph centrality problems, APSP and diameter. In *Proceedings of the Twenty-Sixth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2015, San Diego, CA, USA, January 4-6, 2015*, pages 1681–1697, 2015. doi:10.1137/1.9781611973730.112.
2. Noga Alon, Zvi Galil, and Oded Margalit. On the exponent of the all pairs shortest path problem. *Journal of Computer and System Sciences*, 54(2):255–262, 1997. doi:10.1006/jcss.1997.1388.
3. Noga Alon, Raphael Yuster, and Uri Zwick. Finding and counting given length cycles. *Algorithmica*, 17(3):209–223, Mar 1997. doi:10.1007/BF02523189.
4. Ilya Baran, Erik D. Demaine, and Mihai Pătraşcu. Subquadratic algorithms for 3SUM. *Algorithmica*, 50(4):584–596, Apr 2008. doi:10.1007/s00453-007-9036-3.
5. Hodaya Barr, Tsvi Kopelowitz, Ely Porat, and Liam Roditty. $\{-1, 0, 1\}$-APSP and (min,max)-product problems, 2019. arXiv:1911.06132.
6. Michael A. Bender, Giridhar Pemmasani, Steven Skiena, and Pavel Sumazin. Finding least common ancestors in directed acyclic graphs. In *Proceedings of the Twelfth Annual Symposium on Discrete Algorithms, January 7-9, 2001, Washington, DC, USA.*, pages 845–854, 2001. URL: http://dl.acm.org/citation.cfm?id=365411.365795.
7. David Bremner, Timothy M. Chan, Erik D. Demaine, Jeff Erickson, Ferran Hurtado, John Iacono, Stefan Langerman, and Perouz Taslakian. Necklaces, convolutions, and X + Y. In *Proceedings of the 14th Conference on Annual European Symposium - Volume 14*, ESA'06, pages 160–171, London, UK, UK, 2006. Springer-Verlag. doi:10.1007/11841036_17.
8. Karl Bringmann, Allan Grønlund, and Kasper Green Larsen. A dichotomy for regular expression membership testing. In *58th IEEE Annual Symposium on Foundations of Computer*

Science, FOCS 2017, Berkeley, CA, USA, October 15-17, 2017, pages 307–318, 2017. doi:10.1109/FOCS.2017.36.

9 Karl Bringmann and Tomasz Kociumaka. Personal communication, 2019.

10 Karl Bringmann, Marvin Künnemann, and Karol Wegrzycki. Approximating APSP without scaling: Equivalence of approximate min-plus and exact min-max. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2019, pages 943–954, New York, NY, USA, 2019. ACM. doi:10.1145/3313276.3316373.

11 Timothy M. Chan. More logarithmic-factor speedups for 3SUM, (median, +)-convolution, and some geometric 3SUM-hard problems. In *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2018, New Orleans, LA, USA, January 7-10, 2018*, pages 881–897, 2018. doi:10.1137/1.9781611975031.57.

12 Lijie Chen and Ryan Williams. An equivalence class for orthogonal vectors. In *Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2019, San Diego, California, USA, January 6-9, 2019*, pages 21–40, 2019. doi:10.1137/1.9781611975482.2.

13 Marek Cygan, Holger Dell, Daniel Lokshtanov, Dániel Marx, Jesper Nederlof, Yoshio Okamoto, Ramamohan Paturi, Saket Saurabh, and Magnus Wahlström. On problems as hard as CNF-SAT. *ACM Transactions on Algorithms*, 12(3):41:1–41:24, 2016. doi:10.1145/2925416.

14 Marek Cygan, Marcin Mucha, Karol Wegrzycki, and Michal Wlodarczyk. On problems equivalent to (min,+)-convolution. *ACM Transactions on Algorithms*, 15(1):14:1–14:25, January 2019. doi:10.1145/3293465.

15 Artur Czumaj, Mirosław Kowaluk, and Andrzej Lingas. Faster algorithms for finding lowest common ancestors in directed acyclic graphs. *Theoretical Computer Science*, 380(1):37 – 46, 2007. Automata, Languages and Programming. doi:https://doi.org/10.1016/j.tcs.2007.02.053.

16 Artur Czumaj and Andrzej Lingas. Finding a heaviest triangle is not harder than matrix multiplication. In *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '07, pages 986–994, Philadelphia, PA, USA, 2007. Society for Industrial and Applied Mathematics. URL: http://dl.acm.org/citation.cfm?id=1283383.1283489.

17 Martin Dietzfelbinger. Universal hashing and k-wise independent random variables via integer arithmetic without primes. In *STACS 96*, pages 567–580, Berlin, Heidelberg, 1996. Springer Berlin Heidelberg. doi:10.1007/3-540-60922-9_46.

18 Ran Duan, Ce Jin, and Hongxun Wu. Faster algorithms for all pairs non-decreasing paths problem. In *46th International Colloquium on Automata, Languages, and Programming, ICALP 2019, July 9-12, 2019, Patras, Greece.*, pages 48:1–48:13, 2019. doi:10.4230/LIPIcs.ICALP.2019.48.

19 Ran Duan and Seth Pettie. Fast algorithms for (max, min)-matrix multiplication and bottleneck shortest paths. In *Proceedings of the Twentieth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2009, New York, NY, USA, January 4-6, 2009*, pages 384–391, 2009. doi:10.1137/1.9781611973068.43.

20 Lech Duraj, Krzysztof Kleiner, Adam Polak, and Virginia Vassilevska Williams. Equivalences between triangle and range query problems. *arXiv e-prints*, page arXiv:1908.11819, Aug 2019. arXiv:1908.11819.

21 Michael J Fischer and Albert R Meyer. Boolean matrix multiplication and transitive closure. In *12th Annual Symposium on Switching and Automata Theory, SWAT 1971*, pages 129–131. IEEE, 1971. doi:10.1109/SWAT.1971.4.

22 Anka Gajentaan and Mark H. Overmars. On a class of $O(n^2)$ problems in computational geometry. *Computational Geometry*, 5:165–185, 1995. doi:https://doi.org/10.1016/0925-7721(95)00022-2.

23 Francois Le Gall and Florent Urrutia. Improved rectangular matrix multiplication using powers of the Coppersmith-Winograd tensor. In *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2018, New Orleans, LA, USA, January 7-10, 2018*, pages 1029–1046, 2018. doi:10.1137/1.9781611975031.67.

24 Omer Gold and Micha Sharir. Dominance Product and High-Dimensional Closest Pair under L_∞. In Yoshio Okamoto and Takeshi Tokuyama, editors, *28th International Symposium on Algorithms and Computation (ISAAC 2017)*, volume 92 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 39:1–39:12, Dagstuhl, Germany, 2017. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik. doi:10.4230/LIPIcs.ISAAC.2017.39.

25 Allan Grønlund and Seth Pettie. Threesomes, degenerates, and love triangles. *Journal of the ACM*, 65(4):22:1–22:25, April 2018. doi:10.1145/3185378.

26 MohammadTaghi Hajiaghayi, Silvio Lattanzi, Saeed Seddighin, and Cliff Stein. MapReduce meets fine-grained complexity: Mapreduce algorithms for APSP, matrix multiplication, 3-SUM, and beyond, 2019. arXiv:1905.01748.

27 Tsvi Kopelowitz, Seth Pettie, and Ely Porat. Higher lower bounds from the 3SUM conjecture. In *Proceedings of the Twenty-seventh Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '16, pages 1272–1287, Philadelphia, PA, USA, 2016. Society for Industrial and Applied Mathematics. doi:10.1137/1.9781611974331.ch89.

28 Marvin Künnemann, Ramamohan Paturi, and Stefan Schneider. On the Fine-Grained Complexity of One-Dimensional Dynamic Programming. In Ioannis Chatzigiannakis, Piotr Indyk, Fabian Kuhn, and Anca Muscholl, editors, *44th International Colloquium on Automata, Languages, and Programming (ICALP 2017)*, volume 80 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 21:1–21:15, Dagstuhl, Germany, 2017. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik. doi:10.4230/LIPIcs.ICALP.2017.21.

29 Karim Labib, Przemyslaw Uznanski, and Daniel Wolleb-Graf. Hamming Distance Completeness. In Nadia Pisanti and Solon P. Pissis, editors, *30th Annual Symposium on Combinatorial Pattern Matching (CPM 2019)*, volume 128 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 14:1–14:17, Dagstuhl, Germany, 2019. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik. doi:10.4230/LIPIcs.CPM.2019.14.

30 François Le Gall. Powers of tensors and fast matrix multiplication. In *Proceedings of the 39th International Symposium on Symbolic and Algebraic Computation*, ISSAC '14, pages 296–303, New York, NY, USA, 2014. ACM. doi:10.1145/2608628.2608664.

31 Andrea Lincoln, Virginia Vassilevska Williams, Joshua R. Wang, and R. Ryan Williams. Deterministic Time-Space Trade-Offs for k-SUM. In *43rd International Colloquium on Automata, Languages, and Programming (ICALP 2016)*, volume 55 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 58:1–58:14, Dagstuhl, Germany, 2016. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik. doi:10.4230/LIPIcs.ICALP.2016.58.

32 Jiří Matoušek. Computing dominances in E^n. *Information Processing Letters*, 38(5):277–278, 1991. doi:10.1016/0020-0190(91)90071-O.

33 Mihai Patrascu. Towards polynomial lower bounds for dynamic problems. In *Proceedings of the Forty-second ACM Symposium on Theory of Computing*, STOC '10, pages 603–610, New York, NY, USA, 2010. ACM. doi:10.1145/1806689.1806772.

34 Raimund Seidel. On the all-pairs-shortest-path problem in unweighted undirected graphs. *Journal of Computer and System Sciences*, 51(3):400–403, 1995. doi:https://doi.org/10.1006/jcss.1995.1078.

35 Asaf Shapira, Raphael Yuster, and Uri Zwick. All-pairs bottleneck paths in vertex weighted graphs. *Algorithmica*, 59(4):621–633, 2011. doi:10.1007/s00453-009-9328-x.

36 Volker Strassen. Gaussian elimination is not optimal. *Numerische mathematik*, 13(4):354–356, 1969. doi:10.1007/BF02165411.

37 Virginia Vassilevska. Nondecreasing paths in a weighted graph or: how to optimally read a train schedule. In *Proceedings of the Nineteenth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2008, San Francisco, California, USA, January 20-22, 2008*, pages 465–472, 2008. URL: http://dl.acm.org/citation.cfm?id=1347082.1347133.

38 Virginia Vassilevska, Ryan Williams, and Raphael Yuster. All-pairs bottleneck paths for general graphs in truly sub-cubic time. In *Proceedings of the 39th Annual ACM Symposium*

on Theory of Computing, San Diego, California, USA, June 11-13, 2007, pages 585–589, 2007. doi:10.1145/1250790.1250876.

39 Virginia Vassilevska, Ryan Williams, and Raphael Yuster. All pairs bottleneck paths and max-min matrix products in truly subcubic time. *Theory of Computing*, 5(1):173–189, 2009. doi:10.4086/toc.2009.v005a009.

40 Virginia Vassilevska, Ryan Williams, and Raphael Yuster. Finding heaviest H-subgraphs in real weighted graphs, with applications. *ACM Transactions on Algorithms*, 6(3):44:1–44:23, July 2010. doi:10.1145/1798596.1798597.

41 Virginia Vassilevska Williams. Nondecreasing paths in a weighted graph or: How to optimally read a train schedule. *ACM Transactions on Algorithms*, 6(4):70:1–70:24, 2010. doi:10.1145/1824777.1824790.

42 Virginia Vassilevska Williams. Multiplying matrices faster than Coppersmith-Winograd. In *Proceedings of the Forty-fourth Annual ACM Symposium on Theory of Computing*, STOC '12, pages 887–898, 2012. doi:10.1145/2213977.2214056.

43 Virginia Vassilevska Williams. On some fine-grained questions in algorithms and complexity. In *Proceedings of the International Congress of Mathematicians (ICM 2018)*, pages 3447–3487, 2018. doi:10.1142/9789813272880_0188.

44 Virginia Vassilevska Williams and R. Ryan Williams. Subcubic equivalences between path, matrix, and triangle problems. *Journal of the ACM*, 65(5):27:1–27:38, August 2018. doi:10.1145/3186893.

45 Virginia Vassilevska Williams and Ryan Williams. Finding, minimizing, and counting weighted subgraphs. *SIAM Journal on Computing*, 42(3):831–854, 2013. doi:10.1137/09076619X.

46 Ryan Williams. Faster all-pairs shortest paths via circuit complexity. In *Proceedings of the Forty-sixth Annual ACM Symposium on Theory of Computing*, STOC '14, pages 664–673, New York, NY, USA, 2014. ACM. doi:10.1145/2591796.2591811.

47 Virginia V. Williams. Problem Set 2 in Stanford's class CS367, Oct. 15, 2015. http://theory.stanford.edu/~virgi/cs367/hw2.pdf, 2015.

48 J. W. Wright. The change-making problem. *Journal of the ACM*, 22(1):125–128, January 1975. doi:10.1145/321864.321874.

49 Raphael Yuster. Efficient algorithms on sets of permutations, dominance, and real-weighted APSP. In *Proceedings of the Twentieth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2009, New York, NY, USA, January 4-6, 2009*, pages 950–957, 2009. doi:10.1137/1.9781611973068.103.

50 Uri Zwick. All pairs shortest paths using bridging sets and rectangular matrix multiplication. *Journal of the ACM*, 49(3):289–317, May 2002. doi:10.1145/567112.567114.