Graph Square Roots of Small Distance from Degree One Graphs*

Petr A. Golovach[1], Paloma T. Lima[2], and Charis Papadopoulos[3]

[1]Department of Informatics, University of Bergen, Norway, petr.golovach@uib.no
[2]Department of Computer Science, IT University of Copenhagen, Denmark, palt@itu.dk
[3]Department of Mathematics, University of Ioannina, Greece, charis@uoi.gr

Abstract

Given a graph class $\mathcal{H}$, the task of the $\mathcal{H}$-Square Root problem is to decide whether an input graph G has a square root H from $\mathcal{H}$. We are interested in the parameterized complexity of the problem for classes $\mathcal{H}$ that are composed by the graphs at vertex deletion distance at most k from graphs of maximum degree at most one, that is, we are looking for a square root H such that there is a modulator S of size k such that $H - S$ is the disjoint union of isolated vertices and disjoint edges. We show that different variants of the problems with constraints on the number of isolated vertices and edges in $H - S$ are FPT when parameterized by k by demonstrating algorithms with running time $2^{2^{\mathcal{O}(k)}} \cdot n^5$. We further show that the running time of our algorithms is asymptotically optimal and it is unlikely that the double-exponential dependence on k could be avoided. In particular, we prove that the VC-k Root problem, that asks whether an input graph has a square root with vertex cover of size at most k, cannot be solved in time $2^{2^{o(k)}} \cdot n^{\mathcal{O}(1)}$ unless the Exponential Time Hypothesis fails. Moreover, we point out that VC-k Root parameterized by k does not admit a subexponential kernel unless $\mathsf{P} = \mathsf{NP}$.

1 Introduction

Squares of graphs and square roots constitute widely studied concepts in graph theory, both from a structural perspective as well as from an algorithmic point of view. A graph G is the *square* of a graph H if G can be obtained from H by the addition of an edge between any two vertices of H that are at distance two. In this case, the graph H is called a *square root* of G. It is interesting to notice that there are graphs that admit different square roots, graphs that have a unique square root and graphs that do not have a square root at all. In 1994, Motwani and Sudan [27] proved that the problem of determining if a given graph G has a square root is NP-complete. This problem is known as the Square Root problem.

The intractability of Square Root has been attacked in two different ways. The first one is by imposing some restrictions on the input graph G. In this vein, the Square Root problem has been studied in the setting in which G belongs to a specific class of graphs [4, 12, 11, 21, 26, 25, 28].

Another way of coping with the hardness of the Square Root problem is by imposing some additional structure on the square root H. That is, given the input graph G, the task is to determine whether G has a square root H that belongs to a specific graph class $\mathcal{H}$. This setting is known as the $\mathcal{H}$-Square Root problem and it is the focus of this work. The $\mathcal{H}$-Square Root problem has been shown to be polynomial-time solvable for specific graph classes $\mathcal{H}$

*A preliminary version of this paper appeared in the Proceedings of LATIN 2020 [13]. This work received support from the Research Council of Norway via the project "MULTIVAL".

[18, 21, 22, 19, 20]. To name a few among others, the problem is solved in polynomial time when $\mathcal{H}$ is the class of trees [24], bipartite graphs [17], cactus graphs [12], and, more recently, when $\mathcal{H}$ is the class of cactus block graphs [6], outerplanar graphs [10], and graphs of pathwidth at most 2 [10]. It is interesting to notice that the fact that $\mathcal{H}$-Square Root can be efficiently (say, polynomially) solved for some class $\mathcal{H}$ does not automatically imply that $\mathcal{H}'$-Square Root is efficiently solvable for every subclass $\mathcal{H}'$ of $\mathcal{H}$. On the negative side, $\mathcal{H}$-Square Root remains NP-complete when $\mathcal{H}$ is the class of graphs of girth at least 5 [7], graphs of girth at least 4 [8], split graphs [18], and chordal graphs [18]. The fact that all known NP-hardness constructions involve dense graphs [7, 8, 18, 27] and dense square roots, raised the question of whether $\mathcal{H}$-Square Root is polynomial-time solvable for every sparse graph class $\mathcal{H}$.

We consider this question from the Parameterized Complexity viewpoint for structural parameterizations of $\mathcal{H}$ (we refer to the book of Cygan et al. [5] for an introduction to the field). More precisely, we are interested in graph classes $\mathcal{H}$ that are at *small distance* from a (sparse) graph class for which $\mathcal{H}$-Square Root can be solved in polynomial time. Within this scope, the distance is usually measured either by the number of edge deletions, edge additions or vertex deletions. This approach for the problem was first applied by Cochefert et al. in [3], who considered $\mathcal{H}$-Square Root, where $\mathcal{H}$ is the class of graphs that have a feedback edge set of size at most k, that is, for graphs that can be made forests by at most k edge deletions. They proved that $\mathcal{H}$-Square Root admits a compression to a special variant of the problem with $\mathcal{O}(k^2)$ vertices, implying that the problem can be solved in $2^{\mathcal{O}(k^4)} + \mathcal{O}(n^4 m)$ time, i.e., is fixed-parameter tractable (FPT) when parameterized by k. Herein, we study whether the same complexity behavior occurs if we measure the distance by the number of *vertex deletions* instead of edge deletions.

Towards such an approach, the most natural consideration for $\mathcal{H}$-Square Root is to ask for a square root of feedback *vertex* set of size at most k. The approach used by Cochefert et al. [3] fails if $\mathcal{H}$ is the class of graphs that can be made forests by at most k vertex deletions and the question of the parameterized complexity of our problem for this case is open. In this context, we consider herein the $\mathcal{H}$-Square Root problem when $\mathcal{H}$ is the class of graphs of bounded vertex deletion distance to a disjoint union of isolated vertices and edges. Our main result is that the problem is FPT when parameterized by the vertex deletion distance. Surprisingly, however, we conclude a notable difference on the running time compared to the edge deletion case even on such a relaxed variation: a double-exponential dependency on the vertex deletion distance is highly unavoidable. Therefore, despite the fact that both problems are FPT, the vertex deletion distance parameterization for the $\mathcal{H}$-Square Root problem requires substantial effort. More formally, we are interested in the following problem.

Distance-k-to-$(pK_1 + qK_2)$ Square Root

Input:	A graph G and nonnegative integers p, q, k such that $p + 2q + k = \lvert V(G)\rvert$.
Task:	Decide whether there is a square root H of G such that $H - S$ is a graph isomorphic to $pK_1 + qK_2$, for a set S on k vertices.

Note that when $q = 0$, the problem asks whether G has a square root with a vertex cover of size (at most) k and we refer to the problem as VC-k Root. If $p = 0$, we obtain Distance-k-to-Matching Square Root. Observe also that, given an algorithm solving Distance-k-to-$(pK_1 + qK_2)$ Square Root, then by testing all possible values of p and q such that $p + 2q = \lvert V(G)\rvert - k$, we can solve the Distance-k-to-Degree-One Square Root problem, whose task is to decide whether there is a square root H such that the maximum degree of $H - S$ is at most one for a set S on k vertices. Note that a set of vertices X inducing a graph of maximum degree one is known as a *dissociation* set and the maximum size of a dissociation set is called the *dissociation* number (see, e.g., [30]). Thus, the task of Distance-k-to-Degree-One Square Root is to find a square root H with the dissociation number at least $\lvert V(G)\rvert - k$.

We show that Distance-k-to-(pK_1+qK_2) Square Root can be solved in $2^{2^{\mathcal{O}(k)}} \cdot n^5$ time, that is, the problem is FPT when parameterized by k, the size of the deletion set. We complement this result by showing that the running time of our algorithm is asymptotically optimal in the sense that VC-k Root, i.e., the special case of Distance-k-to-(pK_1+qK_2) Square Root when $q=0$, cannot be solved in $2^{2^{o(k)}} \cdot n^{\mathcal{O}(1)}$ time unless *Exponential Time Hypothesis* (*ETH*) of Impagliazzo, Paturi and Zane [14, 15] fails (see also [5] for an introduction to the algorithmic lower bounds based on ETH). We also prove that VC-k Root does not admit a kernel of size subexponential in k unless $\mathsf{P}=\mathsf{NP}$.

Motivated by the above results, we further investigate the complexity of the $\mathcal{H}$-Square Root problem when $\mathcal{H}$ is the class of graphs of bounded deletion distance to a specific graph class. We show that the problem of testing whether a given graph has a square root of bounded deletion distance to a clique is also FPT parameterized by the size of the deletion set.

2 Preliminaries

Graphs. All graphs considered here are finite undirected graphs without loops and multiple edges. We refer to the textbook by Bondy and Murty [1] for any undefined graph terminology. We denote the vertex set of G by $V(G)$ and the edge set by $E(G)$. We use n to denote the number of vertices of a graph and use m for the number of edges (if this does not create confusion). Given $x \in V(G)$, we denote by $N_G(x)$ the neighborhood of x. The closed neighborhood of x, denoted by $N_G[x]$, is defined as $N_G(x) \cup \{x\}$. For a set $X \subset V(G)$, $N_G(X)$ denotes the set of vertices in $V(G) \setminus X$ that have at least one neighbor in X. Analogously, $N_G[X] = N_G(X) \cup X$. The *distance* between a pair of vertices $u, v \in V(G)$ is the number of edges of a shortest path between them in G. We denote by $N_G^2(u)$ the set of vertices of G that are at distance *exactly* two from u, and $N_G^2[u]$ is the set of vertices at distance at most two from u. Given $S \subseteq V(G)$, we denote by $G-S$ the graph obtained from G by the removal of the vertices of S. If $S=\{u\}$, we also write $G-u$. The *subgraph induced by* S is denoted by $G[S]$, and has S as its vertex set and $\{uv \mid u, v \in S \text{ and } uv \in E(G)\}$ as its edge set. A *clique* is a set $K \subseteq V(G)$ such that $G[K]$ is a complete graph. An *independent set* is a set $I \subseteq V(G)$ such that $G[I]$ has no edges. A *vertex cover* of G is a set $S \subseteq V(G)$ such that $V(G) \setminus S$ is an independent set. A graph is *bipartite* if its vertex set can be partitioned into two independent sets, say A and B, and is *complete bipartite* if it is bipartite and every vertex of A is adjacent to every vertex of B. A *biclique* in a graph G is a set $B \subseteq V(G)$ such that $G[B]$ is a complete bipartite graph. A *matching* in G is a set of edges having no common endpoint. We denote by K_r the complete graph on r vertices. Given two graphs G and G', we denote by $G+G'$ the disjoint union of them. For a positive integer p, pG denotes the disjoint union of p copies of G.

The *square* of a graph H is the graph $G=H^2$ such that $V(G)=V(H)$ and every two distinct vertices u and v are adjacent in G if and only if they are at distance at most two in H. If $G=H^2$, then H is a *square root* of G.

Two vertices u, v are said to be *true twins* if $N_G[u]=N_G[v]$. A *true twin class* of G is a maximal set of vertices that are pairwise true twins. Note that the set of true twin classes of G constitutes a partition of $V(G)$. Let $\mathcal{T}=\{T_1,\ldots,T_r\}$ be the partition of $V(G)$ into true twin classes. We define the *prime-twin* graph $\mathcal{G}$ of G as the graph with the vertex set $\mathcal{T}$ such that two distinct vertices T_i and T_j of $\mathcal{G}$ are adjacent if and only if $uv \in E(G)$ for $u \in T_i$ and $v \in T_j$.

Parameterized Complexity. We refer to the recent book of [5] for an introduction to Parameterized Complexity. Here we only state some basic definitions that are crucial for understanding. In a *parameterized problem*, each instance is supplied with an integer *parameter* k, that is, each instance can be written as a pair (I, k). A parameterized problem is said to be *fixed-parameter* *tractable* (FPT) if it can be solved in time $f(k) \cdot |I|^{\mathcal{O}(1)}$ for some computable function f. A *kernelization* for a parameterized problem is a polynomial time algorithm that maps each instance (I, k) of a parameterized problem to an instance (I', k') of the same problem such that (i) (I, k) is a YES-instance if and only if (I', k') is a YES-instance, and (ii) $|I'| + k'$ is bounded by $f(k)$ for some computable function f. The output (I', k') is called a *kernel*. The function f is said to be the *size* of the kernel.

Integer Programming. We will use integer linear programming as a subroutine in the proof of our main result. In particular, we translate part of our problem as an instance of the following problem.

p-VARIABLE INTEGER LINEAR PROGRAMMING FEASIBILITY

Input: An $m \times p$ matrix A over $\mathbb{Z}$ and a vector $b \in \mathbb{Z}^m$.

Task: Decide whether there is a vector $x \in \mathbb{Z}^p$ such that $Ax \leq b$.

Lenstra [23] and Kannan [16] showed that the above problem is FPT parameterized by p, while Frank and Tardos [9] showed that this algorithm can be made to run also in polynomial space. We will make use of these results, that we formally state next.

Theorem 1 ([9, 16, 23])**.** *p-VARIABLE INTEGER LINEAR PROGRAMMING FEASIBILITY can be solved using $\mathcal{O}(p^{2.5p+o(p)} \cdot L)$ arithmetic operations and space polynomial in L, where L is the number of bits in the input.*

3 Distance-k-to-$(pK_1 + qK_2)$ Square Root

In this section we give an FPT algorithm for the DISTANCE-k-TO-$(pK_1 + qK_2)$ SQUARE ROOT problem, parameterized by k. In the remainder of this section, we use (G, p, q, k) to denote an instance of the problem. Suppose that (G, p, q, k) is a YES-instance and H is a square root of G such that there is $S \subseteq V(G)$ of size k and $H - S$ is isomorphic to $pK_1 + qK_2$. We say that S is a *modulator*, the p vertices of $H - S$ that belong to pK_1 are called *S-isolated* vertices and the q edges that belong to qK_2 are called *S-matching* edges. Slightly abusing notation, we also use these notions when H is not necessarily a square root of G but any graph such that $H - S$ has maximum degree one.

3.1 Structural lemmas

We start by defining the following two equivalence relations on the set of ordered pairs of vertices of G. Two pairs of adjacent vertices (x, y) and (z, w) are called *matched twins*, denoted by $(x, y) \sim_{\text{mt}} (z, w)$, if the following conditions hold:

- $N_G[x] \setminus \{y\} = N_G[z] \setminus \{w\}$, and
- $N_G[y] \setminus \{x\} = N_G[w] \setminus \{z\}$.

A pair of vertices (x, y) is called *comparable* if $N_G[x] \subseteq N_G[y]$. Two comparable pairs of vertices (x, y) and (z, w) are *nested twins*, denoted by $(x, y) \sim_{\text{nt}} (z, w)$, if the following conditions hold:

- $N_G(x) \setminus \{y\} = N_G(z) \setminus \{w\}$, and
- $N_G[y] \setminus \{x\} = N_G[w] \setminus \{z\}$.

Note that, even though similar, the definitions of these two relations differ in two important points. While $\sim_{\text{mt}}$ is defined in terms of comparisons of *closed* neighborhoods of the vertices of pairs, the relation $\sim_{\text{nt}}$ is defined by comparisons of both *open* and *closed* neighborhoods.

Moreover, while $\sim_{\text{mt}}$ is defined over all pairs of vertices, $\sim_{\text{nt}}$ is defined only over comparable pairs of vertices. Next we state and prove properties about matched and nested twins that will be useful for us.

Lemma 1. *Let (x, y) and (z, w) be two distinct pairs of adjacent vertices (resp. comparable pairs) of G that are matched twins (resp. nested twins). Then, the following holds:*

(i) $\{x, y\} \cap \{z, w\} = \emptyset$,
(ii) $xw, zy \notin E(G)$,
(iii) $yw \in E(G)$,
(iv) *if* $(x, y) \sim_{mt} (z, w)$ *then* $xz \in E(G)$,
(v) *if* $(x, y) \sim_{nt} (z, w)$ *then* $xz \notin E(G)$,
(vi) $G - \{x, y\}$ *and* $G - \{z, w\}$ *are isomorphic.*

Proof. For (i), we show that the end-vertices of both pairs are distinct. It is not difficult to see that $(x, y) \nsim_{\text{mt}} (y, x)$ and $(x, y) \nsim_{\text{nt}} (y, x)$, since $x \in N_G[x] \setminus \{y\}$ and $x \notin N_G[y] \setminus \{x\}$. Assume, for the sake of contradiction, that the two pairs share one end-vertex.

- First, we show (i) for $\sim_{\text{mt}}$. Let $(x, y) \sim_{\text{mt}} (z, w)$. Suppose that $y = w$. Then $x \notin N_G[y] \setminus \{x\}$ but $x \in N_G[w] \setminus \{z\}$, that is, $N_G[y] \setminus \{x\} \neq N_G[w] \setminus \{z\}$ contradicting $(x, y) \sim_{\text{mt}} (z, w)$. Assume that $y = z$. Then $z \in N_G[y] \setminus \{w\}$ but $z = y \notin N_G[x] \setminus \{y\}$; a contradiction. The cases $x = z$ and $x = w$ are completely symmetric to the cases considered above.
- Now we prove (i) for $\sim_{\text{nt}}$. Let $(x, y) \sim_{\text{nt}} (z, w)$. Suppose that $y = w$. Then $x \notin N_G[y] \setminus \{x\}$ but $x \in N_G[w] \setminus \{z\}$, that is, $N_G[y] \setminus \{x\} \neq N_G[w] \setminus \{z\}$; a contradiction to $(x, y) \sim_{\text{nt}} (z, w)$. Let $x = z$. Then $y \notin N_G(x) \setminus \{y\}$ but $y \in N_G[z] \setminus \{w\}$, and we get that $N_G(x) \setminus \{y\} \neq N_G(z) \setminus \{w\}$, leading again to a contradiction. Assume that $y = z$. Then $y = z \notin N_G[w] \setminus \{z\}$ but $y \in N_G[y] \setminus \{x\}$ and we again obtain a contradiction. The case $x = w$ is symmetric.

This completes the proof of (i). To show the remaining claims, observe that $N_G[y] \setminus \{x\} = N_G[w] \setminus \{z\}$ holds in both relations.

For (ii), note that if $xw \in E(G)$, then $x \in N_G[w] \setminus \{z\}$ but $x \notin N_G[y] \setminus \{x\}$, a contradiction. So $xw \notin E(G)$. The same follows by a symmetric argument for the edge yz.

For (iii), note that if $yw \notin E(G)$, then $w \notin N_G[y] \setminus \{x\}$, but $w \in N_G[w] \setminus \{z\}$, a contradiction.

To show (iv), observe that if $xz \notin E(G)$, then $x \in N_G[x] \setminus \{y\}$, while $x \notin N_G[z] \setminus \{w\}$, a contradiction.

For (v), if $xz \in E(G)$, then $z \in N_G(x) \setminus \{y\}$, but $z \notin N_G(z) \setminus \{w\}$, a contradiction.

To see (vi), notice that $\{x, y\} \cap \{z, w\} = \emptyset$ by (i). Consider $\alpha\colon V(G) \to V(G)$ such that $\alpha(x) = z$, $\alpha(y) = w$, $\alpha(z) = x$, $\alpha(w) = y$ and $\alpha(v) = v$ for $v \in V(G) \setminus \{x, y, z, w\}$. It is straightforward to see that α is an automorphism of G by the definition of $\sim_{\text{nt}}$ and $\sim_{\text{mt}}$ and the properties (i) and (ii). Hence, $G - \{x, y\}$ and $G - \{z, w\}$ are isomorphic. □

In particular, the properties above allow us to classify pairs of vertices with respect to $\sim_{\text{mt}}$ and $\sim_{\text{nt}}$.

Observation 1. *The relations $\sim_{mt}$ and $\sim_{nt}$ are equivalence relations on pairs of vertices and comparable pairs of vertices, respectively.*

Proof. It is clear that $\sim_{\text{mt}}$ (resp. $\sim_{\text{nt}}$) are reflexive and symmetric on pairs of vertices (resp. comparable vertices). Let (x_1, y_1), (x_2, y_2) and (x_3, y_3) be pairs of vertices. If $N_G[x_1] \setminus \{y_1\} = N_G[x_2] \setminus \{y_2\}$ and $N_G[x_2] \setminus \{y_2\} = N_G[x_3] \setminus \{y_3\}$, then $N_G[x_1] \setminus \{y_1\} = N_G[x_3] \setminus \{y_3\}$. Also if $N_G(x_1) \setminus \{y_1\} = N_G(x_2) \setminus \{y_2\}$ and $N_G(x_2) \setminus \{y_2\} = N_G(x_3) \setminus \{y_3\}$, then $N_G(x_1) \setminus \{y_1\} = N_G(x_3) \setminus \{y_3\}$. This immediately implies that $\sim_{\text{mt}}$ and $\sim_{\text{nt}}$ are transitive, as well. □

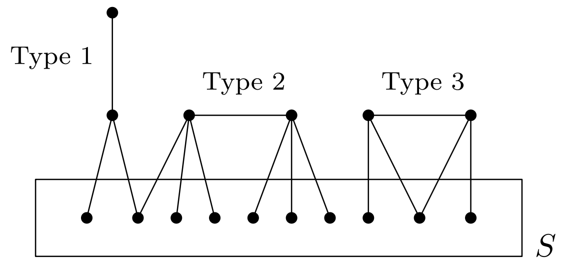

Figure 1: Types of edges of $H - S$.

Let H be a square root of a connected graph G with at least three vertices, such that H is at distance k from $pK_1 + qK_2$, and let S be a modulator. Note that $S \neq \emptyset$, because G is connected and $|V(G)| \geq 3$. Then an S-matching edge ab of H satisfies exactly one of the following conditions:

1. $N_H(a) \cap S = \emptyset$ and $N_H(b) \cap S \neq \emptyset$,
2. $N_H(a) \cap S, N_H(b) \cap S \neq \emptyset$ and $N_H(a) \cap N_H(b) \cap S = \emptyset$,
3. $N_H(a) \cap S, N_H(b) \cap S \neq \emptyset$ and $N_H(a) \cap N_H(b) \cap S \neq \emptyset$.

We refer to them as type 1, 2 and 3 edges, respectively (see Figure 1). We use the same notation for every graph F that has a set of vertices S such that $F - S$ has maximum degree at most one.

In the following three lemmas, we show the properties of the S-matching edges of types 1, 2 and 3 respectively that are crucial for our algorithm. We point out that even though some of the properties presented may be redundant, we state them in the lemmas for clarity of the explanations.

Lemma 2. *Let H be a square root of a connected graph G with at least three vertices such that $H - S$ is isomorphic to $pK_1 + qK_2$ for $S \subseteq V(G)$. If a_1b_1 and a_2b_2 are two type 1 distinct edges such that $N_H(b_1) \cap S = N_H(b_2) \cap S \neq \emptyset$, then the following holds:*

(i) (a_1, b_1) *and* (a_2, b_2) *are comparable pairs,*
(ii) $(a_1, b_1) \sim_{nt} (a_2, b_2)$,
(iii) $(a_1, b_1) \nsim_{mt} (a_2, b_2)$.

Proof. Let $A = N_H(b_1) \cap S = N_H(b_2) \cap S$. Since (a_1, b_1) is a type 1 edge, we have that $N_H(a_1) = \{b_1\}$. Thus, $N_G[a_1] = A \cup \{a_1, b_1\} \subseteq N_H[b_1] \subseteq N_G[b_1]$. The same holds for (a_2, b_2). Hence, the pairs are comparable and (i) is proved.

For (ii), note that since $N_H(b_1) \cap S = N_H(b_2) \cap S = A$, then $N_G[b_1] \setminus \{a_1\} = N_H[A] = N_G[b_2] \setminus \{a_2\}$. Moreover, since $N_H(a_1) = \{b_1\}$ and $N_H(a_2) = \{b_2\}$, we have that $N_G(a_1) \setminus \{b_1\} = A = N_G(a_2) \setminus \{b_2\}$. This shows that $(a_1, b_1) \sim_{\text{nt}} (a_2, b_2)$.

Finally, for (iii), it suffices to notice that by Lemma 1(v), we have that $a_1a_2 \notin E(G)$ and it should be $a_1a_2 \in E(G)$ if $(a_1, b_1) \sim_{\text{mt}} (a_2, b_2)$ by Lemma 1(iv). □

Lemma 3. *Let H be a square root of a connected graph G with at least three vertices such that $H - S$ is isomorphic to $pK_1 + qK_2$ for $S \subseteq V(G)$. If a_1b_1 and a_2b_2 are two distinct type 2 edges such that $N_H(a_1) \cap S = N_H(a_2) \cap S$ and $N_H(b_1) \cap S = N_H(b_2) \cap S$, then the following holds:*

(i) $(a_1, b_1) \sim_{mt} (a_2, b_2)$,
(ii) $(a_1, b_1) \nsim_{nt} (a_2, b_2)$.

Proof. Let $A = N_H(a_1) \cap S = N_H(a_2) \cap S$ and $B = N_H(b_1) \cap S = N_H(b_2) \cap S$. Since a_1b_1 and a_2b_2 are type 2 edges, we have $A \cap B = \emptyset$.

For (i), notice that $N_G[a_1] = N^2_H[a_1] = \{b_1\} \cup B \cup N_H[A]$. Therefore, we have that $N_G[a_1] \setminus \{b_1\} = N^2_H[a_1] \setminus \{b_1\} = B \cup N_H[A]$. By the same arguments, $N_G[a_2] \setminus \{b_2\} = B \cup N_H[A]$ and, therefore, $N_G[a_1] \setminus \{b_1\} = N_G[a_2] \setminus \{b_2\}$. By symmetric arguments, we obtain that $N_G[b_1] \setminus \{a_1\} = N_G[b_2] \setminus \{a_2\}$, which completes the proof that $(a_1, b_1) \sim_{\mathrm{mt}} (a_2, b_2)$.

To prove (ii), notice that $a_1a_2 \in E(G)$ by Lemma 1(iv) and, therefore, $(a_1, b_1) \nsim_{\mathrm{nt}} (a_2, b_2)$ by Lemma 1(v). $\square$

Lemma 4. *Let H be a square root of a connected graph G with at least three vertices such that $H - S$ is isomorphic to $pK_1 + qK_2$ for $S \subseteq V(G)$. If a_1b_1 and a_2b_2 are two distinct type 3 edges such that $N_H(a_1) \cap S = N_H(a_2) \cap S$ and $N_H(b_1) \cap S = N_H(b_2) \cap S$, then the following holds:*

- (i) $(a_1, b_1) \nsim_{mt} (a_2, b_2)$,
- (ii) $(a_1, b_1) \nsim_{nt} (a_2, b_2)$,
- (iii) *a_1 and a_2 (resp. b_1 and b_2) are true twins in G.*

Proof. Let $A = N_H(a_1) \cap S = N_H(a_2) \cap S$ and $B = N_H(b_1) \cap S = N_H(b_2) \cap S$. Since a_1b_1 and a_2b_2 are type 3 edges, $A \cap B \neq \emptyset$.

For (i) and (ii), it suffices to notice that since $A \cap B \neq \emptyset$, then $a_1b_2, b_1a_2 \in E(G)$. By Lemma 1(ii), we conclude that $(a_1, b_1) \nsim_{\mathrm{mt}} (a_2, b_2)$ and $(a_1, b_1) \nsim_{\mathrm{nt}} (a_2, b_2)$.

For (iii), observe that $N_G[a_1] = N^2_H[a_1] = N_H[A] \cup \{b_1\} \cup B$ by the definition. Since $A \cap B \neq \emptyset$, we have that $b_1 \in N_H[A]$. Hence, $N_G[a_1] = N_H[A] \cup B$. By the same arguments, $N_G[a_2] = N_H[A] \cup B$. Then $N_G[a_1] = N_G[a_2]$, that is, a_1 and a_2 are true twins. Clearly, the same holds for b_1 and b_2. $\square$

We also need the following straightforward observation about S-isolated vertices.

Observation 2. *Let H be a square root of a connected graph G with at least three vertices such that $H - S$ is isomorphic to $pK_1 + qK_2$ for $S \subseteq V(G)$. Then every two distinct S-isolated vertices of H with the same neighbors in S are true twins in G.*

Proof. Let u and v be two such S-isolated vertices, that is, $N_H(u) = N_H(v) \subseteq S$. Since G is connected, $N_H(u) \neq \emptyset$ and thus, $uv \in E(G)$. Moreover, since $N_H(u) = N_H(v)$, $N^2_H[u] = N^2_H[v]$. Hence $N_G[u] = N_G[v]$, that is, u and v are true twins in G. $\square$

The next lemma is used to construct reduction rules that allow to bound the size of equivalence classes of pairs of vertices with respect to $\sim_{\mathrm{nt}}$ and $\sim_{\mathrm{mt}}$.

Lemma 5. *Let H be a square root of a connected graph G with at least three vertices such that $H - S$ is isomorphic to $pK_1 + qK_2$ for a modulator $S \subseteq V(G)$ of size k. Let Q be an equivalence class in the set of pairs of comparable pairs of vertices with respect to the relation $\sim_{nt}$ (an equivalence class in the set of pairs of adjacent vertices with respect to the relation $\sim_{mt}$, respectively). If $|Q| \geq 2k + 2^{2k} + 1$, then Q contains two pairs (a_1, b_1) and (a_2, b_2) such that a_1b_1 and a_2b_2 are S-matching edges of type 1 in H satisfying $N_H(b_1) \cap S = N_H(b_2) \cap S \neq \emptyset$ (S-matching edges of type 2 in H satisfying $N_H(a_1) \cap S = N_H(a_2) \cap S$ and $N_H(b_1) \cap S = N_H(b_2) \cap S$, respectively).*

Proof. Let Q be an equivalence class of size at least $2k + 2^{2k} + 1$ with respect to $\sim_{\mathrm{nt}}$ or $\sim_{\mathrm{mt}}$.

By Lemma 1 (i), each vertex of G appears in at most one pair of Q. Since $|S| = k$, there are at most k pairs of Q with at least one element in S. Let

$$Q' = \{(x, y) \in Q \mid x, y \notin S \text{ and } xy \text{ is not a } S\text{-matching edge in } H\}.$$

We now show that $|Q'| \leq k$. Consider $(x, y) \in Q'$. Since $xy \in E(G) \setminus E(H)$, there exists $w \in V(G)$ such that $wx, wy \in E(H)$. Since $H - S$ is isomorphic to $pK_1 + qK_2$, we have that

$w \in S$. Let $(x', y') \in Q' \setminus \{(x, y)\}$. By the same argument, there exists $w' \in S$ such that $w'x', w'y' \in E(H)$. Moreover, it cannot be the case that $w = w'$, since this would imply that $xy', yx' \in E(G)$, which by Lemma 1 (ii) is a contradiction to the fact that $(x, y) \sim_{\mathrm{nt}} (x', y')$ or $(x, y) \sim_{\mathrm{mt}} (x', y')$. That is, for each pair $(x, y) \in Q'$, there is a vertex in S that is adjacent to both elements of the pair and no vertex of S can be adjacent to the elements of more than one pair of Q'. Since $|S| \leq k$, we conclude that $|Q'| \leq k$.

Since $|Q| \geq 2k + 2^{2k} + 1$, there are at least $2^{2k} + 1$ S-matching edges in Q. Given that $|S| \leq k$, by the pigeonhole principle, we have that there are two pairs $(a_1, b_1), (a_2, b_2) \in Q$ such that a_1b_1 and a_2b_2 are S-matched edges in H and $N_H(a_1) \cap S = N_H(a_2) \cap S$ and $N_H(b_1) \cap S = N_H(b_2) \cap S$. In particular, this implies that a_1b_1 and a_2b_2 are of the same type. It cannot be the case that these two edges are of type 3, since by Lemma 4(i) and (ii), these two pairs would not be equivalent with respect to $\sim_{\mathrm{nt}}$ or $\sim_{\mathrm{mt}}$. We now consider the following two cases, one for each of the mentioned equivalence relations.

Suppose that Q is an equivalence class in the set of pairs of comparable pairs of vertices with respect to the relation $\sim_{\mathrm{nt}}$. By Lemma 3(ii), they cannot be of type 2. Hence, a_1b_1 and a_2b_2 are of type 1. In particular, either $N_H(a_1) \cap S = N_H(a_2) \cap S \neq \emptyset$ or $N_H(b_1) \cap S = N_H(b_2) \cap S \neq \emptyset$. If $N_H(a_1) \cap S = N_H(a_2) \cap S \neq \emptyset$, then $a_1a_2 \in E(G)$ contradicting Lemma 1 (v). Hence, a_1b_1 and a_2b_2 are S-matching edges of type 1 in H satisfying $N_H(b_1) \cap S = N_H(b_2) \cap S \neq \emptyset$.

Let now Q be an equivalence class in the set of pairs of adjacent vertices with respect to the relation $\sim_{\mathrm{mt}}$. By Lemma 2(iii), they cannot be of type 1. Hence, a_1b_1 and a_2b_2 are of type 2. This concludes the proof of the lemma. □

3.2 The algorithm

In this section we prove our main result. First, we consider connected graphs. For this, observe that if a connected graph G has a square root H then H is connected as well. We will show the following.

Theorem 2. *Distance-k-to-$(pK_1 + qK_2)$ Square Root can be solved in time $2^{2^{\mathcal{O}(k)}} \cdot n^5$ on connected graphs. Furthermore, given G and k, in time $2^{2^{\mathcal{O}(k)}} \cdot n^5$ one can solve the problem for all pairs of nonnegative integers p and q such that $p + 2q = n - k$.*

For simplicity, in Theorem 2, we assumed that the input graph is connected but it is not difficult to extend the result for general case.

Corollary 1. *Distance-k-to-$(pK_1 + qK_2)$ Square Root can be solved in $2^{2^{\mathcal{O}(k)}} \cdot n^5$ time.*

Proof. Let (G, p, q, k) be an instance of Distance-k-to-$(pK_1 + qK_2)$ Square Root and let $C_1, \ldots, C_\ell$ be the components of G. If $\ell = 1$, we apply Theorem 2. Assume that this is not the case and $\ell \geq 2$. For each $i \in \{1, \ldots, \ell\}$, we use Theorem 2 to solve the instances (C_i, p', q', k') such that $k' \leq k$, $p' \leq p$, $q' \leq q$ and $k' + p' + 2q' = |V(C_i)|$. Then we combine these solutions to solve the input instance using a dynamic programming algorithm.

For $h \in \{1, \ldots, \ell\}$, let G_h be the subgraph of G with the components $C_1, \ldots, C_h$. Clearly, $G_h = G$. For each $h \in \{1, \ldots, \ell\}$, every triple of nonnegative integers k', p', q' such that $k' \leq k$, $p' \leq p$, $q' \leq q$ and $k' + p' + 2q' = |V(G_h)|$, we solve the instance (G_h, p', q', k'). For $h = 1$, this is already done as $G_1 = C_1$. Let $h \geq 2$. Then it is straightforward to observe that (G_h, p', q', k') is a Yes-instance if and only if there are nonnegative integers k_1, p_1, q_1 and k_2, p_2, q_2 such that

- $k_1 + k_2 = k'$, $p_1 + p_2 = p'$, $q_1 + q_2 = q'$, and
- $k_1 + p_1 + 2q_1 = |V(G_{h-1})|$ and $k_2 + p_2 + 2q_2 = |V(C_h)|$,

for which both (G_{h-1}, p_1, q_1, k_1) and (C_h, p_2, q_2, k_2) are Yes-instances. This allows to solve (G_h, p', q', k') in time $\mathcal{O}(n^2)$ if we are given the solutions for (G_{h-1}, p_1, q_1, k_1) and (C_h, p_2, q_2, k_2).

We obtain that, given the tables of solutions for the components of G, we can solve the problem for G in time $\mathcal{O}(n^5)$. Because such tables can be constructed in $2^{2^{\mathcal{O}(k)}} \cdot n^5$ time by Theorem 2, we conclude that the total running time is $2^{2^{\mathcal{O}(k)}} \cdot n^5$. $\square$

Finally, Corollary 1 gives the following statement for the related problems.

Corollary 2. *VC-k Root, Distance-k-to-Matching Square Root and Distance-k-to-Degree-One Square Root can be solved in time* $2^{2^{\mathcal{O}(k)}} \cdot n^5$.

Hence, in the remainder of this section, we focus our attention into providing a proof of Theorem 2. We first provide a sketch of the algorithm and the main ideas behind it. These ideas are then formally presented and proved in Subsections 3.2.2 and 3.2.3.

3.2.1 Sketch of the proof of Theorem 2

Recall that in the Distance-k-to-$(pK_1 + qK_2)$ Square Root problem we want to determine if G has a square root H such that $H - S$ is isomorphic to $pK_1 + qK_2$, for a modulator $S \subset V(G)$ with $|S| = k$. Our algorithm to solve this problem on a connected graph starts with a preprocessing step whose goal is to reduce the number of S-matching edges of type 1 and type 2 in a potential solution. We are able to do that by using the equivalence relations $\sim_{\text{mt}}$ and $\sim_{\text{nt}}$ defined in Section 3.1. More specifically, we show that if there is an equivalence class with respect to $\sim_{\text{mt}}$ or $\sim_{\text{nt}}$ that is of size at least $2k + 2^{2k} + 2$, we can safely delete two vertices of the input graph that form a pair in that equivalence class. After exhaustive application of this reduction rule, we are able to show that the number of S-matching edges of type 1 and type 2 in a potential solution for the obtained instance is now bounded by a function of k. As a result of this preprocessing step, we are also able to show that if our graph is a Yes-instance to the problem, then the number of true twin classes in it is bounded by a function of k. Therefore, if we have too many of such classes, we can safely return No. In summary, at this point in our algorithm we may assume have an instance (G, p, q, k) such that:

(i) $V(G)$ can be partitioned into a small (bounded by a function of k) number of true twin classes;
(ii) if (G, p, q, k) is a Yes-instance, a square root H of G with modulator S attesting this has bounded number of S-matching edges of type 1 and type 2.

Note that the number of S-matching edges of type 3 and the number of S-isolated vertices in a potential solution might be unbounded. To overcome this, we define the notion of a *solution skeleton*. Informally speaking, we obtain the skeleton of a solution H by replacing the set of S-matching edges of type 3 with the same neighborhoods in S by a single representative. We do the same for the S-isolated vertices with the same neighborhood in S. Observe that, by (ii), the skeleton of a potential solution to our instance has bounded size. We enumerate all potential solution skeletons, and test whether our instance has a square root *with that particular skeleton.* This last step is achieved by translating the problem into solving a linear integer program with bounded number of variables, which by Theorem 1, can be solved in time that is FPT by the number of variables.

A formal proof of Theorem 2 is provided in the next two subsections. In Subsection 3.2.2 we deal with the preprocessing steps, while in Subsection 3.2.3 we provide a formal definition of a solution skeleton and show how to solve our problem using them.

3.2.2 Proof of Theorem 2, part I: Reducing the number of type 1 and type 2 edges

Let (G, p, q, k) be an instance of Distance-k-to-$(pK_1 + qK_2)$ Square Root with G being a connected graph. Recall that we want to determine if G has a square root H such that $H - S$ is isomorphic to $pK_1 + qK_2$, for a modulator $S \subset V(G)$ with $|S| = k$, where $p + 2q + k = n$. If G has at most two vertices, then the problem is trivial. Notice also that if $k = 0$, then (G, p, q, k) may be a Yes-instance only if G has at most two vertices, because G is connected. Hence, from now on we assume that $n \geq 3$ and $k \geq 1$.

We exhaustively apply the following rule to reduce the number of type 1 edges in a potential solution. For this, we consider the set $\mathcal{A}$ of comparable pairs of vertices of G and find its partition into equivalence classes with respect to $\sim_{\mathrm{nt}}$. Note that $\mathcal{A}$ contains at most $2m$ elements and can be constructed in time $\mathcal{O}(mn)$. Then the partition of $\mathcal{A}$ into equivalence classes can be found in time $\mathcal{O}(m^2n)$ by checking the neighborhoods of the vertices of each pair.

Rule 2.1. *If there is an equivalence class $Q \subseteq \mathcal{A}$ with respect to $\sim_{nt}$ such that $|Q| \geq 2k + 2^{2k} + 2$, delete two vertices of G that form a pair of Q and set $q := q - 1$.*

The following claim shows that Rule 2.1 is safe.

Claim 2.1. *If G' is the graph obtained from G by the application of Rule 2.1, then G' is connected and (G, p, q, k) and $(G', p, q-1, k)$ are equivalent instances of Distance-k-to-$(pK_1 + qK_2)$ Square Root.*

Proof: Let $G' = G - \{x, y\}$ for a pair $(x, y) \in Q$.

First assume (G, p, q, k) is a Yes-instance to Distance-k-to-$(pK_1 + qK_2)$ Square Root and let H be a square root of G that is a solution to this problem with a modulator S. By Lemma 5, H has two S-matching edges $x'y'$ and $x''y''$ of type 1 such that $(x', y'), (x'', y'') \in Q$ and $N_H(y') \cap S = N_H(y'') \cap S \neq \emptyset$. Note that for any vertex u, if uy' is an edge in H, then uy'' is also an edge in H (except when $u = x'$). Hence, $H' = H - \{x', y'\}$ is a square root of $G'' = G - \{x', y'\}$ with one less S-matching edge. Moreover, H' is connected, because H is connected and $N_H(y') \setminus \{x'\} = N_{H'}(y'') \setminus \{x''\}$. This implies that G'' is connected as well. We conclude that $(G'', p, q-1, k)$ is a Yes-instance with G'' be a connected graph. Because G' and G'' are isomorphic by Lemma 1(vi), we have that $(G', p, q-1, k)$ is a Yes-instance as well and G' is connected.

Now assume $(G', p, q-1, k)$ is a Yes-instance to Distance-k-to-$(pK_1 + qK_2)$ Square Root and let H' be a square root of G' that is a solution to this problem with a modulator S. Recall that Q consists of pairs of vertices whose end-vertices are pairwise distinct by Lemma 1(i). Hence, $Q' = Q \setminus \{(x, y)\}$ contains at least $2k + 2^{2k} + 1$ elements. By the definition of $\sim_{\mathrm{nt}}$, every two pairs of $Q' = Q \setminus \{(x, y)\}$ are equivalent with respect to the relation for G'. Thus, by Lemma 5, there are $(x', y'), (x'', y'') \in Q'$ such that $x'y'$ and $x''y''$ are S-matching edges of type 1 in H' and $N_H(y') \cap S = N_H(y'') \cap S \neq \emptyset$. We construct a square root H for G by adding the edge xy to H' as an S-matching edge of type 1 with $N_H(y) \cap S = N_H(y') \cap S$. To see that H is indeed a square root for G, note that since H' was a square root for G', we have $H'^2 = G'$. Now we argue about the edges of G that are incident to x and y. Since $(x, y), (x', y') \in Q$, we have that $N_G(x) \setminus \{y\} = N_G(x') \setminus \{y'\}$ and $N_G[y] \setminus \{x\} = N_G[y'] \setminus \{x'\}$. This means that if $w \neq x$ is a neighbor of y in G, then w is also a neighbor of y'. Since H' is a square root of G', we have that either $y'w \in E(H')$ or y' and w are at distance two in H'. Since $N_H(y) \cap S = N_H(y') \cap S$, the same holds for y: it is either adjacent to w or it is at distance two from w in H. A symmetric argument holds for any edge incident to x in G. Hence, we conclude that H is indeed a square root of G. ⌟

We also want to reduce the number of type 2 edges in a potential solution. Let $\mathcal{B}$ be the set of pairs of adjacent vertices. We construct the partition of $\mathcal{B}$ into equivalence classes with respect to $\sim_{\text{mt}}$. We have that $|\mathcal{B}| = 2m$ and, therefore, the partition of $\mathcal{B}$ into equivalence classes can be found in time $\mathcal{O}(m^2 n)$ by checking the neighborhoods of the vertices of each pair. We exhaustively apply the following rule.

Rule 2.2. *If there is an equivalence class $Q \subseteq \mathcal{B}$ with respect to $\sim_{mt}$ such that $|Q| \geq 2k+2^{2k}+2$, delete two vertices of G that form a pair of Q and set $q := q-1$.*

The following claim shows that Rule 2.1 is safe.

Claim 2.2. *If G' is the graph obtained from G by the application of Rule 2.2, then G' is connected and (G, p, q, k) and $(G', p, q-1, k)$ are equivalent instances of* Distance-k-to-$(pK_1 + qK_2)$ Square Root.

Proof: The proof of this claim follows the same lines as the proof of Claim 2.1. Let $G' = G - \{x, y\}$ for $(x, y) \in Q$.

Let (G, p, q, k) be a Yes-instance to Distance-k-to-$(pK_1 + qK_2)$ Square Root and let H be a square root of G that is a solution to this problem with a modulator S. By Lemma 5, H has two S-matching edges $x'y'$ and $x''y''$ of type 2 such that $(x', y'), (x'', y'') \in Q$ and $N_H(x') \cap S = N_H(x'') \cap S$ and $N_H(y') \cap S = N_H(y'') \cap S$. Note that for any vertex u, if ux' (resp. uy') is an edge in H, then ux'' (resp. uy'') is also an edge in H, except when $u = y'$ (resp. $u = x'$). Thus, $H' = H - \{x', y'\}$ is a square root of $G'' = G - \{x', y'\}$ with one less S-matching edge. We also have that H' is connected, because H is connected. This implies that G'' is also connected. We conclude that $(G'', p, q-1, k)$ is a Yes-instance with G'' be a connected graph. Because G' and G'' are isomorphic by Lemma 1 (vi), we have that $(G', p, q-1, k)$ is a Yes-instance as well and G' is connected.

Now assume $(G', p, q-1, k)$ is a Yes-instance to Distance-k-to-$(pK_1 + qK_2)$ Square Root and let H' be a square root of G' that is a solution to this problem with a modulator S. Recall that Q consists of pairs of vertices whose end-vertices are pairwise distinct by Lemma 1(i). Hence, $Q' = Q \setminus \{(x, y)\}$ contains at least $2k + 2^{2k} + 1$ elements. By the definition of $\sim_{\text{mt}}$, every two pairs of $Q' = Q \setminus \{(x, y)\}$ are equivalent with respect to the relation for G'. Thus, by Lemma 5, there are $(x', y'), (x'', y'') \in Q'$ such that $x'y'$ and $x''y''$ are S-matching edges of type 2 in H' with $N_H(x') \cap S = N_H(x'') \cap S$ and $N_H(y') \cap S = N_H(y'') \cap S$. We construct a square root H for G by adding the edge xy to H' as a S-matching edge of type 2 with $N_H(x) \cap S = N_H(x') \cap S$ and $N_H(y) \cap S = N_H(y') \cap S$. To see that H is indeed a square root for G, note that since H' was a square root for G', we have $H'^2 = G'$. Now we argue about the edges of G that are incident to x and y. Since $(x, y), (x', y') \in Q$, we have that $N_G[x] \setminus \{y\} = N_G[x'] \setminus \{y'\}$ and $N_G[y] \setminus \{x\} = N_G[y'] \setminus \{x'\}$. This means that if $w \neq x$ is a neighbor of y in G, then w is also a neighbor of y'. Since H' is a square root of G', we have that either $y'w \in E(H')$ or y' and w are at distance two in H'. Since $N_H(y) \cap S = N_H(y') \cap S$, the same holds for y: it is either adjacent to w in H or it is at distance two from w in H. A symmetric argument holds for any edge incident to x in G. Hence, we conclude that H is indeed a square root of G. ⌟

After exhaustive application of Rules 2.1 and 2.2 we obtain the following bounds on the number of S-matching edges of types 1 and 2 in a potential solution.

Claim 2.3. *Let (G', p, q', k) be the instance of* Distance-k-to-$(pK_1 + qK_2)$ Square Root *after exhaustive applications of Rules 2.1 and 2.2. Then G' is a connected graph and a potential solution H to the instance has at most $2^k(2k + 2^{2k} + 1)$ S-matching edges of type 1 and $2^{2k}(2k + 2^{2k} + 1)$ S-matching edges of type 2.*

Proof: Clearly, G' is connected by Claims 2.1 and 2.2.

By Lemma 2(ii) and Lemma 3(ii), if two S-matching edges xy and $x'y'$ of a potential solution behave in the same way with respect to S, that is, if $N_H(x) \cap S = N_H(x') \cap S$ and $N_H(y) \cap S =$

$N_H(y') \cap S$, then they belong to the same equivalence class (either with respect to $\sim_{\mathrm{nt}}$ or to $\sim_{\mathrm{mt}}$). Hence, after exhaustive application of Rule 2.1, for each set $A \subseteq S$, there are at most $2k + 2^{2k} + 1$ S-matching edges xy such that $N_H(y) \cap S = A$. Thus, there are at most $2^k(2k + 2^{2k} + 1)$ S-matching edges of type 1 in H. Analogously, after exhaustive application of Rule 2.1, for each $A, B \subseteq S$, there are at most $2k + 2^{2k} + 1$ S-matching edges xy such that $N_H(x) \cap S = A$ and $N_H(y) \cap S = B$. Hence, there are at most $2^{2k}(2k + 2^{2k} + 1)$ S-matching edges of type 2. ⌟

For simplicity, we call (G, p, q, k) again the instance obtained after exhaustive applications of Rules 2.1 and 2.2. Notice that G can be constructed in polynomial time, since the equivalence classes according to $\sim_{\mathrm{mt}}$ and $\sim_{\mathrm{nt}}$ can be computed in time $\mathcal{O}(m^2 n)$.

By Claim 2.3, in a potential solution, the number of S-matching edges of types 1 and 2 is bounded by a function of k. We will make use of this fact to make further guesses about the structure of a potential solution. To do so, we first consider the classes of true twins of G and show the following.

Claim 2.4. *Let $\mathcal{T} = \{T_1, \ldots, T_r\}$ be the partition of $V(G)$ into classes of true twins. If (G, p, q, k) is a* Yes*-instance to our problem, then $r \leq 2(2^k + 2^{2k})(2k + 2^{2k} + 1) + k + 2^k + 2 \cdot 2^{2k}$.*

Proof: Assume (G, p, q, k) is a Yes-instance to our problem and let H be a square root of G containing a modulator S of size k such that $H - S$ is isomorphic to $pK_1 + qK_2$. Let X be the set of vertices of G that are endpoints of type 1 and type 2 S-matching edges in H. By Claim 2.3, $|X| \leq 2(2^k + 2^{2k})(2k + 2^{2k} + 1)$. Note that if two S-isolated vertices of H have the same neighborhood in S, they are true twins in G by Observation 2. Moreover, by Lemma 4(iii), if xy and $x'y'$ are two type 3 S-matching edges in H satisfying $N_H(x) \cap S = N_H(x') \cap S$ and $N_H(y) \cap S = N_H(y') \cap S$, then x and x' (resp. y and y') are true twins in G. As already explained, there are no other types of edges in $H - S$. Thus, we have at most $2(2^k + 2^{2k})(2k + 2^{2k} + 1)$ distinct classes of true twins among the vertices of X, at most k classes among the vertices of S, at most 2^k classes among the S-isolated vertices and at most $2 \cdot 2^{2k}$ classes among the vertices that are endpoints of type 3 k-matching edges. This shows that $r \leq 2(2^k + 2^{2k})(2k + 2^{2k} + 1) + k + 2^k + 2 \cdot 2^{2k}$. ⌟

Observe that the partition $\mathcal{T} = \{T_1, \ldots, T_r\}$ of $V(G)$ into classes of true twins can be constructed in linear time [29]. Using Claim 2.4, we apply the following rule.

Rule 2.3. *If $|\mathcal{T}| > 2(2^k + 2^{2k})(2k + 2^{2k} + 1) + k + 2^k + 2 \cdot 2^{2k}$, then return* No *and stop.*

From now, we assume that we do not stop by Rule 2.3. This means that $|\mathcal{T}| = \mathcal{O}(2^{4k})$.

3.2.3 Proof of Theorem 2, part II: Searching for a solution with given skeleton

Suppose that (G, p, q, k) is a Yes-instance to Distance-k-to-$(pK_1 + qK_2)$ Square Root and let H be a square root of G that is a solution to this instance with a modulator S. We say that F is the *skeleton* of H with respect to S if F is obtained from H be the exhaustive application of the following rules:

- (i) if H has two distinct type 3 S-matching edges xy and $x'y'$ with $N_H(x) \cap S = N_H(x') \cap S$ and $N_H(y) \cap S = N_H(y') \cap S$, then delete x and y,
- (ii) if H has two distinct S-isolated vertices x and y with $N_H(x) = N_H(y)$, then delete x.

In other words, we replace the set of S-matching edges of type 3 with the same neighborhoods on the end-vertices in S by a single representative and we replace the set of S-isolated vertices with the same neighborhoods by a single representative.

We say that a graph F is a *potential solution skeleton* with respect to a set $S \subseteq V(F)$ of size k for (G, p, q, k) if the following conditions hold:

(i) $F - S$ has maximum degree one, that is, $F - S$ is isomorphic to $sK_1 + tK_2$ for some nonnegative integers s and t,

(ii) for every two distinct S-isolated vertices x and y of F, $N_F(x) \neq N_F(y)$,

(iii) for every two distinct S-matching edges xy and $x'y'$ of type 3, either $N_F(x)\cap S \neq N_H(x')\cap S$ or $N_F(y) \cap S \neq N_H(y') \cap S$,

(iv) for every $A, B \subseteq S$ such that $A \cap B = \emptyset$ and at least one of A and B is nonempty, $\{xy \in E(F - S) \mid N_F(x) \cap S = A \text{ and } N_F(y) \cap S = B\}$ has size at most $2k + 2^{2k} + 1$.

Note that (iv) means that the number of type 1 and type 2 S-matched edges with the same neighbors in S is upper bounded by $2k + 2^{2k} + 1$. Since Rules 2.1 and 2.2 cannot be applied to (G, p, q, k), we obtain the following claim by Lemmas 2(ii) and 3(ii).

Claim 2.5. *Every skeleton of a solution to (G, p, q, k) is a potential solution skeleton for this instance with respect to the modulator S.*

We observe that each potential solution skeleton has bounded size.

Claim 2.6. *For every potential solution skeleton F for (G, p, q, k),*

$$|V(F)| \leq k + 2^k + 2 \cdot 2^{2k} + 2 \cdot 2^{2k}(2k + 2^{2k} + 1).$$

Proof: By the definition, F has k vertices in S, at most 2^k S-isolated vertices and at most $2 \cdot 2^{2k}$ end-vertices of S-matching edges of type 3. For each $A, B \subseteq S$ such that $A \cap B = \emptyset$ and at least one of A and B is nonempty, F has at most $k + 2^{2k} + 1$ S-matching edges xy of type 1 or type 2 with $N_F(x) \cap S = A$ and $N_F(y) \cap S = B$. Then we have at most $2 \cdot 2^{2k}(2k + 2^{2k} + 1)$ end-vertices of these edges. ⌟

Moreover, we can construct the family $\mathcal{F}$ of all potential solution skeletons together with their modulators.

Claim 2.7. *The family $\mathcal{F}$ of all pairs (F, S), where F is a potential solution skeleton and $S \subseteq V(F)$ is a modulator of size k, has size at most $2^{\binom{k}{2}} + 2^{2^k} + 2^{2^{2k}} + (2k + 2^{2k} + 2)^{2^{2k}}$ and can be constructed in time $2^{2^{\mathcal{O}(k)}}$.*

Proof: There are at most $2^{\binom{k}{2}}$ distinct subgraph with the set of vertices S of size k. We have at most 2^{2^k} distinct sets of S-isolated vertices and there are at most $2^{2^{2k}}$ distinct sets of S-matching edges of type 3. For each $A, B \subseteq S$ such that $A \cap B = \emptyset$ and at least one of A and B is nonempty, there are $k + 2^{2k} + 2$ possible sets of S-matching edges xy of type 1 or type 2 such that $N_F(x) \cap S = A$ and $N_F(y) \cap S = B$. Therefore, we have at most $(2k + 2^{2k} + 2)^{2^{2k}}$ distinct sets of type 1 or type 2. Then $|\mathcal{F}| \leq 2^{\binom{k}{2}} + 2^{2^k} + 2^{2^{2k}} + (2k + 2^{2k} + 2)^{2^{2k}}$. Finally, it is straightforward to see that $\mathcal{F}$ can be constructed in $2^{2^{\mathcal{O}(k)}}$ time. ⌟

Using Claim 2.7, we construct $\mathcal{F}$, and for every $(F, S) \in \mathcal{F}$, we check whether there is a solution H to (G, p, q, k) with a modulator S', whose skeleton is isomorphic to F with an isomorphism that maps S to S'. If we find such a solution, then (G, p, q, k) is a YES-instance. Otherwise, Claims 2.5 guarantees that (G, p, q, k) is a NO-instance.

Assume that we are given $(F, S) \in \mathcal{F}$ for the instance (G, p, q, k).

Recall that we have the partition $\mathcal{T} = \{T_1, \ldots, T_r\}$ of $V(G)$ into true twin classes of size at most $2(2^k + 2^{2k})(2k + 2^{2k} + 1) + k + 2^k + 2 \cdot 2^{2k}$ by Rule 2.3. Recall also that the prime-twin graph $\mathcal{G}$ of G is the graph with the vertex set $\mathcal{T}$ such that two distinct vertices T_i and T_j of $\mathcal{G}$ are adjacent if and only if $uv \in E(G)$ for $u \in T_i$ and $v \in T_j$. Clearly, given G and $\mathcal{T}$, $\mathcal{G}$ can be constructed in linear time. For an induced subgraph R of G, we define $\tau_R \colon V(R) \to \mathcal{T}$ to be a mapping such that $\tau_R(v) = T_i$ if $v \in T_i$ for $T_i \in \mathcal{T}$.

Let $\varphi \colon V(F) \to \mathcal{T}$ be a surjective mapping. We say that φ is *$\mathcal{G}$-compatible* if every two distinct vertices u and v of F are adjacent in F^2 if and only if $\varphi(u)$ and $\varphi(v)$ are adjacent in $\mathcal{G}$.

Claim 2.8. *Let F be the skeleton of a solution H to (G, p, q, k). Then $\tau_F\colon V(F) \to \mathcal{T}$ is a $\mathcal{G}$-compatible surjection.*

Proof: Recall that $H^2 = G$ and F is an induced subgraph of H. Then the definition of F and Lemma 4 (iii) immediately imply that τ_F is a $\mathcal{G}$-compatible surjection. ⌟

Our next step is to reduce our problem to solving a system of linear integer inequalities. Let $\varphi\colon V(F) \to \mathcal{T}$ be a $\mathcal{G}$-compatible surjective mapping. Let X_1, X_2 and X_3 be the sets of end-vertices of the S-matching edges of type 1, type 2 and type 3 respectively in F. Let also Y be the set of S-isolated vertices of F. For every vertex $v \in V(F)$, we introduce an integer variable x_v. Informally, x_v is the number of vertices of a potential solution H that correspond to a vertex v in the solution skeleton (recall that a single representative was kept in a skeleton for any set of S-matching edges of type 3 with the same neighborhood in S; the same holds for a set of S-isolated vertices with the same neighborhood in S).

$$\begin{cases} x_v = 1 & \text{for } v \in S \cup X_1 \cup X_2, \\ x_v \geq 1 & \text{for } v \in Y \cup X_3, \\ x_u - x_v = 0 & \text{for every type 3 edge } uv, \\ \sum_{v \in Y} x_v = p, & \\ \sum_{v \in X_1 \cup X_2 \cup X_3} x_v = 2q, & \\ \sum_{v \in \varphi^{-1}(T_i)} x_v = |T_i| & \text{for } T_i \in \mathcal{T}. \end{cases} \tag{1}$$

Note that the total number of variables is $|V(F)|$, which by Claim 2.6, is bounded by $2^{O(k)}$. The following claim is crucial for our algorithm.

Claim 2.9. *The instance (G, p, q, k) has a solution H with a modulator S' such that there is an isomorphism $\psi\colon V(F) \to V(F')$ for the skeleton F' of H mapping S to S' if and only if there is a $\mathcal{G}$-compatible surjective mapping $\varphi\colon V(F) \to \mathcal{T}$ such that the system (1) has a solution.*

Proof: Suppose that there is a solution H to (G, p, q, k) with a modulator S', whose skeleton F' is isomorphic to F with an isomorphism that maps S to S'. To simplify the notation, we identify F and F' and identify S and S'. We set $\varphi = \tau_F$. By Claim 2.8, φ is a $\mathcal{G}$-compatible surjection. For $v \in Y$, we define the value of

$$x_v = |\{u \in V(H) \mid u \text{ is an } S\text{-isolated and } N_H(u) = N_H(v)\}|.$$

For each S-matching edge uv of type 3 of F,

$$\begin{aligned} x_u = x_v = |\{xy \in E(H) \mid & xy \text{ is an } S\text{-matching edge}, \\ & N_H(x) \cap S = N_H(u) \cap S \text{ and } N_H(y) \cap S = N_H(v) \cap S\}|. \end{aligned}$$

This defines the value of the variables x_v for $v \in X_3$. Recall that for all $v \notin Y \cup X_3$, $x_v = 1$ by the definition of (1). It is straightforward to verify that the constructed assignment of the variables gives a solution of (1) for φ.

For the opposite direction, let $\varphi\colon V(F) \to \mathcal{T}$ be a $\mathcal{G}$-compatible surjective mapping such that the system (1) has a solution. Assume that the variables x_v have values that satisfy (1). We construct the graph $\hat{F}$ from F and the extension $\hat{\varphi}$ of φ as follows.

- For every S-isolated vertex v of F, replace v by x_v copies that are adjacent to the same vertices as v and define $\hat{\varphi}(x) = \varphi(v)$ for the constructed vertices.
- For every S-matching edge uv of type 3, replace u and v by $x_u = x_v$ copies of pairs of adjacent vertices x and y, make x and y adjacent to the same vertices of S as u and v respectively, and define $\hat{\varphi}(x) = \varphi(u)$ and $\hat{\varphi}(y) = \varphi(v)$ respectively.

· Set $\hat{\varphi}(v) = \varphi(v)$ for the remaining vertices.

Observe that by the construction and the assumption that the values of the variables x_v satisfy (1), $\hat{F}$ has p S-isolated vertices, q S-matching edges, and for every $T_i \in \mathcal{T}$, $|\{v \in V(\hat{F}) \mid v \in \hat{\varphi}^{-1}(T_i)\}| = |T_i|$. We define $\psi\colon V(G) \to V(G)$ by mapping $|T_i|$ vertices of $\{v \in V(\hat{F}) \mid v \in \hat{\varphi}^{-1}(T_i)\}$ arbitrarily into distinct vertices of $T_i \subseteq V(G)$ for each $T_i \in \mathcal{T}$. Clearly, ψ is a bijection. Notice that by Lemma 4 (iii) and Observation 2, the sets of vertices of $\hat{F}$ constructed from S-isolated vertices and the end-vertices of S-matching edges are sets of true twins in $\hat{F}^2$. Also we have that, because φ is $\mathcal{G}$-compatible, two distinct vertices $u, v \in V(\hat{F})$ are adjacent in $\hat{F}^2$ if and only if either $\hat{\varphi}(u) = \hat{\varphi}(v)$ or $\hat{\varphi}(u) \neq \hat{\varphi}(v)$ and $\hat{\varphi}(u)\hat{\varphi}(v) \in E(\mathcal{G})$. This implies that ψ is an isomorphism of $\hat{F}^2$ and G, which means that G has a square root isomorphic to $\hat{F}$. Clearly, $\psi|_{V(F)}$ is an isomorphism of F into the skeleton of H mapping S to S'. ⌟

By Claim 2.9, we can state our task as follows: verify whether there is a $\mathcal{G}$-compatible surjection $\varphi\colon V(F) \to \mathcal{T}$ such that (1) has a solution.

For this, we consider all at most $|V(F)|^{|\mathcal{T}|} = 2^{2^{\mathcal{O}(k)}}$ surjections $\varphi\colon V(F) \to \mathcal{T}$. For each φ, we verify whether it is $\mathcal{G}$-compatible. Clearly, it can be done in time $\mathcal{O}(|V(F)|^3)$. If φ is $\mathcal{G}$-compatible, we construct the system (1) with $|V(F)| = 2^{\mathcal{O}(k)}$ variables in time $\mathcal{O}(|V(F)|^2)$. Then we solve it by applying Theorem 1 in $2^{2^{\mathcal{O}(k)}} \log n$ time. This completes the description of the algorithm and its correctness proof.

3.2.4 Running time analysis

To evaluate the total running time, notice that the preprocessing step described in Subsection 3.2.2, that is, the exhaustive application of Rules 2.1 and 2.2 is done in polynomial time. In particular, constructing sets $\mathcal{A}$ and $\mathcal{B}$ can be done on $\mathcal{O}(m^2 n)$ time. Then Rules 2.1 and 2.2 can be applied in $\mathcal{O}(m)$ time. Then the construction of $\mathcal{T}$, $\mathcal{G}$ is and the application of Rule 2.3 is done in linear time. By Claim 2.7, the family $\mathcal{F}$, described in Subsection 3.2.3, is constructed in time $2^{2^{\mathcal{O}(k)}}$. The final steps, that is, constructing φ and systems (1) and solving the systems, can be done in time $2^{2^{\mathcal{O}(k)}} \log n$. Therefore, the total running time is $2^{2^{\mathcal{O}(k)}} \cdot n^5$.

It remains to show that, given G and k, we can solve the problem for all pairs of nonnegative integers p and q such that $p + 2q = n - k$ in $2^{2^{\mathcal{O}(k)}} \cdot n^5$ time. Notice that the polynomial factor in the running time is dominated by the time needed to construct $\mathcal{A}$ and $\mathcal{B}$. If we wish to solve the problem for different values of p and q, we can construct these sets just once. Furthermore, we can construct $\mathcal{T}$ once as well. This implies that we can solve the problem for a family of instances that differ only by the values of p and q in $2^{2^{\mathcal{O}(k)}} \cdot n^5$ time.

4 A lower bound for Distance-k-to-$(pK_1 + qK_2)$ Square Root

In this section, we show that the running time of our algorithm for Distance-k-to-(pK_1+qK_2) Square Root given in Section 3 (see Theorem 2) cannot be significantly improved. In fact, we show that the Distance-k-to-(pK_1+qK_2) Square Root problem admits a double-exponential lower bound, even for the special case $q = 0$, that is, in the case of VC-k Root.

To provide a lower bound for the VC-k Root problem, we will give a parameterized reduction from the Biclique Cover problem. This problem takes as input a bipartite graph G and a nonnegative integer k, and the task is to decide whether the edges of G can be covered by at most k complete bipartite subgraphs. Chandran et al. [2] showed the following two results about the Biclique Cover problem that will be of interest to us.

Theorem 3 ([2]). *Biclique Cover cannot be solved in time $2^{2^{o(k)}} \cdot n^{\mathcal{O}(1)}$ unless ETH is false.*

Theorem 4 ([2]). *Biclique Cover does not admit a kernel of size $2^{o(k)}$ unless $\mathsf{P} = \mathsf{NP}$.*

Lemma 6. *There exists a polynomial time algorithm that, given an instance (B, k) for* Biclique Cover*, produces an equivalent instance $(G, k+4)$ for VC-k* Root*, with $|V(G)| = |V(B)|+k+6$.*

Proof. Let (B, k) be an instance of Biclique Cover where (X, Y) is the bipartition of $V(B)$. Let $X = \{x_1, \ldots, x_p\}$ and $Y = \{y_1, \ldots, y_q\}$. We construct the instance $(G, k + 4)$ for VC-k Root such that $V(G) = X \cup Y \cup \{z_1, \ldots, z_k\} \cup \{u, v, w, u', v', w'\}$. Denote by Z the set $\{z_1, \ldots, z_k\}$. The edge set of G is defined in the following way: $G[X \cup Z \cup \{u\}]$, $G[X \cup \{v\}]$, $\{u, v, w\}$, $G[Y \cup Z \cup \{u'\}]$, $G[Y \cup \{v'\}]$ and $\{u', v', w'\}$ are cliques and $x_iy_j \in E(G)$ if and only if $x_iy_j \in E(B)$. The construction of G is shown in Figure 2.

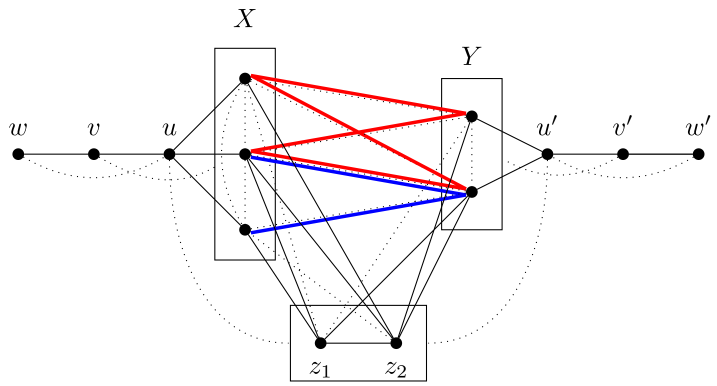

Figure 2: Illustrating the graphs G and H considered in the proof of Lemma 6. The sets X and Y form the bipartition of an instance of Biclique Cover and the two colored completed bipartite subgraphs correspond to a solution of the problem, where $k = 2$. The constructed graph G of VC-k Root is depicted by the solid and dotted black edges, whereas the graph spanned by the solid black edges corresponds to the square root H of G.

For the forward direction, suppose (B, k) is a Yes-instance for Biclique Cover. We will show that $(G, k + 4)$ is a Yes-instance for VC-k Root. Note that if B has a biclique cover of size strictly less than k, we can add arbitrary bicliques to this cover and obtain a biclique cover for B of size exactly k. Let $\mathcal{C} = \{C_1, \ldots, C_k\}$ be such a biclique cover. We construct the following square root candidate H for G with $V(H) = V(G)$. Add the edges uv, vw, $u'v'$ and $v'w'$ to H, and also all the edges between u and X, all the edges between u' and Y and all the edges in $G[Z]$. Finally, for each $1 \leq i \leq k$, add to H all the edges between z_i and the vertices of C_i.

Claim 4.1. *The constructed graph H is indeed a square root of G.*

Proof: We show that $xy \in E(G)$ if and only if $d_H(x, y) \leq 2$. For the forward direction, let $xy \in E(G)$ be such that $xy \notin E(H)$. If $xy = uw$, note that $uv, vw \in E(H)$. If $xy = vx_i$ for some i, note that $uv, ux_i \in E(G)$. If $xy = x_ix_j$, then $ux_i, ux_j \in E(H)$. If $xy = uz_i$, let x_j be a vertex of C_i and note that $ux_j, x_jz_i \in E(H)$. If $xy = x_jz_i$, let ℓ be such that $x_j \in C_\ell$ and observe that $x_jz_\ell, z_\ell z_i \in E(H)$. Symmetric arguments apply for the edges in $G[Y \cup Z \cup \{u', v', w'\}]$. Finally, if $xy = x_iy_j$, let C_ℓ be the biclique containing the edge x_iy_j in $\mathcal{C}$ and note that $x_iz_\ell, y_jz_\ell \in E(H)$. For the other direction, note that, by construction, if $xy \in E(H)$, then $xy \in E(G)$. It is easy to check that if $d_H(x, y) = 2$, then $xy \in E(G)$ (see the dotted edges in Figure 2). ⌟

We conclude that $(G, k + 4)$ is a Yes-instance for VC-k Root by Claim 4.1 together with the fact that $Z \cup \{u, v, u', v'\}$ is a vertex cover of H of size $k + 4$.

Before we show the reverse direction of the theorem, we state the next three claims, that concern the structure of any square root of the graph G.

Claim 4.2. *The edges uv, vw, $u'v'$ and $v'w'$ belong to any square root of G.*

Proof: Suppose for a contradiction that the graph G has a square root H such that $vw \notin E(H)$. In this case, it holds that $vu, uw \in E(H)$, since u is the only common neighbor of v and w. However, since $wx_i \notin E(G)$ for every $1 \leq i \leq p$ and $wz_j \notin E(G)$ for every $1 \leq j \leq k$, then $ux_i, uz_j \notin E(H)$. Therefore, there must exist an induced P_3 in H with endpoints, for instance, u and z_ℓ, for some ℓ. However, since $ux_i, uz_j \notin E(H)$ for every $1 \leq i \leq p$ and every $1 \leq j \leq k$ and $N_G(u) = X \cup Z \cup \{v, w\}$, either v or w have to be the middle vertex of the P_3. This is a contradiction, since $vz_\ell, wz_\ell \notin E(G)$.

Now suppose for a contradiction that the graph G has a square root H such that $uv \notin E(H)$. If there exists ℓ such that $ux_\ell, vx_\ell \in E(H)$, then we have a contradiction, since this would imply that no edge incident to w can be in H, given that $wx_\ell \notin E(G)$. We can then conclude that $vw, uw \in E(H)$. We can now reach a contradiction by the same argument as used in the previous paragraph.

The claim follows by a symmetric argument for the edges $u'v'$ and $v'w'$. ⌟

Claim 4.3. *The edges $\{ux_i, u'y_j \mid 1 \leq i \leq p, 1 \leq j \leq q\}$ belong to any square root of G.*

Proof: Suppose for a contradiction that G has a square root H such that $ux_i \notin E(H)$ for some $1 \leq i \leq p$. By Claim 4.2, $uv, vw \in E(H)$. This implies that $vx_i \notin E(H)$, since $wx_i \notin E(G)$. Since $ux_i \notin E(H)$ by assumption, there must exist j such that x_j is the middle vertex of a P_3 in H with endpoints v and x_i. However, this is a contradiction, since $wx_j \notin E(G)$. The claim follows by a symmetric argument for the edges of the form $u'y_j$. ⌟

Claim 4.4. *The edges $\{x_iy_j \mid 1 \leq i \leq p, 1 \leq j \leq q\}$ do not belong to any square root of G.*

Proof: Suppose for a contradiction that G has a square root H such that $x_iy_j \in E(H)$ for some $1 \leq i \leq p$ and $1 \leq j \leq q$. By Claim 4.3, we have that $ux_i \in E(H)$, which is a contradiction since $uy_j \notin E(G)$. ⌟

Now, for the reverse direction of the theorem, assume that G has a square root H that has a vertex cover of size at most $k+4$. By Claim 4.4, for every edge of G of the form x_iy_j, it holds that $x_iy_j \notin E(H)$. This implies that, for every such edge, there exists an induced P_3 in H having x_i and y_j as its endpoints. Since $N_G(x_i) \cap N_G(y_j) = Z$, only vertices of Z can be the middle vertices of these paths. For $1 \leq \ell \leq k$, let $C_\ell = N_H(z_\ell) \cap (X \cup Y)$. We will now show that $\mathcal{C} = \{C_1, \ldots, C_k\}$ is a biclique cover of B. First, note that since for every edge x_iy_j, there exists $z_h \in Z$ such that $z_hx_i, z_hy_j \in E(H)$, we conclude that $x_iy_j \in C_h$, which implies that $\mathcal{C}$ is an edge cover of B. Furthermore, for a given ℓ, since every vertex of C_ℓ is adjacent to z_ℓ in H, $G[C_\ell]$ is a clique and, therefore, $B[C_\ell]$ is a biclique. This implies that $\mathcal{C}$ is indeed a biclique cover of B of size k, which concludes the proof of the theorem. ☐

From Theorem 3 and Lemma 6 we obtain the following theorems.

Theorem 5. *VC-k Root cannot be solved in time $2^{2^{o(k)}} \cdot n^{\mathcal{O}(1)}$ unless ETH is false.*

Moreover, from Theorem 4 and Lemma 6 we can also conclude the following corollary.

Theorem 6. *VC-k Root does not admit a kernel of size $2^{o(k)}$ unless* $\mathsf{P} = \mathsf{NP}$.

Proof. Assume that VC-k Root has a kernel of size $2^{o(k)}$. Since VC-k Root is in NP and Biclique Cover is NP-complete, there is an algorithm $\mathcal{A}$ that in time $\mathcal{O}(n^c)$ reduces VC-k Root to Biclique Cover, where c is a positive constant. Then combining the reduction from Lemma 6, the kernelization algorithm for VC-k Root and $\mathcal{A}$, we obtain a kernel for Biclique Cover of size $(2^{o(k)})^c$ that is subexponential in k. By Theorem 4, this is impossible unless $\mathsf{P} = \mathsf{NP}$. Equivalently, we can observe that Chandran et al. [2], in fact, proved a stronger claim. Their proof shows that Biclique Cover does not admit a *compression* (we refer to [5] for the definition of the notion) of subexponential in k size to any problem in NP. ☐

5 Distance-k-to-Clique Square Root

In this section, we consider the complexity of testing whether a graph admits a square root of bounded deletion distance to a clique. More formally, we consider the following problem:

DISTANCE-k-TO-CLIQUE SQUARE ROOT

Input: A graph G and nonnegative integer k.

Task: Decide whether there is a square root H of G such that $H - S$ is a complete graph for a set S on k vertices.

We give an algorithm running in FPT-time parameterized by k, the size of the deletion set. That is, we prove the following theorem.

Theorem 7. *DISTANCE-k-TO-CLIQUE SQUARE ROOT can be solved in time $2^{2^{\mathcal{O}(k)}} \cdot (n+m)$.*

Proof. Let (G, k) be an instance to DISTANCE-k-TO-CLIQUE SQUARE ROOT. We start by computing the number of classes of true twins in G. Recall that this can be done in linear time [29]. If G has at least $2^k + k + 1$ classes of true twins, then G is a NO-instance to the problem, as we show in the following claim.

Claim 7.1. *Let G be a graph and H be a square root of G such that $H - S$ is a complete graph, with $|S| = k$. Let $T_1, \dots, T_t$ be a partition of $V(G)$ into classes of true twins. Then $t \le 2^k + k$.*

Proof: Let $C = V(H) \setminus S$. Note that if $u, v \in C$ and $N_H(u) \cap S = N_H(v) \cap S$, then u and v are true twins in G. Thus, we have at most 2^k distinct classes of true twins among the vertices of C, and at most k among the vertices of S. ⌟

Hence, from now on we assume that G has at most $2^k + k$ classes of true twins. We exhaustively apply the following rule in order to decrease the size of each class of true twins in G.

Rule 7.1. *If $|T_i| \ge 2^k + k + 1$ for some i, delete a vertex from T_i.*

The following claim shows that Rule 7.1 is safe.

Claim 7.2. *If G' is the graph obtained from G by the application of Rule 7.1, then (G, k) and (G', k) are equivalent instances of DISTANCE-k-TO-CLIQUE SQUARE ROOT.*

Proof: Let $G' = G - v$. First assume (G, k) is a YES-instance to DISTANCE-k-TO-CLIQUE SQUARE ROOT and let H be a square root of G that is a solution to this problem. Since $|T_i| \ge 2^k + k + 1$ and G has at most $2^k + k$ classes of true twins, by the pigeonhole principle there are two vertices $x, y \in T_i$ such that, in H, $x, y \notin S$ and $N_H[x] \cap S = N_H[y] \cap S$. That is, x and y are true twins in H also. Thus, $H' = H - x$ is a square root for $G'' = G - x$ such that $H' - S$ is a complete graph. Since G' and G'' are isomorphic, we have that (G', k) is a YES-instance as well.

Now assume (G', k) is a yes-instance to DISTANCE-k-TO-CLIQUE SQUARE ROOT and let H' be a square root of G' that is a solution to the problem. Note that $T_i \setminus \{v\}$ is a true twin class of G' of size at least $2^k + k$. Thus, there exists $u \in T_i \setminus \{v\}$ such that, in H', $u \notin S$. We can add v to H' as a true twin of u and obtain a square root H for G such that $H - S$ is a complete graph. ⌟

After exhaustive application of Rule 7.1, we obtain an instance (G', p', k) such that G' contains at most $(2^k + k)^2$ vertices, since it has at most $2^k + k$ twin classes, each of size at most $2^k + k$. Moreover, (G', k) and (G, k) are equivalent instances of DISTANCE-k-TO-CLIQUE SQUARE ROOT. We can now check by brute force whether (G', k) is YES-instance to the problem. Since G' has $2^{\mathcal{O}(k)}$ vertices, this can be done in time $2^{2^{\mathcal{O}(k)}}$. We obtain that the total running time is $2^{2^{\mathcal{O}(k)}} \cdot (n+m)$, which concludes the proof of the theorem. □

6 Conclusion

In this work, we showed that Distance-k-to-$(pK_1 + qK_2)$ Square Root and its variants can be solved in $2^{2^{\mathcal{O}(k)}} \cdot n^{\mathcal{O}(1)}$ time. We also proved that the double-exponential dependence on k is unavoidable up to Exponential Time Hypothesis, that is, the problem cannot be solved in $2^{2^{o(k)}} \cdot n^{\mathcal{O}(1)}$ time unless ETH fails. We also proved that the problem does not admit a kernel of subexponential in k size unless $\mathsf{P} = \mathsf{NP}$. We believe that it would be interesting to further investigate the parameterized complexity of $\mathcal{H}$-Square Root for sparse graph classes $\mathcal{H}$ under structural parameterizations. The natural candidates are the Distance-k-to-Linear-Forest Square Root and Feedback-Vertex Set-k Square Root problems, whose tasks are to decide whether the input graph has a square root that can be made a linear forest, that is, a union of paths, and a forest respectively by (at most) k vertex deletions. Recall that the existence of an FPT algorithm for $\mathcal{H}$-Square Root does not imply the same for subclasses of $\mathcal{H}$. However, it can be noted that the reduction from Lemma 6 implies that our complexity lower bounds still hold and, therefore, we cannot expect that these problems would be easier.

Parameterized complexity of $\mathcal{H}$-Square Root is widely open for other, not necessarily sparse, graph classes. We considered the Distance-k-to-Clique Square Root problem and proved that it is FPT when parameterized by k. What can be said if we ask for a square root that is at deletion distance (at most) k form a *cluster graph*, that is, the disjoint union of cliques? We believe that our techniques allows to show that this problem is FPT when parameterized by k if the number of cliques is a fixed constant. Is the problem FPT without this constraint?

References

[1] J. A. Bondy and U. S. R. Murty, *Graph Theory*, Springer, 2008.

[2] S. Chandran, D. Issac, and A. Karrenbauer, *On the parameterized complexity of biclique cover and partition*, in Proceedings of IPEC 2016, vol. 63 of LIPIcs, 2017, pp. 11:1–11:13.

[3] M. Cochefert, J. Couturier, P. A. Golovach, D. Kratsch, and D. Paulusma, *Parameterized algorithms for finding square roots*, Algorithmica, 74 (2016), pp. 602–629.

[4] M. Cochefert, J. Couturier, P. A. Golovach, D. Kratsch, D. Paulusma, and A. Stewart, *Computing square roots of graphs with low maximum degree*, Discrete Applied Mathematics, 248 (2018), pp. 93–101.

[5] M. Cygan, F. V. Fomin, L. Kowalik, D. Lokshtanov, D. Marx, M. Pilipczuk, M. Pilipczuk, and S. Saurabh, *Parameterized Algorithms*, Springer, 2015.

[6] G. Ducoffe, *Finding cut-vertices in the square roots of a graph*, Discrete Applied Mathematics, 257 (2019), pp. 158–174.

[7] B. Farzad and M. Karimi, *Square-root finding problem in graphs, a complete dichotomy theorem*, CoRR, abs/1210.7684 (2012).

[8] B. Farzad, L. C. Lau, V. B. Le, and N. N. Tuy, *Complexity of finding graph roots with girth conditions*, Algorithmica, 62 (2012), pp. 38–53.

[9] A. Frank and E. Tardos, *An application of simultaneous diophantine approximation in combinatorial optimization*, Combinatorica, 7 (1987), pp. 49–65.

[10] P. A. Golovach, P. Heggernes, D. Kratsch, P. T. Lima, and D. Paulusma, *Algorithms for outerplanar graph roots and graph roots of pathwidth at most 2*, Algorithmica, 81 (2019), pp. 2795–2828.

[11] P. A. Golovach, D. Kratsch, D. Paulusma, and A. Stewart, *A linear kernel for finding square roots of almost planar graphs*, Theor. Comput. Sci., 689 (2017), pp. 36–47.

[12] P. A. Golovach, D. Kratsch, D. Paulusma, and A. Stewart, *Finding cactus roots in polynomial time*, Theory of Computing Systems, 62 (2018), pp. 1409–1426.

[13] P. A. Golovach, P. T. Lima, and C. Papadopoulos, *Graph square roots of small distance from degree one graphs*, in Proceedings of LATIN 2020, 2020, pp. 116–128.

[14] R. Impagliazzo and R. Paturi, *On the complexity of k-sat*, J. Comput. Syst. Sci., 62 (2001), pp. 367–375.

[15] R. Impagliazzo, R. Paturi, and F. Zane, *Which problems have strongly exponential complexity?*, J. Comput. Syst. Sci., 63 (2001), pp. 512–530.

[16] R. Kannan, *Minkowski's convex body theorem and integer programming*, Mathematics of Operations Research, 12 (1987), pp. 415–440.

[17] L. C. Lau, *Bipartite roots of graphs*, ACM Transactions on Algorithms, 2 (2006), pp. 178–208.

[18] L. C. Lau and D. G. Corneil, *Recognizing powers of proper interval, split, and chordal graph*, SIAM J. Discrete Math., 18 (2004), pp. 83–102.

[19] V. B. Le, A. Oversberg, and O. Schaudt, *Polynomial time recognition of squares of ptolemaic graphs and 3-sun-free split graphs*, Theoretical Computer Science, 602 (2015), pp. 39–49.

[20] ——, *A unified approach for recognizing squares of split graphs*, Theoretical Computer Science, 648 (2016), pp. 26–33.

[21] V. B. Le and N. N. Tuy, *The square of a block graph*, Discrete Mathematics, 310 (2010), pp. 734–741.

[22] ——, *A good characterization of squares of strongly chordal split graphs*, Inf. Process. Lett., 111 (2011), pp. 120–123.

[23] H. W. Lenstra, *Integer programming with a fixed number of variables*, Mathematics of Operations Research, 8 (1983), pp. 538–548.

[24] Y.-L. Lin and S. Skiena, *Algorithms for square roots of graphs*, SIAM J. Discrete Math., 8 (1995), pp. 99–118.

[25] M. Milanic, A. Oversberg, and O. Schaudt, *A characterization of line graphs that are squares of graphs*, Discrete Applied Mathematics, 173 (2014), pp. 83–91.

[26] M. Milanic and O. Schaudt, *Computing square roots of trivially perfect and threshold graphs*, Discrete Applied Mathematics, 161 (2013), pp. 1538–1545.

[27] R. Motwani and M. Sudan, *Computing roots of graphs is hard*, Discrete Applied Mathematics, 54 (1994), pp. 81–88.

[28] N. V. Nestoridis and D. M. Thilikos, *Square roots of minor closed graph classes*, Discrete Applied Mathematics, 168 (2014), pp. 34–39.

[29] M. Tedder, D. Corneil, M. Habib, and C. Paul, *Simpler linear-time modular decomposition via recursive factorizing permutations*, in Proceedings of ICALP 2008, 2008, pp. 634–645.

[30] M. Yannakakis, *Node-deletion problems on bipartite graphs*, SIAM Journal on Computing, 10 (1981), pp. 310–327.