



**HAL**  
open science

# On the minimum number of inversions to make a digraph $k$ -(arc-)strong

Julien Duron, Frédéric Havet, Florian Hörsch, Clément Rambaud

► **To cite this version:**

Julien Duron, Frédéric Havet, Florian Hörsch, Clément Rambaud. On the minimum number of inversions to make a digraph  $k$ -(arc-)strong. EUROCOMB 2023 - European Conference on Combinatorics, Graph Theory and Applications, Aug 2023, Prague, Czech Republic. pp.386-392, 10.5817/CZ.MUNI.EUROCOMB23-054 . hal-04352250

**HAL Id: hal-04352250**

**<https://hal.science/hal-04352250v1>**

Submitted on 19 Dec 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# On the minimum number of inversions to make a digraph $k$ -(arc-)strong.

Julien Duron<sup>1</sup>, Frédéric Havet<sup>2</sup>, Florian Hörsch<sup>3</sup>, and Clément Rambaud<sup>2,4</sup>

<sup>1</sup>École normale supérieure de Lyon, LIP, France

<sup>2</sup>Université Côte d'Azur, CNRS, Inria, I3S, Sophia Antipolis, France

<sup>3</sup>CISPA Saarbrücken, Germany

<sup>4</sup>DIENS, École normale supérieure, CNRS, PSL University, Paris, France

## Abstract

The *inversion* of a set  $X$  of vertices in a digraph  $D$  consists of reversing the direction of all arcs of  $D\langle X \rangle$ . We study  $\text{sinv}'_k(D)$  (resp.  $\text{sinv}_k(D)$ ) which is the minimum number of inversions needed to transform  $D$  into a  $k$ -arc-strong (resp.  $k$ -strong) digraph and  $\text{sinv}'_k(n) = \max\{\text{sinv}'_k(D) \mid D \text{ is a } 2k\text{-edge-connected digraph of order } n\}$ . We show :

- (i)  $\frac{1}{2} \log(n - k + 1) \leq \text{sinv}'_k(n) \leq \log n + 4k - 3$  ;
- (ii) for any fixed positive integers  $k$  and  $t$ , deciding whether a given oriented graph  $\vec{G}$  satisfies  $\text{sinv}'_k(\vec{G}) \leq t$  (resp.  $\text{sinv}_k(\vec{G}) \leq t$ ) is NP-complete ;
- (iii) if  $T$  is a tournament of order at least  $2k + 1$ , then  $\text{sinv}'_k(T) \leq \text{sinv}_k(T) \leq 2k$ , and  $\text{sinv}'_k(T) \leq \frac{4}{3}k + o(k)$ ;
- (iv)  $\frac{1}{2} \log(2k + 1) \leq \text{sinv}'_k(T) \leq \text{sinv}_k(T)$  for some tournament  $T$  of order  $2k + 1$ ;
- (v) if  $T$  is a tournament of order at least  $19k - 2$  (resp.  $11k - 2$ ), then  $\text{sinv}'_k(T) \leq \text{sinv}_k(T) \leq 1$  (resp.  $\text{sinv}_k(T) \leq 3$ );
- (vi) for every  $\epsilon > 0$ , there exists  $C$  such that  $\text{sinv}'_k(T) \leq \text{sinv}_k(T) \leq C$  for every tournament  $T$  on at least  $2k + 1 + \epsilon k$  vertices.

**Keywords:** inversion; tournament;  $k$ -strong;  $k$ -arc-strong.

## 1 Introduction

Notation not given below is consistent with [BJG09]. We denote by  $[k]$  the set  $\{1, 2, \dots, k\}$ . A **digraph** may contain **digons**, that are, pairs of arcs in opposite direction between the same two vertices, but no parallel arcs or loops while an **oriented graph** is a digraph without digons.

A **feedback arc set** in a digraph is a set of arcs whose reversal results in an acyclic digraph. Finding a minimum cardinality feedback arc set is an important and heavily studied problem. It is one of the first problems shown to be NP-hard listed by Karp in [Kar72]. Furthermore, it is hard to approximate. For arbitrary digraphs, the best known ratio is  $\mathcal{O}(\log n \log \log n)$  due to Even et al. [ENSS95]. It was proven to be APX-hard by Kann [Kan92]. This was later strengthened by Dinur and Safra who showed that, unless P is equal to NP, it cannot be approximated within a factor of better than 1.36 [DS05]. For **tournaments**, which are orientations of complete graphs, the problem was proven to be remain NP-complete, independently by Alon [Alo06] and Charbit, Thomassé et Yeo [CTY07]. On the other hand, a polynomial-time approximation scheme was provided by Kenyon-Mathieu and Schudy [KMS07], improving on an earlier 3-approximation algorithm by Ailon Charikar and Newman [ACN08].

To make a digraph  $D$  acyclic, one can use a different operation from arc reversal, called inversion. The **inversion** of a set  $X$  of vertices consists in reversing the direction of all arcs of  $D\langle X \rangle$ . We say that we **invert**  $X$  in  $D$ . The

resulting digraph is denoted by  $\text{Inv}(D; X)$ . If  $(X_i)_{i \in I}$  is a family of subsets of  $V(D)$ , then  $\text{Inv}(D; (X_i)_{i \in I})$  is the digraph obtained after inverting the  $X_i$  one after another. Observe that this is independent of the order in which we invert the  $X_i$ :  $\text{Inv}(D; (X_i)_{i \in I})$  is obtained from  $D$  by reversing the arcs such that an odd number of the  $X_i$  contain its two endvertices. A **decycling family** of a digraph  $D$  is a family  $(X_i)_{i \in I}$  of subsets of  $V(D)$  such that  $\text{Inv}(D; (X_i)_{i \in I})$  is acyclic. A digraph admits a decycling family if and only if it does not contain any digon. Indeed, observe that an inversion operation changes the orientation of either none or both of the arcs in a digon. Hence, a digraph containing a digon cannot be made acyclic by inversions. Conversely, in an oriented graph, the pairs of endvertices of the arcs of a minimal feedback arc-set form a decycling family. The **inversion number** of an oriented graph  $D$ , denoted by  $\text{inv}(D)$ , is the minimum number of inversions needed to transform  $D$  into an acyclic oriented graph, that is, the minimum cardinality of a decycling family. It was first introduced by Belkhechine et al. in [BBBBP10] and then studied in several papers [BJdSH22, TSKP22, AHH<sup>+</sup>22, APS<sup>+</sup>22]. In particular, Belkhechine et al. in [BBBBP10] proved that, for any fixed integer  $k$ , deciding whether a given tournament has inversion number at most  $k$  is polynomial-time solvable. In contrast, Bang-Jensen et al. [BJdSH22] proved that deciding whether a given digraph has inversion number 1 is NP-complete. This was generalized by Alon et al. [APS<sup>+</sup>22]: for any fixed positive integer  $k$ , deciding whether a given digraph has inversion number at most  $k$  is NP-complete. The maximum  $\text{inv}(n)$  over all inversion numbers of digraphs of order  $n$  has also been investigated. Independently, Aubian et al. [AHH<sup>+</sup>22] and Alon et al. [APS<sup>+</sup>22] proved  $n - 2\sqrt{n \log n} \leq \text{inv}(n) \leq n - \lceil \log(n + 1) \rceil$ .

The main purpose of this article is to study the possibilities of applying the inversion operation to obtain a different objective than the resulting digraph being acyclic. Instead of making a digraph acyclic, we are interested in making it satisfy a prescribed connectivity property. A digraph  $D$  is **strongly connected** or simply **strong** (resp.  **$k$ -arc-strong** for some positive integer  $k$ ), if for any partition  $(V_1, V_2)$  of  $V(D)$  with  $V_1, V_2 \neq \emptyset$  there is an arc (resp. at least  $k$  arcs) with tail in  $V_1$  and head in  $V_2$ . Similarly, a multigraph  $G$  is **connected** (resp.  **$k$ -edge-connected**) for some positive integer  $k$ , if for any partition  $(V_1, V_2)$  of  $V(G)$  with  $V_1, V_2 \neq \emptyset$  there is an edge (resp. at least  $k$  edges) with one endvertex in  $V_1$  and one endvertex in  $V_2$ . We further say that  $G$  is  **$k$ -connected** if  $|V(G)| \geq k + 1$  and  $G - S$  is connected for every  $S \subseteq V(G)$  with  $|S| \leq k - 1$ . For a given digraph  $D$ , we denote by  $\text{UG}(D)$  the undirected (multi)graph that we obtain by suppressing the orientations of the arcs. A digraph is  **$k$ -connected** (resp.  **$k$ -edge-connected**) if its underlying (multi)graph is. Clearly, a digraph  $D$  can be made  $k$ -arc-strong by reversing some arcs if and only if the edges of  $\text{UG}(D)$  can be oriented such that the resulting digraph is  $k$ -arc-strong. Robbins' Theorem [Rob39] asserts that a graph admits a strong orientation if and only if it is 2-edge-connected, and more generally, Nash–Williams' weak orientation theorem [NW60], asserts that a graph admits a  $k$ -arc-strong orientation if and only if it is  $2k$ -edge-connected. We can hence decide in polynomial time whether a given digraph can be transformed into a  $k$ -arc-strong digraph by applying arc reversals, applying standard flow algorithms to its underlying graph. Furthermore, it is well-known that, if this is the case, by reducing to a minimum-cost submodular flow problem, one can determine, in polynomial time, a minimum set of arcs in  $D$  whose reversal gives a  $k$ -arc-strong digraph, see Section 8.8.4 of [BJG09] for details. It is easy to see that the number of necessary arc reversals to make a  $2k$ -edge-connected digraph  $D$   $k$ -arc-strong cannot be bounded by a function depending only on  $k$ . For example, one can consider digraphs that contain a number of sinks which is linear in the number of vertices of the graph, where a **sink** is a vertex with no outgoing arc. However, Bang-Jensen and Yeó [BJY04] proved that for tournaments, the size of such a set is always bounded by a quadratic function of  $k$ . Precisely, they showed that every tournament on at least  $2k + 1$  vertices can be made  $k$ -arc-strong by reversing at most  $\frac{1}{2}k(k - 1)$  arcs. This result is tight for the transitive tournaments.

We are interested in the problem of using inversions to make a digraph  $k$ -arc-strong. A  **$k$ -arc-strengthening family** of a digraph  $D$  is a family  $(X_i)_{i \in I}$  of subsets of  $V(D)$  such that  $\text{Inv}(D; (X_i)_{i \in I})$  is  $k$ -arc-strong. The  **$k$ -arc-strong inversion number** of a digraph  $D$ , denoted by  $\text{sinv}'_k(D)$ , is the minimum number of inversions needed to transform  $D$  into a  $k$ -arc-strong digraph, that is, the minimum cardinality of a  $k$ -strengthening family. We first deal with the extremal behaviour of  $\text{sinv}'_k(D)$  for some fixed  $k$ , that is, we deal with the question of finding the maximum number of necessary inversions to make a  $2k$ -edge-connected digraph on a fixed number of vertices  $k$ -arc-strong. To this end, we define  $\text{sinv}'_k(n) = \max\{\text{sinv}'_k(D) \mid D \text{ } 2k\text{-edge-connected digraph of order } n\}$ . It turns out that  $\text{sinv}'_k(n)$  is an unbounded, but slowly growing, function. We are able to determine  $\text{sinv}'_k(n)$  up to a constant multiplicative factor of roughly 2. More precisely, we show the following result:

**Theorem 1.1.** *For any pair of positive integers  $k, n$  with  $n \geq 2k + 1$ , we have*

$$\frac{1}{2} \log(n - k + 1) \leq \text{sinv}'_k(n) \leq \log n + 4k - 3.$$

Observe that the condition  $n \geq 2k + 1$  is necessary for  $\text{sinv}'_k(n)$  to be well-defined.

Next, we consider the algorithmic complexity of computing  $\text{sinv}'_k(D)$  algorithmically for a given digraph  $D$  and a fixed integer  $k$ . We show that this problem is NP-hard. More precisely, we show the following, slightly stronger, result.

**Theorem 1.2.** *Deciding whether a given oriented graph  $D$  satisfies  $\text{sinv}'_k(D) \leq t$  is NP-complete for all pairs of positive integers  $k$  and  $t$ .*

Furthermore we show that there is little hope to approximate  $\text{sinv}'_k(D)$  within a factor better than 2.

**Theorem 1.3.** *Unless  $P=NP$ , for any positive integer  $k$ , there is no  $(2 - \epsilon)$ -approximation algorithm for computing  $\text{sinv}'_k(D)$  for oriented graphs for any  $\epsilon > 0$ .*

As a related problem, one may also want to make a digraph  $k$ -strong. A digraph  $D$  is  $k$ -strong if  $|V(D)| \geq k + 1$  and for any set  $S \subseteq V(D)$  with less than  $k$  vertices,  $D - S$  is strong. A digraph which can be made  $k$ -strong by reversing arcs is  $k$ -strengthenable. The 1-strengthenable digraphs are the 2-edge-connected ones, because being 1-strong is equivalent to being strong or 1-arc-strong. Thomassen [Tho15] proved that the 2-strengthenable digraphs are the 4-edge-connected digraphs  $D$  such that  $D - v$  is 2-edge-connected for every vertex  $v \in V(D)$ . On the other hand, it is NP-hard to compute the minimum number of arc reversals needed to make a given digraph 2-strong, as shown by Bang-Jensen et al. in [BJHK23]. Furthermore, in contrast to the analogous problem for  $k$ -arc-strengthenable digraphs, for  $k \geq 3$ , it is NP-complete to decide whether a digraph is  $k$ -strengthenable. Indeed, Durand de Gevigney [Dur20] proved that it is NP-complete to decide whether an undirected graph has a  $k$ -strong orientation for any  $k \geq 3$ .

It is also natural to use inversions to make a digraph  $k$ -strong. A  $k$ -strengthening family of a digraph  $D$  is a family  $(X_i)_{i \in I}$  of subsets of  $V(D)$  such that  $\text{Inv}(D; (X_i)_{i \in I})$  is  $k$ -strong. The  $k$ -strong inversion number of a  $k$ -strengthenable digraph  $D$ , denoted by  $\text{sinv}_k(D)$ , is the minimum number of inversions needed to transform  $D$  into a  $k$ -strong digraph, that is, the minimum cardinality of a  $k$ -strengthening family. In the light of the complexity result of Durand de Gevigney [Dur20], it seems difficult to obtain an extremal result in the shape of Theorem 1.1 for  $k$ -strong digraphs. However, we give the following complexity results which are the natural analogues of Theorems 1.2 and 1.3.

**Theorem 1.4.** *Deciding whether a given  $k$ -strengthenable oriented graph  $D$  satisfies  $\text{sinv}_k(D) \leq t$  is NP-complete for all pairs of positive integers  $k$  and  $t$ .*

**Theorem 1.5.** *Unless  $P=NP$ , for any positive integer  $k$ , there is no  $(2 - \epsilon)$ -approximation algorithm for computing  $\text{sinv}_k(D)$  for oriented graphs for any  $\epsilon > 0$ .*

In the remainder of this article, we focus on a particular kind of oriented graphs, namely tournaments. It is not hard to show that every tournament of order at least  $2k + 1$  is  $k$ -strengthenable and that it can be made  $k$ -strong by reversing the orientation of at most  $\frac{1}{4}(4k - 2)(4k - 3)$  arcs, see e.g. [BJG09], p. 379.

Further, the following improvement was conjectured by Bang-Jensen. (See [BJG09].)

**Conjecture 1.6** (Bang-Jensen, 1994). *Every tournament on at least  $2k + 1$  vertices can be made  $k$ -strong by reversing at most  $\frac{1}{2}k(k + 1)$  arcs.*

It would be tight as shown by the transitive tournaments. Bang-Jensen, Johansen, and Yeo [BJJY23] proved this conjecture for tournaments of order at least  $3k - 1$ .

It is then natural to ask whether or not we can make a tournament  $k$ -strong or  $k$ -arc-strong using significantly less than  $\frac{1}{2}k(k + 1)$  inversions. This leads to consider  $M_k = \max\{\text{sinv}_k(T) \mid T \text{ tournament of order at least } 2k + 1\}$  and  $M'_k = \max\{\text{sinv}'_k(T) \mid T \text{ tournament of order at least } 2k + 1\}$ . We show that these numbers are indeed significantly smaller than  $\frac{1}{2}k(k + 1)$ , but cannot be bounded by a constant independent of  $k$ . More precisely, we show the following result.

**Theorem 1.7.** For every sufficiently large integer  $k$ , we have  $\frac{1}{2} \log(2k + 1) \leq M'_k \leq M_k \leq 2k$ .

The lower bound is obtained by a counting argument.

We also prove that  $M_1 = M'_1 = 1$  and  $M_2 = M'_2 = 2$  showing that the upper bound is not tight for  $k = 1, 2$ . We also prove a better upper for  $M'_k$  for large value of  $k$ .

**Theorem 1.8.**  $M'_k \leq \frac{4}{3}k + o(k)$ .

However, we believe that this bound, as well as the others, is not tight.

**Problem 1.9.** Find better bounds on  $M_k$  and  $M'_k$ .

We further study the question of how the parameters  $\text{sinv}_k$  and  $\text{sinv}'_k$  behave when fixing  $k$  and considering tournaments whose size is significantly larger than  $2k + 1$ . As a first result, we prove that every sufficiently large tournament can be made  $k$ -strong (and thus also  $k$ -arc-strong) in one inversion.

**Theorem 1.10.** If  $n \geq (2k - 1)2^{2k}$ , then  $\text{sinv}_k(T) \leq 1$ .

This leads us to the study of the functions  $m_k(n) = \max\{\text{sinv}_k(T) \mid T \text{ tournament of order } n\}$  and  $m'_k(n) = \max\{\text{sinv}'_k(T) \mid T \text{ tournament of order } n\}$  for all  $n \geq 2k + 1$ . We believe that  $m_k$  and  $m'_k$  have the following monotonic behaviour.

**Conjecture 1.11.**

- (i)  $m_k$  and  $m'_k$  are non-increasing mappings.
- (ii)  $M_k = M'_k = m_k(2k + 1) = m'_k(2k + 1)$ .

Note that (i) implies (ii). This conjecture is motivated by the fact that one can easily prove that  $m_k$  and  $m'_k$  are non-increasing for  $n > 4k - 2$ . Indeed, it is well-known that every tournament  $T$  of order  $n > 4k - 2$  has a vertex  $v$  with  $d_T^-(v) \geq k$  and  $d_T^+(v) \geq k$ . Moreover adding a vertex with in- and out-degree at least  $k$  to a  $k$ -(arc-)strong digraph results in a  $k$ -(arc-)strong digraph. Hence any  $k$ -(arc-)strengthening family of  $T - v$  is also a  $k$ -(arc-)strengthening family of  $T$ . Thus  $\text{sinv}_k(T) \leq \text{sinv}_k(T - v) \leq m_k(n - 1)$  and  $\text{sinv}'_k(T) \leq \text{sinv}'_k(T - v) \leq m'_k(n - 1)$ . Hence  $m_k(n) \leq m_k(n - 1)$  and  $m'_k(n) \leq m'_k(n - 1)$ . Therefore, in order to approach Problem 1.9, it is sufficient to consider tournaments whose order is in the range from  $2k + 1$  to  $4k + 2$ .

Theorem 1.10 implies that, for every pair of positive integers  $k$  and  $i$ , there is a smallest integer  $N_k(i)$  such that  $m_k(n) \leq i$  for all  $n \geq N_k(i)$ . Similarly, for every pair of positive integers  $k$  and  $i$ , there is a smallest integer  $N'_k(i)$  such that  $m'_k(n) \leq i$  for all  $n \geq N'_k(i)$ . Since  $m'_k(n) \leq m_k(n)$  for all  $k$  and  $n$ , we have  $N'_k(i) \leq N_k(i)$  for all  $k$  and  $i$ . It is natural to ask the following questions.

**Problem 1.12.** What is the minimum integer  $N_k(i)$  such that  $\text{sinv}_k(T) \leq i$  for every tournament  $T$  of order at least  $N_k(i)$  ?

What is the minimum integer  $N'_k(i)$  such that  $\text{sinv}'_k(T) \leq i$  for every tournament  $T$  of order at least  $N'_k(i)$  ?

Our most important result on Problem 1.12 is the following significant improvement on Theorem 1.10.

**Theorem 1.13.** For any positive integer  $k$ , we have  $N'_k(1) \leq N_k(1) \leq 19k - 2$ .

Concerning lower bounds for  $N_k(1)$ , we have the following result.

**Proposition 1.14.** For any positive integer  $k$ , we have  $N_k(1) \geq 5k - 2$ .

We further show that a significantly smaller tournament can still be dealt with three inversions.

**Theorem 1.15.** For any positive integer  $k$ , we have  $N'_k(3) \leq N_k(3) \leq 11k - 2$ .

Finally, using probabilistic methods, we manage to prove that every tournament that is a constant factor bigger than  $2k$  can be made  $k$ -strong by a constant number of inversions. More precisely, we prove the following result.

**Theorem 1.16.** *There exists a function  $f : \mathbb{R}_{>0} \rightarrow \mathbb{N}$  such that for every  $\epsilon > 0$  and every positive integer  $k$ , if  $T$  is an  $n$ -vertex tournament with  $n \geq 2k + 1 + \epsilon k$ , then  $\text{sinv}_k(T) \leq f(\epsilon)$ .*

The fact that  $m_k(n) = 1$  for  $n$  sufficiently large implies that the set  $\mathcal{F}_k$  of tournaments  $T$  such that  $\text{sinv}_k(T) > 1$  is finite. This implies that computing  $\text{sinv}_k$  and  $\text{sinv}'_k$  for fixed  $k$  can be done in polynomial time for tournaments.

**Corollary 1.17.** *Let  $k$  be a positive integer. We can compute  $\text{sinv}_k(T)$  (resp.  $\text{sinv}'_k(T)$ ) for a given tournament  $T$  on  $n$  vertices in  $O(n^{7/2})$  time (resp.  $O(n^2)$  time).*

*Proof.* We first check in constant time, whether  $T$  is in  $\mathcal{F}_k$ . If yes, then we can return  $\text{sinv}_k(T)$  and  $\text{sinv}'_k(T)$  that may be stored in some precomputed table. If not, then  $\text{sinv}'_k(T) \leq \text{sinv}_k(T) \leq 1$ . We then check in  $O(n^{7/2})$  using algorithms due to Even and Tarjan [ET75] or Galil [Gal80] based on flow (resp.  $O(n^2)$  using Mansour and Schieber [MS89] algorithm based on flow or Gabow's algorithm [Gab91] based on matroids) whether  $T$  is  $k$ -strong (resp.  $k$ -arc-strong) to determine whether  $\text{sinv}_k(T)$  (resp.  $\text{sinv}'_k(T)$ ) is equal to 0 or 1.  $\square$

This article is structured as follows. We give some notation and a collection of preliminary results in Section 2. In Section 3, we prove Theorem 1.1. In Section 4, we prove the complexity results, namely Theorems 1.2 to 1.5. The results on tournaments are contained in Section 5, in which we prove Theorem 1.7, and in Section 6, in which we prove Theorems 1.10, 1.13, 1.15 and 1.16 and Proposition 1.14.

## 2 Preliminaries

In Section 2.1, we introduce some formal notation and in Section 2.2, we give a collection of auxiliary results.

### 2.1 Notation

A **mixed graph**  $G = (V, E, A)$  is a triple consisting of a set  $V$  of elements called **vertices**, a set  $E$  of unordered pairs of vertices called **edges**, and a set  $A$  of ordered pairs of vertices called **arcs**. Hence a graph can be seen as a mixed graph whose arc set is empty, while a digraph can be seen as a mixed graph whose edge set is empty. Given a mixed graph  $G$ , its vertex set is denoted by  $V(G)$ , its edge set is denoted by  $E(G)$  and its arc set by  $A(G)$ . The **underlying graph**  $\text{UG}(G)$  of a mixed graph  $G$  is the undirected multigraph that we obtain by suppressing the orientations of the arcs (i.e. replacing each arc by an edge between its two endvertices). An **orientation** of a mixed graph  $G$  is a digraph obtained by giving an orientation to every edge, i.e. replacing each edge by one of the two arcs between its endvertices.

Let  $S$  be a set of vertices in a mixed graph  $G$ . We denote by  $\delta_G(S)$  the set of edges of  $E(G)$  with exactly one endvertex in  $S$ . We also denote by  $\delta_G^+(S)$  (resp.  $\delta_G^-(S)$ ) the set of arcs in  $A(G)$  with tail (resp. head) in  $S$  and head (resp. tail) out of  $S$ . The **degree** (resp. **out-degree**, **in-degree**) of  $S$  in  $G$  is  $d_G(S) = |\delta_G(S)|$  (resp.  $d_G^+(S) = |\delta_G^+(S)|$ ,  $d_G^-(S) = |\delta_G^-(S)|$ ). Let  $N_G(S)$  (resp.  $N_G^+(S)$ ,  $N_G^-(S)$ ) denote the set of vertices in  $V(G) \setminus S$  that are incident to at least one edge (arc) in  $\delta_G(S)$  (resp.  $\delta_G^+(S)$ ,  $\delta_G^-(S)$ ). The **degree** (resp. **out-degree**, **in-degree**) of a vertex  $v$  in  $G$  is  $d_G(v) = d_G(\{v\})$  (resp.  $d_G^+(v) = d_G^+(\{v\})$ ,  $d_G^-(v) = d_G^-(\{v\})$ ). Furthermore, for two disjoint sets  $S_1, S_2$ , let  $d_G(S_1, S_2)$  be the number of edges in  $E$  with one endvertex in  $S_1$  and one endvertex in  $S_2$ , let  $d_G^+(S_1, S_2)$  be the number of arcs in  $A$  whose tail is in  $S_1$  and whose head is in  $S_2$ , and let  $d_G^-(S_1, S_2)$  be  $d_G^+(S_2, S_1)$ . For simplicity, when the mixed graph is clear from the context, the subscript  $G$  is omitted.

Let  $D$  be a digraph. A **sink** (resp. **source**) in a digraph is a vertex with out-degree 0 (resp. in-degree 0). We say that  $D$  is **eulerian** if  $d_D^-(v) = d_D^+(v)$  for all vertex  $v$ . Let  $A$  and  $B$  be two sets in  $D$ . If  $ab$  is an arc for all pair  $(a, b)$  in  $A \times B$ , then we write  $A \Rightarrow B$ .

Let  $u, v$  two distinct vertices in  $D$ . The **strong-connectivity** from  $u$  to  $v$  in  $D$ , denoted by  $\kappa_D(u, v)$ , is the maximal number  $\alpha$  such that  $D - X$  contains a  $(u, v)$ -path for every  $X \subseteq V(D) \setminus \{u, v\}$  with  $|X| \leq \alpha - 1$ . For some  $S \subseteq V(D)$  and positive integer  $k$ , we say that  $S$  is  **$k$ -strong in  $D$**  if  $\kappa_D(u, v) \geq k$  for all  $u, v \in S$ .

## 2.2 Preliminary results

We use several times without explicitly mentioning it that for every positive integer  $k$ , there is a  $k$ -strong tournament on  $2k + 1$  vertices. One example for such a tournament is the so-called rotative tournament whose vertex set is  $[2k + 1]$  and where an arc is oriented from  $i$  to  $j$  if  $i - j \in [k](\text{mod } 2k + 1)$ .

We need the following simple result that allows us to extend a set which is  $k$ -strong in a digraph.

**Proposition 2.1.** *Let  $D$  be a digraph, let  $S$  be a  $k$ -strong set in  $D$  and let  $v \in V(D) \setminus S$ . If  $v$  has at least  $k$  in-neighbours in  $S$  and at least  $k$  out-neighbours in  $S$ , then  $S \cup \{v\}$  is  $k$ -strong in  $D$ .*

*Proof.* Suppose for the sake of a contradiction that there is a set  $X \subseteq V(D)$  of size at most  $k - 1$  and a pair  $(s, t)$  of vertices in  $S \cup \{v\} \setminus X$  such that there is no  $(s, t)$ -path in  $D - X$ . If  $s, t \in S$ , then this contradicts the fact that  $D$  is  $k$ -strong in  $S$ . Thus exactly one of  $s$  and  $t$  is  $v$ . Without loss of generality, suppose  $s \in S$  and  $t = v$ . Since  $|N^-(v) \cap S| \geq k$ ,  $v$  has an in-neighbour  $t_0 \in S - X$ , and as  $D$  is  $k$ -strong in  $D$ , there is an  $(s, t_0)$ -path in  $D - X$ , and so there is an  $(s, t)$ -path in  $D - X$ , a contradiction.  $\square$

The following result is helpful for applying Proposition 2.1.

**Proposition 2.2.** *Let  $T$  be a tournament on at least  $4k - 2$  vertices for some positive integer  $k$ . Then there exists some  $v \in V(T)$  with  $\min\{d_T^+(v), d_T^-(v)\} \geq k$ .*

We need the following simple characterization of eulerian tournaments.

**Proposition 2.3** (Folklore). *A tournament on  $2k + 1$  vertices is  $k$ -strong if and only if it is eulerian.*

We need the following orientation property of mixed graphs that can be found in [FF10].

**Proposition 2.4.** *Let  $H$  be a mixed graph whose underlying graph is eulerian. Then  $H$  has an eulerian orientation if and only if  $d_{E(H)}(S) \geq d_{A(H)}^+(S) - d_{A(H)}^-(S)$  for all  $S \subseteq V(H)$ .*

Finally, we state two basic tools from probability theory.

**Proposition 2.5** (Union Bound). *Let  $E_1, \dots, E_\ell$  be a set of events in a random experiment and  $E$  the event that at least one of  $E_1, \dots, E_\ell$  occurs. Then  $\Pr(E) \leq \sum_{i=1}^{\ell} \Pr(E_i)$ .*

**Proposition 2.6** (Chernoff's Bound). *If  $X$  is a random variable following a binomial law with parameters  $p \in [0, 1]$  and  $n \geq 0$ , then for every  $\epsilon \in [0, 1]$*

$$\Pr[X \leq (1 - \epsilon)pn] \leq \exp\left(-\frac{\epsilon^2}{2}pn\right).$$

We refer the reader to [MR02, Part II, Section 5] for an introduction to the probabilistic method, including proofs of Propositions 2.5 and 2.6.

## 3 Bounds on $\text{sinv}'_k(n)$

This section is dedicated to the extremal results we have on the  $k$ -arc-strong inversion number. In particular, we prove Theorem 1.1, which we restate.

**Theorem 1.1.** *For any pair of positive integers  $k, n$  with  $n \geq 2k + 1$ , we have*

$$\frac{1}{2} \log(n - k + 1) \leq \text{sinv}'_k(n) \leq \log n + 4k - 3.$$

Let  $d$  be a positive integer. A multigraph  $G$  is said to be  $d$ -**degenerate** if every submultigraph of  $G$  has a vertex of degree at most  $d$ . Every  $d$ -degenerate graph admits a  $d$ -**degenerate ordering**, that is an ordering  $(v_1, \dots, v_n)$  of the vertices of  $G$  such that every vertex has at most  $d$  neighbours with lower indices.

We shall need the following proposition, which, as observed in [BJHK23], is a direct consequence of Corollary 2 in [KS06].

**Proposition 3.1.** *Let  $G$  be a multigraph that has a  $k$ -arc-connected orientation for some positive integer  $k$  and let  $(e_1, f_1), \dots, (e_t, f_t)$  be a collection of pairwise disjoint pairs of parallel edges in  $G$ . Then  $G$  has a  $k$ -arc-connected orientation in which  $e_i$  and  $f_i$  are oriented in opposite directions for  $i = 1, \dots, t$ .*

The following result is part of a draft of a forthcoming paper by a group containing the authors.

**Theorem 3.2** ([IWG]). *Let  $G$  be an  $n$ -vertex  $d$ -degenerate graph  $G$ . For any two orientations  $\vec{G}_1, \vec{G}_2$  of  $G$ , one can transform  $\vec{G}_1$  into  $\vec{G}_2$  by inverting at most  $\log n + 2d - 1$  sets.*

For sake of completeness, we give the proof of this theorem. We denote by  $\mathbb{F}_2$  the field with two elements 0 and 1. Given two vectors  $\vec{x}, \vec{y} \in \mathbb{F}_2^t$ , we denote by  $\vec{x} \cdot \vec{y}$  their inner product, that is  $\sum_{i=1}^t \vec{x}_i \vec{y}_i \in \mathbb{F}_2$ .

*Proof.* Set  $t = \log n + 2d - 1$ . Let  $\vec{G}_1, \vec{G}_2$  be two orientations of  $G$ . For every edge  $uv \in E(G)$ , we define  $c(uv) \in \mathbb{F}_2$  such that  $c(uv) = 0$  if and only if  $uv$  is oriented the same in  $\vec{G}_1$  and in  $\vec{G}_2$ . Let  $(v^1, \dots, v^n)$  be a  $d$ -degenerate ordering of  $G$ . We will iteratively build vectors  $\vec{v}^1, \dots, \vec{v}^n \in \mathbb{F}_2^t$  such that, for every  $i \in [n]$ , the following conditions hold:

- $A^i$  : for every  $\ell > i$ , the vectors of  $\{\vec{v}^j \mid 1 \leq j \leq i \text{ and } v^j \in N(v^\ell)\}$  are linearly independent, and
- $B^i$  : for every edge  $v^j v^{j'}$  with  $j, j' \leq i$ ,  $\vec{v}^j \cdot \vec{v}^{j'} = c(v^j v^{j'})$ .

This will prove the theorem. Indeed, given such vectors, we can define the inversions  $X_i = \{u \mid \vec{u}_i = 1\}$  for every  $i \in \{1, \dots, t\}$ . Then the arc  $uv$  is reversed by  $X_1, \dots, X_t$  if and only if  $\vec{u} \cdot \vec{v} = 1$ . Hence Condition  $B^n$  implies that  $\text{Inv}(\vec{G}_1; X_1, \dots, X_t) = \vec{G}_2$ .

We now build this family of vectors as follows. For  $i = 1$ , we take  $\vec{v}^1 \in \mathbb{F}_2^t \setminus \{\vec{0}\}$  arbitrarily. Now suppose that we have  $\vec{v}^1, \dots, \vec{v}^i$  satisfying  $A^i$  and  $B^i$ . We shall now choose  $\vec{v}^{i+1}$ . Satisfying  $A^{i+1}$  forbids a vector space of dimension at most  $d - 1$  for each neighbour  $v^\ell$  of  $v^{i+1}$  with  $\ell > i + 1$ , because  $v^\ell$  has at most  $d - 1$  neighbours in  $\{v^1, \dots, v^i\}$  by  $d$ -degeneracy of  $G$ . So, in total, satisfying  $A^{i+1}$  forbids the union of those vector spaces, whose size is at most  $1 + n(2^{d-1} - 1)$ . Satisfying  $B^{i+1}$  allows at least  $2^{t-d}$  values for  $\vec{v}^{i+1}$ . Indeed, according to  $B^i$  the vectors of  $\{\vec{v}^j \mid 1 \leq j \leq i \text{ and } v^j \in N(v^{i+1})\}$  are linearly independent, so the affine space of the solutions of the system of equations  $\vec{v}^j \cdot \vec{v}^{i+1} = c(v^j v^{i+1})$  for the at most  $d$  indices  $j$  such that  $1 \leq j \leq i$  and  $v^j \in N(v^{i+1})$ , has dimension at least  $t - d$ . Since  $2^{t-d} = 2^{d-1+\log n} > 1 + n(2^{d-1} - 1)$ , there is a vector  $\vec{v}^{i+1} \in \mathbb{F}_2^t$  such that both  $A^{i+1}$  and  $B^{i+1}$  are satisfied. This concludes the proof of the theorem.  $\square$

In a multigraph  $G$ , for every pair  $(u, v)$  of vertices, a  $(u, v)$ -**cut** is a set  $F$  of edges such that every  $(u, v)$ -path in  $G$  intersects  $F$ . A **cut** is a  $(u, v)$ -cut for some  $u, v$ . Note that a multigraph is  $k$ -edge-connected if and only if for every pair  $u, v$  of vertices, there is no  $(u, v)$ -cut of size at most  $k - 1$  in  $G$ .

A multigraph  $G$  is **minimally  $k$ -edge-connected** if it is  $k$ -edge-connected and  $G \setminus e$  is not  $k$ -edge-connected for any edge  $e \in E(G)$ . Equivalently, every edge is in a cut of size exactly  $k$ .

**Lemma 3.3.** *Let  $k$  be a positive integer and let  $G$  be an  $n$ -vertex multigraph. If for every edge  $uv \in E(G)$  there is a  $(u, v)$ -cut of size at most  $k$  in  $G$ , then  $G$  has at most  $k(n - 1)$  edges.*

*Proof.* Let  $F_1, \dots, F_k$  be a set of  $k$  edge-disjoint spanning forests in  $G$  such that  $\sum_{i=1}^k |E(F_i)|$  is maximized. If there is an edge  $e = uv \in E(G) \setminus \bigcup_{i=1}^k E(F_i)$ , then, as  $e$  cannot be added to  $E(F_i)$ , we obtain that  $F_i$  contains a  $(u, v)$ -path for  $i \in [k]$ . Hence  $G \setminus e$  does not contain any  $(u, v)$ -cut whose size is smaller than  $k$ , a contradiction to the fact  $e$  is contained in a cut of size at most  $k$ . This yields  $|E(G)| = \sum_{i=1}^k |E(F_i)| \leq k(n - 1)$ .  $\square$

We are now ready to prove the upper bound in Theorem 1.1.

**Lemma 3.4.** *For every  $2k$ -edge-connected  $n$ -vertex digraph  $D$ ,  $\text{sinv}'_k(D) \leq \log n + 4k - 3$ .*

*Proof.* Without loss of generality, suppose that  $D$  is minimally  $2k$ -edge-connected. Then by Lemma 3.3, every subgraph of  $\text{UG}(D)$  has average degree smaller than  $2k$ . This implies that  $\text{UG}(D)$  is  $(2k - 1)$ -degenerate. Let  $D_0$  be the subdigraph obtained from  $D$  by removing all digons. By Proposition 3.1, there is an orientation  $D'_0$  of  $\text{UG}(D_0)$  such that together with the digons of  $D$ , this digraph is  $k$ -arc-strong. By Theorem 3.2, there is a set  $\mathcal{X}$  of at most  $\log n + 2(2k - 1) - 1$  inversions transforming  $D_0$  into  $D'_0$ . Since digons are preserved by inversions, we deduce that  $\text{Inv}(D, \mathcal{X})$  is  $k$ -arc-strong, and so  $\text{sinv}'_k(D) \leq \log n + 4k - 3$ .  $\square$



In order to prove the lower bound in Theorem 1.1, we first need the following intermediate result.

**Lemma 3.5.** *For every positive integer  $t$ , there exists a 2-edge-connected digraph  $D$  on  $2^{t-1} + 1 + \binom{2^{t-1}+1}{2}$  vertices with a specified vertex  $s$  such that in each digraph obtained from  $D$  by applying at most  $t - 1$  inversions contains a sink or source distinct from  $s$ .*

*Proof.* Let first  $S$  be a set of  $2^{t-1} + 1$  vertices. Now we obtain the digraph  $D$  by adding a vertex  $v_{\{s_1, s_2\}}$  as well as the arcs  $s_1 v_{\{s_1, s_2\}}$  and  $s_2 v_{\{s_1, s_2\}}$  for every  $\{s_1, s_2\} \subseteq S$ . It is easy to see that  $D$  is 2-edge-connected and that the number of vertices of  $D$  is  $2^{t-1} + 1 + \binom{2^{t-1}+1}{2}$ . Now let  $D'$  be obtained from  $D$  by inverting a collection of  $t - 1$  sets  $X_1, \dots, X_{t-1} \subseteq V(D)$ . As  $|S| = 2^{t-1} + 1$ , there exist vertices  $s_1, s_2 \in S$  such that for  $i = 1, \dots, t - 1$ , we have either  $\{s_1, s_2\} \subseteq X_i$  or  $\{s_1, s_2\} \cap X_i = \emptyset$ . Then in each of the  $t - 1$  inversions, either both or none of the arcs incident to  $v_{\{s_1, s_2\}}$  are inverted. We obtain that  $v_{\{s_1, s_2\}}$  is either a source or a sink in  $D'$ . The statement hence follows for an arbitrary  $s \in S$ .  $\square$

The next lemma gives a construction of graphs of arbitrary size that need a significant amount of inversions to become strong.

**Lemma 3.6.** *For every positive integer  $n \geq 3$ , there is a 2-edge-connected digraph  $D$  on  $n$  vertices such that any digraph obtained from  $D$  by applying at most  $\frac{1}{2} \lceil \log n \rceil - 1$  inversions contains a sink or a source.*

*Proof.* For  $n = 3, 4$ , the statement is clearly true. We may therefore suppose that  $n \geq 5$  and hence  $\frac{3}{2\sqrt{2}\sqrt{n}} \leq \frac{1}{2}$ . Let  $n' = 2^{\frac{1}{2} \lceil \log n \rceil - 1} + 1 + \binom{2^{\frac{1}{2} \lceil \log n \rceil - 1} + 1}{2}$ . By Lemma 3.5, there is a digraph  $D'$  on  $n'$  vertices together with a vertex  $s \in V(D')$  such that every graph obtained from  $D'$  by applying at most  $\frac{1}{2} \lceil \log n \rceil - 1$  inversions contains a sink or source distinct from  $s$ .

Next observe that

$$\begin{aligned}
n' &= 2^{\frac{1}{2} \lceil \log n \rceil - 1} + 1 + \binom{2^{\frac{1}{2} \lceil \log n \rceil - 1} + 1}{2} \\
&= 2^{\frac{1}{2} \lceil \log n \rceil - 1} + 1 + \frac{1}{2} (2^{\frac{1}{2} \lceil \log n \rceil - 1} + 1) 2^{\frac{1}{2} \lceil \log n \rceil - 1} \\
&\leq 2^{\frac{1}{2} \log n - \frac{1}{2}} + 1 + \frac{1}{2} (2^{\frac{1}{2} \log n - \frac{1}{2}} + 1) 2^{\frac{1}{2} \log n - \frac{1}{2}} \\
&= \frac{\sqrt{n}}{\sqrt{2}} + 1 + \frac{1}{2} \left( \frac{\sqrt{n}}{\sqrt{2}} + 1 \right) \frac{\sqrt{n}}{\sqrt{2}} \\
&= \frac{n}{4} + \frac{3}{2\sqrt{2}} \sqrt{n} + 1 \\
&= \frac{n}{4} + \frac{3}{2\sqrt{2}\sqrt{n}} n + 1 \\
&\leq \frac{n}{4} + \frac{n}{2} + \frac{n}{4} \\
&= n.
\end{aligned}$$

We now obtain  $D$  from  $D'$  by adding a new set of  $n - n'$  vertices and for each of them, adding a digon linking it to  $s$ . By construction,  $D$  has  $n$  vertices. Further, as the same property holds for  $D'$ , in any graph obtained from  $D$  by at most  $\frac{1}{2} \lceil \log n \rceil - 1$  inversions, one of the vertices in  $V(D') \setminus \{s\}$  is a source or a sink.  $\square$

We are now ready to prove the lower bound in Theorem 1.1.

**Lemma 3.7.** *For every pair of positive integers  $n, k$  with  $n \geq 2k + 1$ , there is a  $2k$ -edge-connected digraph  $D$  on  $n$  vertices such that  $\text{sinv}'_k(D) \geq \frac{1}{2} \log(n - k + 1)$ .*

*Proof.* By assumption, we have  $n - k + 1 \geq k + 2 \geq 3$ . Hence, by Lemma 3.6, there exists a 2-edge-connected digraph  $D_0$  on  $n - k + 1$  vertices such that any digraph obtained from  $D_0$  by inverting less than  $\frac{1}{2} \log(n - k + 1)$  sets has a sink or a source. Now let  $D$  be the digraph obtained from  $D_0$  by adding a set  $S$  of  $k - 1$  vertices, for any pair of vertices in  $S$  a digon linking them and for every vertex in  $S$  and every vertex in  $V(D_0)$  a digon linking these vertices. One can check that  $D$  has  $n$  vertices and is  $2k$ -edge-connected.

Consider now a family of subsets  $\mathcal{X}$  such that  $D' = \text{Inv}(D; \mathcal{X})$  is  $k$ -arc-strong. Any vertex  $v \in V(D_0)$  is linked to the vertices of  $S$  by digons in  $D$ . Since an inversion transforms a digon into a digon, it is also connected to the vertices of  $S$  by digons in  $D'$ . Hence, in  $D'$ ,  $v$  has at least one in-neighbour and one out-neighbour in  $V(D_0)$ . In other words, the subdigraph of  $D'$  induced by  $V(D_0)$  has no source and no sink. Thus  $|\mathcal{X}| \geq \frac{1}{2} \log(n - k + 1)$ .

Therefore  $\text{sinv}'_k(D) \geq \frac{1}{2} \log(n - k + 1)$ .  $\square$

Finally, Lemmas 3.4 and 3.7 imply Theorem 1.1.

## 4 Complexity of computing $\text{sinv}_k$ and $\text{sinv}'_k$

In this section, we deal with the complexity of computing the parameters  $\text{sinv}_k(D)$  and  $\text{sinv}'_k(D)$  for a given digraph  $D$ . More concretely, we prove Theorems 1.2, 1.3, 1.4, and 1.5.

In Subsection 4.1, we prove Theorems 1.2 and 1.4, for  $t = 1$  and every positive integer  $k$ . In Subsection 4.3 (resp. Subsection 4.2), we prove Theorem 1.4 (resp. Theorem 1.2), in the remaining cases, that is for every  $t \geq 2$  and every positive integer  $k$ . Finally, in Subsection 4.4 we prove Theorems 1.5 and 1.3. Note that all the studied decision problems are in NP since a  $k$ -strengthening (resp.  $k$ -arc-strengthening) family  $\mathcal{X}$  is a certificate as checking whether  $\text{Inv}(D; \mathcal{X})$  is  $k$ -strong (resp.  $k$ -arc-strong) can be done in polynomial time. Therefore we just need to prove their hardness.

### 4.1 One single inversion

In this section, we show hardness results for the case that a connectivity property is supposed to be achieved by a single inversion.

We need the following problem. For some integer  $k \geq 1$ , an instance of the MONOTONE EQUITABLE  $k$ -SAT problem ( $\text{ME}k\text{SAT}$ ) consists of a set of variables  $X$  and a set of clauses  $\mathcal{C}$  each of which contains exactly  $2k + 1$  nonnegated variables and the question is whether there is a truth assignment  $\phi : X \rightarrow \{\text{true}, \text{false}\}$  such that every clause in  $\mathcal{C}$  contains at least  $k$  true and  $k$  false variables with respect to  $\phi$ .

**Proposition 4.1.**  *$\text{ME}k\text{SAT}$  is NP-hard for every  $k \geq 1$ .*

*Proof.* For  $k = 1$ , we exactly have the MNAE3SAT problem, which is well-known to be NP-hard, see for example [Sch78]. We now proceed by induction. We fix some integer  $k \geq 1$ , assume that  $\text{ME}k\text{SAT}$  is NP-hard and show through a reduction from  $\text{ME}k\text{SAT}$  that so is  $\text{ME}(k + 1)\text{SAT}$ .

Let  $(X, \mathcal{C})$  be an instance of  $\text{ME}k\text{SAT}$ . We now create an instance  $(X', \mathcal{C}')$  of  $\text{ME}(k + 1)\text{SAT}$ . First, let  $Y = \{y_1, \dots, y_{k+3}\}$  and  $Z = \{z_1, \dots, z_{k+3}\}$  be sets of new variables. Now let  $\mathcal{C}_1$  the the set of all clauses  $C \subseteq Y \cup Z$  with  $|C| = 2k + 3$  and  $k + 1 \leq |C \cap Y| \leq k + 2$ . Moreover, we define  $\mathcal{C}_2 = \{C \cup \{y_1, z_1\} \mid C \in \mathcal{C}\}$ . Finally, we set  $X' = X \cup Y \cup Z$  and  $\mathcal{C}' = \mathcal{C}_1 \cup \mathcal{C}_2$ .

Before giving the main proof that our reduction works indeed, we show the following intermediate result.

**Claim 4.1.1.** *For any mapping  $\phi : Y \cup Z \rightarrow \{\text{true}, \text{false}\}$  that satisfies  $(Y \cup Z, \mathcal{C}_1)$ , we have  $\phi(y_1) = \dots = \phi(y_{k+3}) \neq \phi(z_1) = \dots = \phi(z_{k+3})$ .*

*Proof of claim.* First suppose that both  $Y$  and  $Z$  contain both true and false variables with respect to  $\phi$ , say  $\phi(y_1) = \phi(z_1) = \text{true}$  and  $\phi(y_2) = \phi(z_2) = \text{false}$ . Let  $\mathcal{C}_1 = Y \cup Z \setminus \{y_2, z_2, z_3\}$  and  $\mathcal{C}_2 = Y \cup X \setminus \{y_1, z_1, z_3\}$ . Further, let  $\alpha_1, \alpha_2$  be the number of true variables with respect to  $\phi$  in  $\mathcal{C}_1$  and  $\mathcal{C}_2$ , respectively. As  $\mathcal{C}_1, \mathcal{C}_2 \in \mathcal{C}_1$ , we have  $k + 1 \leq \alpha_i \leq k + 2$  for  $i = 1, 2$ . On the other hand, by assumption, we have  $\alpha_2 = \alpha_1 - 2$ . This yields  $k + 1 \leq \alpha_2 = \alpha_1 - 2 \leq (k + 2) - 2 = k$ , a contradiction.

We may hence suppose by symmetry that  $\phi(y_1) = \dots = \phi(y_{k+3}) = \text{true}$ . Suppose that  $Z$  contains a true variable with respect to  $\phi$ , say  $\phi(z_1) = \text{true}$ . Then consider the clause  $C = Y \cup Z \setminus \{y_1, z_2, z_3\}$ . We obtain that  $C$  contains at most  $k$  negative variables with respect to  $\phi$ , a contradiction, as  $C \in \mathcal{C}_1$ . This proves the claim.  $\diamond$

We now show that  $(X', \mathcal{C}')$  is a positive instance of  $\text{ME}(k+1)\text{SAT}$  if and only if  $(X, \mathcal{C})$  is a positive instance of  $\text{ME}k\text{SAT}$ .

First suppose that  $(X, \mathcal{C})$  is a positive instance of  $\text{ME}k\text{SAT}$ , so there is an assignment  $\phi : X \rightarrow \{\text{true}, \text{false}\}$  such that every clause in  $\mathcal{C}$  contains at least  $k$  true and at least  $k$  false variables with respect to  $\phi$ . Define  $\phi' : X' \rightarrow \{\text{true}, \text{false}\}$  by  $\phi'(x) = \phi(x)$  for all  $x \in X$  and  $\phi'(y_i) = \text{true}$  and  $\phi'(z_i) = \text{false}$  for all  $i \in [k+3]$ . By construction, every clause in  $\mathcal{C}_1$  contains at least  $k+1$  true and at least  $k+1$  false variables with respect to  $\phi'$ . Further, for every clause  $C \cup \{y_1, z_1\} \in \mathcal{C}_2$ , the clause  $C$  contains at least  $k$  true and at least  $k$  false variables with respect to  $\phi$ . As  $\phi'(y_1) = \text{true}$  and  $\phi'(z_1) = \text{false}$ , we obtain that  $C \cup \{y_1, z_1\}$  contains at least  $k+1$  true and at least  $k+1$  false variables with respect to  $\phi'$ . Hence  $(X', \mathcal{C}')$  is a positive instance of  $\text{ME}(k+1)\text{SAT}$ .

Now suppose that  $(X', \mathcal{C}')$  is a positive instance of  $\text{ME}(k+1)\text{SAT}$ , so there is an assignment  $\phi' : X' \rightarrow \{\text{true}, \text{false}\}$  such that every clause in  $\mathcal{C}'$  contains at least  $k+1$  true and at least  $k+1$  false variables with respect to  $\phi'$ . By Claim 4.1.1, we have  $\phi'(y_1) = \dots = \phi'(y_{k+3}) \neq \phi'(z_1) = \dots = \phi'(z_{k+3})$ .

By symmetry and Claim 4.1.1, we may suppose that  $\phi'(y_i) = \text{true}$  and  $\phi'(z_i) = \text{false}$  for all  $i \in [k+3]$ . Let  $\phi$  be the restriction of  $\phi'$  to  $X$ . For every  $C \in \mathcal{C}$ , the clause  $C \cup \{y_1, z_1\}$  contains at least  $k+1$  true and at least  $k+1$  false variables with respect to  $\phi'$ . As  $\phi'(y_1) = \text{true}$  and  $\phi'(z_1) = \text{false}$ , we obtain that  $C$  contains at least  $k$  true and at least  $k$  false variables with respect to  $\phi$ . Hence  $(X, \mathcal{C})$  is a positive instance of  $\text{ME}k\text{SAT}$ . This proves the lemma.  $\square$

We can now prove the following results.

**Theorem 4.2.** *Deciding whether  $\text{sinv}'_k(D) \leq 1$  for a given oriented graph  $D$  is NP-complete.*

**Theorem 4.3.** *Deciding whether  $\text{sinv}_k(D) \leq 1$  for a given  $k$ -strengthenable oriented graph  $D$  is NP-complete.*

The proof of Theorems 4.2 and 4.3 is a reduction from  $\text{ME}k\text{SAT}$  which is NP-hard by Proposition 4.1. Fix some  $k \geq 2$  and let  $(X, \mathcal{C})$  be an instance of  $\text{ME}k\text{SAT}$ . We construct an oriented graph  $D(X, \mathcal{C})$  as follows. We start from the disjoint union of a  $k$ -strong tournament  $T$  with vertex set  $\{s_1, \dots, s_{2k+1}\}$  and two stable sets  $V = \{v_x \mid x \in X\}$  and  $W = \{w_C \mid C \in \mathcal{C}\}$ . Next, for every  $i \in [k]$  and every  $x \in X$ , we add an arc from  $s_i$  to  $v_x$  and for every  $i = \{k+1, \dots, 2k\}$  and every  $x \in X$ , we add an arc from  $v_x$  to  $s_i$ . Finally, for every  $x \in X$  and  $C \in \mathcal{C}$  with  $x \in C$ , we add an arc from  $v_x$  to  $w_C$ .

Observe that the digraph  $D(X, \mathcal{C})$  is  $k$ -strengthenable. Indeed  $D(X, \mathcal{C}) \langle S \rangle$  is  $k$ -strong and hence  $S$  is  $k$ -strong in  $D(X, \mathcal{C})$ . Hence by Proposition 2.1,  $S \cup V$  is  $k$ -strong in  $D(X, \mathcal{C})$ . Now, for every  $C \in \mathcal{C}$ , we reverse  $k$  arcs between  $w_C$  and its  $2k+1$  neighbours in  $V$ . Then each  $w_C$  has at least  $k$  in-neighbours and  $k$ -out-neighbours in  $S \cup V$ . So, by Proposition 2.1 again, the resulting digraph is  $k$ -strong.

The main technical part of the reduction is contained in the proof of the following result.

**Claim 4.3.1.** *The following are equivalent:*

- (i)  $(X, \mathcal{C})$  is a positive instance of  $\text{ME}k\text{SAT}$ ;
- (ii)  $D(X, \mathcal{C})$  can be made  $k$ -strong by inverting a single set  $Z$ ;
- (iii)  $D(X, \mathcal{C})$  can be made  $k$ -arc-strong by inverting a single set  $Z$ .

*Proof of claim.* Set  $D = D(X, \mathcal{C})$ .

Suppose that (i) holds, so there is an assignment  $\phi : X \rightarrow \{\text{true}, \text{false}\}$  such that every clause in  $\mathcal{C}$  contains at least  $k$  true and at least  $k$  false variables with respect to  $\phi$ . Let  $Z = W \cup \{v_x \mid x \in X, \phi(x) = \text{true}\}$  and let  $D'$  be the oriented graph obtained from  $D$  by inverting  $Z$ . Clearly,  $S$  is  $k$ -strong in  $D'$ . Next, by Lemma 2.1, we obtain that  $S \cup V$  is  $k$ -strong in  $D'$ . Now consider some  $C \in \mathcal{C}$ . As  $\phi$  is equitable,  $C$  contains at least  $k$  variables  $x$  which satisfy  $\phi(x) = \text{true}$ . Hence, by construction,  $w_C$  has at least  $k$  out-neighbours in  $S \cup V$  in  $D'$ . Further, as  $\phi$  is equitable,  $C$  contains at least  $k$  variables  $x$  which satisfy  $\phi(x) = \text{false}$ . Hence, by construction,  $w_C$  has at least  $k$  in-neighbours in  $S \cup V$  in  $D'$ . By Lemma 2.1,  $D'$  is  $k$ -strong, so (ii) is satisfied.

Clearly, (ii) implies (iii).

Now suppose that (iii) holds, so  $D$  can be transformed into a  $k$ -arc-strong oriented graph  $D'$  by inverting a single set  $Z$ . We now define an assignment  $\phi : X \rightarrow \{\text{true}, \text{false}\}$  in the following way: if  $v_x \in Z$ , we set  $\phi(x) = \text{true}$ , otherwise we set  $\phi(x) = \text{false}$ . In order to prove that  $\phi$  has the desired properties, consider some clause  $C \in \mathcal{C}$ . As  $D'$  is  $k$ -arc-strong,  $w_C$  has at least  $k$  outgoing arcs in  $D'$ . It follows that  $w_C \in Z$  and there are at least  $k$  variables in  $X$  with  $x \in C$  and  $v_x \in Z$ . Hence  $C$  contains at least  $k$  variables  $x$  with  $\phi(x) = \text{true}$ . Next, as  $D'$  is  $k$ -arc-strong,  $w_C$  has at least  $k$  incoming arcs in  $D'$ . As  $w_C \in Z$ , it follows that there are at least  $k$  variables in  $X$  with  $x \in C$  and  $v_x \notin Z$ . Hence  $C$  contains at least  $k$  variables  $x$  with  $\phi(x) = \text{false}$ . We obtain that  $\phi$  has the desired properties, so (iii) is satisfied.  $\diamond$

Clearly, the size of  $D(X, \mathcal{C})$  is polynomial in the size of  $(X, \mathcal{C})$  for fixed  $k$ , and  $D(X, \mathcal{C})$  can be constructed in polynomial time from  $(X, \mathcal{C})$ . Hence Claim 4.3.1 implies Theorems 4.2 and 4.3.

## 4.2 Several inversions to become $k$ -arc-strong

In order to prove Theorems 1.4 and 1.2 for  $t \geq 2$  and all  $k$ , we shall use another well-studied graph parameter. Let  $G$  be a graph. A **cut cover** of  $G$  is a collection  $X_1, \dots, X_t$  of subsets of  $V(G)$  such that  $\bigcup_{i=1}^t \delta_G(X_i) = E(G)$ . The **cut covering number** of  $G$ , denoted by  $\text{cc}(G)$ , is the minimum integer  $t$  such that there is a cut cover of size  $t$ . The following well-known result shows a close relationship between the cut covering number and the classical chromatic number.

**Proposition 4.4.** *For any graph  $G$ , we have  $\text{cc}(G) = \lceil \log(\chi(G)) \rceil$ .*

It is well-known that the problem of deciding whether a given graph can be coloured with  $t$  colours is NP-hard for any  $t \geq 3$ , see [Sch78]. It hence follows from Proposition 4.4 that deciding whether the cut covering number of a given graph is at most  $t$  is NP-hard for any integer  $t \geq 2$ . We show a reduction from this problem to the one of deciding whether  $\text{sinv}'_k(D) \leq t$  for a given oriented graph  $D$ .

Given a digraph  $D$  and a set  $X$  of vertices in  $D$ , we denote by  $\partial_D(X)$  the set of edges of  $\text{UG}(D)$  with exactly one endvertex in  $X$ :  $\partial_D(X) = \delta_{\text{UG}(D)}(X)$ . The main technical part of the reduction is contained in the following Lemma that will be reused in Section 4.4.0

**Lemma 4.5.** *Given a graph  $G$  and a positive integer  $k$ , one can construct in polynomial time some oriented graph  $D$  such that  $\text{sinv}'_k(D) = \text{cc}(G)$  and  $|V(D)| = |V(G)| + (2k + 1)|E(G)| + 2k + 1$ .*

*Proof.* Let  $<$  be an arbitrary ordering on  $V(G)$ . We let  $V(D)$  contain  $V(G)$ , a set of vertices  $Z_e = \{z_e^1, \dots, z_e^{2k+1}\}$  for every  $e \in E(G)$  and a set of vertices  $X = \{x_1, \dots, x_{2k+1}\}$ . We add arcs to  $D$  such that  $D\langle X \rangle$  is a  $k$ -arc-strong tournament. We let  $A(D)$  contain an arc from  $v$  to  $x_i$  for every  $v \in V(G)$  and every  $i = 1, \dots, k$  and an arc from  $x_i$  to  $v$  for every  $v \in V(G)$  and every  $i = k + 1, \dots, 2k$ . Next for every  $e \in E(G)$ , we add arcs to  $D$  such that  $D\langle Z_e \rangle$  is a  $k$ -arc-strong tournament. Finally, for every  $e = uv \in E(G)$  with  $u < v$ , we add the arcs  $uz_e^i$  for  $i = 1, \dots, k$ , the arcs  $z_e^i v$  for  $i = 2, \dots, k$  and the arc  $vz_e^1$ . This finishes the description of  $D$ . Observe that  $|V(D)| = |V(G)| + (2k + 1)|E(G)| + 2k + 1$ .

We now show that  $\text{sinv}'_k(D) = \text{cc}(G)$ .

We first show that  $\text{sinv}'_k(D) \leq \text{cc}(G)$ . Let  $(X_1, \dots, X_{\text{cc}(G)})$  be an optimal cut cover of  $G$ . Now for every  $e = uv \in E(G)$ , there is a smallest  $\alpha(e)$  such that exactly one of  $u$  and  $v$  is contained in  $X_{\alpha(e)}$ . Now for  $i \in [\text{cc}(G)]$ , let  $X'_i = X_i \cup \{z_e^1 : \alpha(e) = i\}$ . Let  $D'$  be the digraph obtained from  $D$  by inverting  $X'_1, \dots, X'_{\text{cc}(G)}$ . We will show that  $D'$  is  $k$ -arc-strong. Let  $S \subseteq V(D)$  and suppose that  $\min\{d_{D'}^-(S), d_{D'}^+(S)\} < k$ . By symmetry, we may suppose that  $S \cap X \neq \emptyset$ . As  $D'\langle X \rangle = D\langle X \rangle$  is  $k$ -arc-strong, we obtain  $X \subseteq S$ . Next observe that in  $D'$  every  $v \in V(G)$  is incident to at least  $k$  arcs coming from  $X$  and  $k$  arcs going to  $X$ . Thus  $V(G) \subseteq S$ . Now consider some  $e = uv \in E(G)$  with  $u < v$ . As  $D'$  by construction contains  $k$  arcs from  $V(G)$  to  $Z_e$  and  $k$  arcs from  $Z_e$  to  $V(G)$ , we obtain that  $Z_e \subseteq S$ . Observe that either  $vz_e^1, z_e^1 u \in A(D')$  (if  $u \in X_{\alpha(e)}$ ) or  $z_e^1 v, uz_e^1 \in A(D')$  (if  $v \in X_{\alpha(e)}$ ). Thus  $D'$  contains  $k$  arcs from  $V(G)$  to  $Z_e$  and  $k$  arcs from  $Z_e$  to  $V(G)$ , and we obtain that  $Z_e \subseteq S$ . This yields  $S = V(D')$ , so  $D'$  is  $k$ -arc-strong. We hence have  $\text{sinv}'_k(D) \leq \text{cc}(G)$ .

We now show that  $\text{cc}(G) \leq \text{sinv}'_k(D)$ . Let  $(X_1, \dots, X_t)$  be a smallest collection of sets whose inversion makes  $D$   $k$ -arc-strong. Let  $D' = \text{Inv}(D; (X_i)_{i \in [t]})$  and for  $i \in [t]$ , let  $X'_i = X_i \cap V(G)$ .

**Claim 4.5.1.** *Let  $e = uv \in E(G)$ . Then there is some  $i \in [t]$  such that  $X'_i$  contains exactly one of  $u$  and  $v$ .*

*Proof of claim.* Suppose otherwise. As  $D'$  is  $k$ -arc-strong and  $d_D(Z_e) = 2k$ , we have  $d_{D'}^-(Z_e) = d_{D'}^+(Z_e) = k$ . For every  $i \in [t]$  either both  $u$  and  $v$  or none of  $u$  and  $v$  are contained in  $X_i$ . Thus we deduce that for every  $j \in [k]$ , either both the arcs corresponding to edges in  $\partial_D(Z_e) \cap \partial_D(z_e^j)$  are inverted or none of them are. Hence, in  $D'$ , there is exactly one arc entering  $Z_e$  incident to  $z_e^j$  for  $j = 2, \dots, k$ , and there are either zero or two arcs entering  $Z_e$  incident to  $z_e^1$ . Thus  $d_{D'}^-(Z_e) \in \{k-1, k+1\}$ , a contradiction.  $\diamond$

By Claim 4.5.1, we obtain that  $(X'_1, \dots, X'_t)$  is a cut cover for  $G$ . Hence  $\text{cc}(G) \leq t = \text{sinv}'_k(D)$ .  $\square$

Lemma 4.5 and the NP-hardness of deciding whether a given graph  $G$  satisfies  $\text{cc}(G) \leq t$  for all  $t \geq 2$  directly imply the following.

**Corollary 4.6.** *Let  $t$  be an integer greater than 1 and  $k$  a positive integer. Deciding whether  $\text{sinv}'_k(D) \leq t$  for a given oriented graph  $D$  is NP-complete.*

Clearly, Theorem 4.2 and Corollary 4.6 imply Theorem 1.2.

### 4.3 Several inversions to become $k$ -strong

In this subsection, we give a reduction from the cut covering problem to the problem of deciding whether  $\text{sinv}_k(D) \leq t$  for a given oriented graph  $D$ . The main technical part of the reduction is contained in the following lemma that will be reused in Section 4.4.

**Lemma 4.7.** *Given a graph  $G$  and a positive integer  $k$ , one can construct in polynomial time a  $k$ -strengthenable oriented graph  $D$  such that  $\text{sinv}_k(D) = \text{cc}(G)$  and  $|V(D)| = |V(G)| + (2k+1)|E(G)| + 2k+1$ .*

*Proof.* We let  $V(D)$  contain  $V(G)$ , a set of vertices  $Z_e = \{z_e^1, \dots, z_e^{2k+1}\}$  for every  $e \in E(G)$  and a set of vertices  $W = \{w_1, \dots, w_{2k+1}\}$ . We add arcs to  $D$  such that  $D\langle W \rangle$  is a  $k$ -strong tournament. We let  $A(D)$  contain an arc from  $v$  to  $w_i$  for every  $v \in V(G)$  and every  $i \in [k]$  and an arc from  $w_i$  to  $v$  for every  $v \in V(G)$  and every  $i \in \{k+1, \dots, 2k\}$ . Next for every  $e \in E(G)$ , we add arcs to  $D$  such that  $D\langle Z_e \rangle$  is a  $k$ -strong tournament. Further, for every  $e \in E(G)$ ,  $i \in [k]$  and  $j \in [k-1]$ , we add an arc from  $z_e^i$  to  $w_j$  and for every  $e \in E(G)$ ,  $i \in \{k+1, \dots, 2k\}$  and  $j \in [k-1]$ , we add an arc from  $w_j$  to  $z_e^i$ . Finally, for every  $e = uv \in E(G)$ , we add the arcs  $uz_e^1$  and  $vz_e^1$ . This finishes the description of  $D$ . Observe that  $|V(D)| = |V(G)| + (2k+1)|E(G)| + 2k+1$ .

We now show that  $\text{sinv}_k(D) = \text{cc}(G)$ . Note that this implies that  $\text{sinv}_k(D)$  is finite and thus  $D$  is  $k$ -strengthenable.

We first show that  $\text{sinv}_k(D) \leq \text{cc}(G)$ . Let  $(X_1, \dots, X_t)$  be an optimal cut cover of  $G$ . For every  $e = uv \in E(G)$ , there is a smallest integer  $\alpha(e)$  such that exactly one of  $u$  and  $v$  is contained in  $X_{\alpha(e)}$ . Now for every  $i \in [t]$ , let  $X'_i = X_i \cup \{z_e^1 : \alpha(e) = i\}$ . Let  $D'$  be the digraph obtained from  $D$  by inverting  $X'_1, \dots, X'_t$ . We will show that  $D'$  is  $k$ -strong. Let  $Y \subseteq V(D)$  with  $|Y| \leq k-1$ . We need to show that  $D - Y$  is strongly connected. As  $D\langle W \rangle$  is  $k$ -strong, we obtain that  $W \setminus Y$  is contained in a single strongly connected component  $S$  of  $D' - Y$ . Next observe that every  $v \in V(G) \setminus Y$  has an in-neighbour and an out-neighbour in  $S$ , so  $V(G) \setminus Y \subseteq S$ . Now consider some  $e = uv \in E(G)$ . As  $D'\langle Z_e \rangle = D\langle Z_e \rangle$  is  $k$ -strong, we obtain that  $Z_e \setminus Y$  is contained in a single strongly connected component of  $D' - Y$ . If  $Y \neq \{w_1, \dots, w_{k-1}\}$ , then there is at least one arc from  $Z_e \setminus Y$  to  $W \setminus Y$  and there is at least one arc from  $W \setminus Y$  to  $Z_e \setminus Y$ , so  $Z_e \setminus Y \subseteq S$ . Finally, if  $Y = \{w_1, \dots, w_{k-1}\}$ , observe that, by construction, either  $D'$  contains the arcs  $uz_e^1$  and  $z_e^1v$  or the arcs  $vz_e^1$  and  $z_e^1u$ . Again, we obtain  $Z_e \setminus Y \subseteq S$ . This yields that  $D'$  is  $k$ -strong. We hence have  $\text{sinv}_k(D) \leq t = \text{cc}(G)$ .

We now show that  $\text{cc}(G) \leq \text{sinv}_k(D)$ . Let  $(X_1, \dots, X_t)$  be a smallest collection of sets whose inversion makes  $D$   $k$ -strong. For  $i = 1, \dots, t$ , let  $X'_i = X_i \cap V(G)$ .

**Claim 4.7.1.** *Let  $e = uv \in E(G)$ . Then there is some  $i \in \{1, \dots, t\}$  such that  $X_i$  contains exactly one of  $u$  and  $v$ .*

*Proof of claim.* Suppose otherwise, that is for every  $i \in [t]$  either both  $u$  and  $v$  or none of  $u$  and  $v$  are in  $X_i$ . Thus both the arcs corresponding to edges in  $\partial_D(Z_e) \cap \partial_D(z_e^1)$  are inverted or none of them. This yields that in  $D' - \{w_1, \dots, w_{k-1}\}$ , there are either zero or two arcs entering  $Z_e$ . Hence  $D' - \{w_1, \dots, w_{k-1}\}$  is not strongly connected, a contradiction.  $\diamond$

By Claim 4.7.1, we obtain that  $(X'_1, \dots, X'_t)$  is a cut cover for  $G$ . Thus  $\text{cc}(G) \leq t = \text{sinv}_k(D)$ , and we conclude that  $\text{cc}(G) = \text{sinv}_k(D)$ .  $\square$

Lemma 4.7 and the NP-hardness of deciding whether a given graph  $G$  satisfies  $\text{cc}(G) \leq t$  for all  $t \geq 2$  directly imply the following.

**Corollary 4.8.** *Let  $t$  be an integer greater than 1 and  $k$  a positive integer. Deciding whether  $\text{sinv}_k(D) \leq t$  for a given oriented graph  $D$  is NP-complete.*

Clearly, Theorem 4.3 and Corollary 4.8 imply Theorem 1.2.

## 4.4 Unapproximability

In order to prove Theorems 1.3 and 1.5, we need the following result of Zuckerman [Zuc07].

**Proposition 4.9** ([Zuc07]). *Unless  $P=NP$ , there is no polynomial-time  $\mathcal{O}(n^{1-\epsilon})$ -approximation algorithm for computing  $\chi(G)$  for any  $\epsilon > 0$ .*

As a consequence, we obtain a negative result concerning the approximation of the cut covering number.

**Proposition 4.10.** *Unless  $P=NP$ , there is no polynomial-time  $(2 - \epsilon)$ -approximation algorithm for computing the cut covering number of a given graph for any  $\epsilon > 0$ .*

*Proof.* Suppose that there exists a polynomial-time  $(2 - \epsilon)$ -approximation algorithm  $A$  for computing the cut covering number of a given graph for some  $\epsilon > 0$ . Now choose  $\epsilon'$  with  $0 < \epsilon' < \epsilon$  and consider a graph  $G$  on  $n \geq 4^{\frac{1}{\epsilon - \epsilon'}}$  vertices. Let  $A'$  be the algorithm that takes  $G$  as input and returns  $2^\alpha$  where  $\alpha$  is the output of  $A$  applied to  $G$ . Clearly, as  $A$  runs in polynomial time, so does  $A'$ . Further, by Proposition 4.4, we have  $2^\alpha \geq 2^{\text{cc}(G)} \geq 2^{\log(\chi(G))} = \chi(G)$ . Finally, by Proposition 4.4, we have

$$\begin{aligned} 2^\alpha &\leq 2^{(2-\epsilon)\lceil \log(\chi(G)) \rceil} \\ &\leq 2^{(2-\epsilon)(\log(\chi(G))+1)} \\ &= 2^{2-\epsilon} 2^{\log(\chi(G))(1-\epsilon)} 2^{\log(\chi(G))} \\ &\leq 4\chi(G)^{1-\epsilon} \chi(G) \\ &\leq 4n^{1-\epsilon} \chi(G) \\ &\leq n^{1-\epsilon'} \chi(G). \end{aligned}$$

Hence  $A'$  is a  $n^{1-\epsilon'}$ -approximation algorithm for the chromatic number, so the statement follows by Proposition 4.9.  $\square$

Now Proposition 4.10 and Lemma 4.5 imply Theorem 1.3, and Proposition 4.10 and Lemma 4.7 imply Theorem 1.5.

## 5 Bounds on $M'_k$ and $M_k$

In this section, we prove Theorem 1.7, that is  $\frac{1}{2} \log(2k+1) \leq M_k \leq 2k$  for all sufficiently large  $k$ . The left hand-side inequality is proved in Theorem 5.2, and the right-hand side one in Theorem 5.8.

## 5.1 Lower bound on $m'_k(2k+1)$

We shall first show the lower bound  $m'_k(2k+1) = \Omega(\log k)$ . We need the following result.

**Theorem 5.1** (McKay [McK90]). *The number of labelled eulerian tournaments on  $n$  vertices is*

$$\left(\frac{2^{n+1}}{\pi n}\right)^{\frac{n-1}{2}} \sqrt{\frac{n}{e}}(1+o(1)).$$

We are now ready to prove the lower bound in Theorem 1.7.

**Theorem 5.2.** *For every sufficiently large  $k$ ,  $m'_k(2k+1) \geq \frac{1}{2} \log(2k+1)$ .*

*Proof.* By Theorem 5.1, we may choose an integer  $k_0$  such that for all  $k \geq k_0$  and for  $n = 2k+1$ , the number of labelled tournaments on  $n$  vertices is at most  $\left(\frac{2^{n+1}}{\pi n}\right)^{\frac{n-1}{2}} \sqrt{n}$  and  $\frac{n-1}{2}(\log \pi - 1) > \log n$ . Now fix some  $k \geq k_0$  and let  $n = 2k+1$ . By Proposition 2.3, all the  $k$ -arc-strong tournaments on  $n$  vertices are eulerian. For two tournaments  $T, T'$  on the same vertex set, we say that  $T'$  is **reachable** from  $T$  by  $t$  inversions for some positive integer  $t$ , if there is a family of sets  $(X_1, \dots, X_t)$  such that  $\text{Inv}(T; X_1, \dots, X_t) = T'$ . Observe that there are  $2^n$  possibilities to choose  $X_i$  for  $i = 1, \dots, t$ , hence the number of tournaments reachable from  $T$  by  $t$  inversions is at most  $(2^n)^t = 2^{nt}$ . Therefore, the number of labelled  $n$ -vertex tournaments that are reachable by  $t$  inversions from an eulerian one is at most

$$\begin{aligned} 2^{nt} \left(\frac{2^{n+1}}{\pi n}\right)^{\frac{n-1}{2}} \sqrt{n} &= 2^{\binom{n}{2}} \cdot 2^{nt} \left(\frac{2}{\pi n}\right)^{\frac{n-1}{2}} \sqrt{n} \\ &= 2^{\binom{n}{2}} \cdot 2^{nt + \frac{n-1}{2}(1 - \log(\pi) - \log n) + \frac{1}{2} \log n} \\ &< 2^{\binom{n}{2}} \cdot 2^{nt - \frac{1}{2}n \log n} \end{aligned}$$

If  $t \leq \frac{1}{2} \log n$ , then  $nt - \frac{1}{2}n \log n \leq 1$ , so the number of such tournaments is less than  $2^{\binom{n}{2}}$ . It follows that there is at least one tournament  $T^*$  on  $2k+1$  vertices which cannot be reached from an eulerian tournament by at most  $t$  inversions. Thus  $\text{sinv}'_k(T^*) > t$ .  $\square$

Theorem 5.2 can be slightly generalised to tournaments on  $2k+c$  vertices for small integers  $c$ . To do so, we will need the following generalisation of Theorem 5.1.

For a positive integer  $n$  and a collection of integers  $\alpha_1, \dots, \alpha_n$ , we denote by  $NT(n; \alpha_1, \dots, \alpha_n)$  the number of labelled tournaments on  $[n]$  in which the vertex  $i$  satisfies  $d^+(i) - d^-(i) = \alpha_i$ .

**Theorem 5.3** (Spencer [Spe74] and McKay [McK90]). *Let  $n$  be a positive odd integer and let  $\alpha_1, \dots, \alpha_n$  be integers. Then*

$$NT(n, \alpha_1, \dots, \alpha_n) \leq \left(\frac{2^{n+1}}{\pi n}\right)^{(n-1)/2} \sqrt{\frac{n}{e}} \exp\left(-\frac{1+o(1)}{2n} \sum_{i=1}^n \alpha_i^2\right).$$

**Corollary 5.4.** *For all integers  $k$  and  $c$ , the number of tournaments  $T$  on  $n = 2k+c$  vertices such that every vertex has in- and out-degree at least  $k$  is at most  $2^{\log(2c)n} \left(\frac{2^{n+1}}{\pi n}\right)^{(n-1)/2} \sqrt{\frac{n}{e}}$  for  $k$  large enough.*

*Proof.* By Theorem 5.3, the number of tournaments  $T$  on  $n = 2k+c$  vertices such that every vertex has in- and out-degree at least  $k$  is at most

$$\left(\frac{2^{n+1}}{\pi n}\right)^{(n-1)/2} \sqrt{\frac{n}{e}} \sum_{\alpha_1, \dots, \alpha_n \in \{-c+1, \dots, c-1\}^n} \exp\left(-\frac{1+o(1)}{2n} \sum_{i=1}^n \alpha_i^2\right).$$

Further, we have

$$\begin{aligned}
\sum_{\alpha_1, \dots, \alpha_n \in \{-c+1, \dots, c-1\}^n} \exp\left(-\frac{1+o(1)}{2n} \sum_{i=1}^n \alpha_i^2\right) &= \prod_{i=1}^n \left( \sum_{\alpha=-c+1}^{c-1} \exp\left(-\frac{1+o(1)}{2n} \alpha^2\right) \right) \\
&\leq \prod_{i=1}^n \left( \sum_{\alpha=-c+1}^{c-1} 1 + o(1) \right) \\
&= (2c-1+o(1))^n \\
&\leq 2^{\log(2c)n}
\end{aligned}$$

for  $n$  large enough, and the result follows.  $\square$

**Theorem 5.5.** *For every positive integer  $c$  fixed, and for every positive integer  $k$  large enough compared to  $c$ , there exists a tournament  $T$  on  $2k+c$  vertices such that  $\text{sinv}'_k(T) > \frac{1}{2} \log(2k+c) - \log(2c)$ . In particular,  $m'_k(2k+c)$  is unbounded for every fixed  $c$ .*

*Proof.* By Corollary 5.4, the number of tournaments  $T$  on  $n = 2k+c$  vertices with  $\text{sinv}'_k(T) \leq t$  is for  $k$  large enough at most

$$2^{nt} 2^{\log(2c)n} \left(\frac{2^{n+1}}{\pi n}\right)^{(n-1)/2} \sqrt{\frac{n}{e}} \leq 2^{\binom{n}{2}} 2^{nt - \frac{1}{2}n \log n + \log(2c)n}$$

which is smaller than  $2^{\binom{n}{2}}$  if  $t \leq \frac{1}{2} \log n - \log(2c)$ . Hence there exists a tournament  $T^*$  with  $\text{sinv}'_k(T^*) > \frac{1}{2} \log(2k+c) - \log(2c)$ .  $\square$

## 5.2 Values of $M'_1, M_1, M'_2$ and $M_2$

We here provide the exact values of  $M_i$  and  $M'_i$  for  $i \in \{1, 2\}$ .

**Proposition 5.6.** *Let  $T$  be a tournament of order  $n \geq 3$ . We have  $\text{sinv}_1(T) = \text{sinv}'_1(T) = 0$  if  $T$  is strong and  $\text{sinv}_1(T) = \text{sinv}'_1(T) = 1$  otherwise. In particular,  $m_1(n) = m'_1(n) = 1$  for all  $n \geq 3$  and  $M_1 = M'_1 = 1$ .*

*Proof.* (i) Trivially, if  $T$  is strong, then  $\text{sinv}_1(T) = 0$ . If  $T$  is not strong, then  $\text{sinv}_1(T) \geq 1$ . Now consider a hamiltonian path of  $T$ . Such a path does exist by Redei's Theorem (see e.g. Theorem 1.4.2 in [BJG09]). Let  $x$  be its initial vertex and  $y$  its terminal vertex. Then inverting  $\{x, y\}$  yields a tournament with a directed hamiltonian cycle because  $V(T) \setminus \{x, y\} \neq \emptyset$ , so a strong tournament. Hence  $\text{sinv}_1(T) = 1$ .  $\square$

A non-strong tournament is said to be **reducible**. By definition, a reducible tournament has a **reduction**  $T_1 \Rightarrow T_2$  that is two subtournaments  $T_1, T_2$  such that  $(V(T_1), V(T_2))$  is a partition of  $V(T)$  and  $V(T_1) \Rightarrow V(T_2)$ .

Let  $S_4$  be the unique strong tournament of order 4. Its vertex set is  $\{a, b, c, d\}$  and its arc set is  $\{ab, bc, cd, da, ca, db\}$ .

**Proposition 5.7.**  $M_2 = M'_2 = 2$ .

*Proof.* The rotative tournament  $R_5$  of order 5 is the only 2-arc-strong tournament of order 5. As observed in [BBBBP10], we have  $\text{inv}(R_5) = 2$ , so  $\text{sinv}'_2(TT_5) = 2$ . Hence  $M_2 \geq M'_2 \geq 2$ .

Let us now prove that  $M_2 \leq 2$ . We shall prove by induction on  $n$  that every tournament  $T$  of order at least 5 satisfies  $\text{sinv}_2(T) \leq 2$ .

Assume first that  $T$  is a tournament of order 5. If  $T$  is strong, then, by Camion's theorem [Cam59], it has a hamiltonian cycle  $v_1 v_2 v_3 v_4 v_5 v_1$ . Let  $A^+ = A(T) \cap \{v_1 v_3, v_2 v_4, v_3 v_5, v_4 v_1, v_5 v_2\}$  and  $A^- = A(T) \cap \{v_3 v_1, v_4 v_2, v_5 v_3, v_1 v_4, v_2 v_5\}$ . We have  $|A^+| + |A^-| = 5$ , so one of the two sets  $A^+, A^-$  has at most two arcs. Reversing the arcs of this set, one after another, yields the 2-strong tournament  $R_5$ .

Assume now that  $T$  is not strong. Then it must be isomorphic to one of the following tournaments or their converse, where the **converse** of a digraph  $D$  is defined to be  $\text{Inv}(D, V(D))$ .



- $TT_5$  with hamiltonian path  $v_1v_2v_3v_4v_5$ . Then inverting  $\{v_1, v_2, v_4, v_5\}$  and  $\{v_1, v_5\}$  yields  $R_5$ .
- $S_4 \Rightarrow \{x\}$ . Then inverting  $\{c, d, x\}$  and  $\{c, d\}$  yields  $R_5$ .
- $\{x\} \Rightarrow \vec{C}_3 \Rightarrow \{y\}$  with  $\vec{C}_3 = abca$ . Then inverting  $\{a, x, y\}$  yields  $R_5$ .
- $\{x\} \Rightarrow \{y\} \Rightarrow \vec{C}_3$  with  $\vec{C}_3 = abca$ . Then inverting  $\{a, b, c, y\}$  and  $\{a, x, y\}$  yields  $R_5$ .

As  $\text{sinv}_2(\text{Inv}(D, V(D))) = \text{sinv}_2(D)$  for digraph  $D$ , the statement follows. Assume now that  $T$  has at least 6 vertices.

Assume first  $T$  has a vertex  $v$  such that  $\min\{d^+(v), d^-(v)\} \geq 2$ . By the induction hypothesis,  $\text{sinv}_2(T - v) \leq 2$ , so there is a family  $\mathcal{X}$  of at most two subsets of  $V(T - v)$  such that  $\text{Inv}(T - v; \mathcal{X})$  is 2-strong. Now  $\text{Inv}(T; \mathcal{X}) - v = \text{Inv}(T - v; \mathcal{X})$  and  $\min\{d^+(v), d^-(v)\} \geq 2$ . Thus, by Lemma 2.1,  $\text{Inv}(T; \mathcal{X})$  is 2-strong, and it follows that  $\text{sinv}_2(T) \leq |\mathcal{X}| \leq 2$ .

Assume now that every vertex has either in-degree at most 1 or out-degree at most 1. Then necessarily  $T$  must be the tournament  $\vec{C}_3 \Rightarrow \vec{C}_3$ . Let  $V(T) = \{a, b, c, d, e, f\}$  with  $\{a, b, c\} \Rightarrow \{d, e, f\}$ . Then inverting  $\{a, b, c, d\}$  and  $\{a, d, e, f\}$  transforms  $T$  into a 2-strong tournament.  $\square$

### 5.3 Upper bounds on $M_k$ and transitive tournaments

In this section, we give some upper bounds on  $M_k$  for all positive integers  $k$ . Most notably, we prove the upper bound in Theorem 1.7.

**Theorem 5.8.**  $M_k \leq 2k$ .

*Proof.* Let  $D$  be a tournament with  $V(D) = \{v_1, \dots, v_n\}$  with  $n \geq 2k + 1$ . Further, let  $T$  be a  $k$ -strong tournament on  $\{v_1, \dots, v_{2k+1}\}$ . We now define sets  $X_1, \dots, X_{2k}$ . Suppose that the sets  $X_1, \dots, X_{i-1}$  have already been created and let  $D_{i-1}$  be the graph obtained from  $D$  by inverting  $X_1, \dots, X_{i-1}$ . Now let  $X_i = \{v_i\} \cup A_i \cup B_i$ , where  $A_i$  is the set of vertices  $v_j$  with  $j \in \{i+1, \dots, 2k+1\}$  for which the edge  $v_i v_j$  has a different orientation in  $T$  and  $D_{i-1}$ , and  $B_i$  is, when  $i \leq k$  (resp.  $i \geq k+1$ ), the set of vertices  $v_j$  with  $j \geq 2k+2$  for which  $D_{i-1}$  contains the arc  $v_i v_j$  (resp.  $v_j v_i$ ).

We still need to show that  $D_{2k}$  is  $k$ -strong. Observe that  $D_{2k} \langle \{v_1, \dots, v_{2k+1}\} \rangle = T$  which is  $k$ -strong by assumption. However, for any  $j \geq 2k+2$ ,  $D_{2k}$  contains the arcs  $v_j v_i$  for  $i = 1, \dots, k$  and the arcs  $v_i v_j$  for  $i = k+1, \dots, 2k$ . Hence, by Lemma 2.1,  $D_{2k}$  is  $k$ -strong.  $\square$

Theorems 5.2 and 5.8 directly imply Theorem 1.7.

It is tempting to improve the upper bound in Theorem 1.7. While we are not able provide an improvement for the general case, we show in the following that a small improvement can be achieved in a seemingly critical case, namely when the size of the tournament is exactly  $2k + 1$ . We first need the following result.

**Theorem 5.9.** Let  $n, k$  be integers with  $n \geq 2k + 1$ .

$$\text{sinv}_k(TT_n) = \text{sinv}'_k(TT_n) = \begin{cases} 2 & \text{if } 2k + 1 \leq n < 3k \\ 1 & \text{if } 3k \leq n. \end{cases}$$

*Proof.* Let  $(v_1, \dots, v_n)$  be an ordering of  $V(TT_n)$  such that  $v_i v_j \in A(TT_n)$  for every  $i < j$ . Let  $X_0 = \{v_i \mid i \text{ even}\}$  and  $X_1 = \{v_i \mid i \text{ odd}\}$ . We claim that  $T' = \text{Inv}(TT_n; X_0, X_1)$  is  $k$ -strong. Consider a set  $Y$  of at most  $k - 1$  vertices of  $T'$ . Let  $i_0$  (resp.  $j_0$ ) be the smallest (resp. largest) even integer  $\ell$  such that  $v_\ell \notin Y$ , and let  $i_1$  (resp.  $j_1$ ) be the smallest (resp. largest) odd integer  $\ell$  such that  $v_\ell \notin Y$ . Observe that  $j_0 > i_1$  and  $j_1 > i_0$  since  $|Y| < k$  and  $n \geq 2k + 1$ . In particular,  $v_{i_1} v_{j_0}, v_{i_0} v_{j_1} \in A(T')$ . Since  $T' \langle X_0 \rangle - Y$  (resp.  $T' \langle X_1 \rangle - Y$ ) is a transitive tournament with source  $v_{j_0}$  and sink  $v_{i_0}$  (resp. source  $v_{j_1}$  and sink  $v_{i_1}$ ), there are hamiltonian paths  $P_0$  from  $v_{j_0}$  to  $v_{i_0}$  in  $T' \langle X_0 \rangle - Y$  and  $P_1$  from  $v_{j_1}$  to  $v_{i_1}$  in  $T' \langle X_1 \rangle - Y$ . Finally  $P_0 v_{i_0} v_{j_1} P_1 v_{i_1} v_{j_0}$  is an hamiltonian directed cycle of  $T' - Y$ , and so  $T' - Y$  is strong. Hence  $T'$  is  $k$ -strong, and  $\text{sinv}_k(TT_n) \leq 2$ .

Let us now prove that  $\text{sinv}'_k(TT_n) \geq 2$  if  $n \leq 3k - 1$ . Suppose for a contradiction that there exists  $X \subseteq V(TT_n)$  such that  $T' = \text{Inv}(TT_n; X)$  is  $k$ -arc-strong. Then, for every  $i \leq k$ , since  $v_i$  has in-degree  $i - 1$  in  $TT_n$ , we have  $v_i \in X$ . Similarly, for every  $i \geq n - k + 1$ , since  $v_i$  has out-degree  $n - i$  in  $TT_n$ , we have  $v_i \in X$ . But then every

out-neighbour of  $v_1$  in  $T'$  is in  $\{v_{k+1}, \dots, v_{n-k}\}$  and so  $d_{T'}^-(v_1) \leq n - 2k < k$ . Hence  $T'$  is not  $k$ -arc-strong, a contradiction.

Finally, let us show that  $\text{inv}_k(TT_n) \leq 1$  if  $n \geq 3k$ . Let  $(v_1, \dots, v_n)$  be an ordering of  $V(TT_n)$  such that  $v_i v_j \in A(TT_n)$  for every  $i < j$ . Let  $A = \{v_1, \dots, v_k\}$ ,  $B = \{v_{k+1}, \dots, v_{2k}\}$  and  $C = \{v_{2k+1}, \dots, v_n\}$ . In  $T' = \text{Inv}(TT_n; A \cup C)$ , we have  $A \Rightarrow B$ ,  $B \Rightarrow C$  and  $C \Rightarrow A$ . Since  $|A|, |B|, |C| \geq k$ ,  $T'$  is  $k$ -strong.  $\square$

As mentioned in the introduction, it was proven independently by Alon et al. [APS<sup>+</sup>22] and Aubian et al. [AHH<sup>+</sup>22] that  $\text{inv}(n) \leq n - \lceil \log(n+1) \rceil$ . Hence, every tournament of order  $2k+1$  can be made acyclic in at most  $2k - \lceil \log(k+1) \rceil$  inversions, and  $k$ -strong in two more inversions by Theorem 5.9. Thus we have the following.

**Corollary 5.10.**  $m_k(2k+1) \leq 2k - \lceil \log(k+1) \rceil + 2$ .

## 5.4 Upper bound on $M'_k$

This section is dedicated to proving Theorem 1.8. First, we give some auxiliary results in Section 5.4.1. Next, in Section 5.4.2, we show the result when restricting to tournaments of size exactly  $2k+1$ . We finally generalize this to bigger tournaments in Section 5.4.3.

### 5.4.1 Preliminaries

We here collect some auxiliary results we need for the proof of Theorem 1.8. We first need a basic result combining probability and linear algebra.

**Proposition 5.11.** *Let  $A \in \mathbb{F}_2^{q \times q'}$  be a matrix whose rank is  $q$  for some integers  $q, q'$  with  $q \leq q'$ . Further, let  $\vec{v} \in \mathbb{F}_2^q$  be a fixed vector and let another vector  $\vec{w} \in \mathbb{F}_2^{q'}$  be drawn uniformly at random. Then  $\Pr(A\vec{w} = \vec{v}) = 2^{-q}$ .*

The simple proof of the following result is similar to the one of Theorem 5.8. It will appear fully in [IWG].

**Proposition 5.12.** *Let  $T_1, T_2$  be tournaments with  $V(T_1) = V(T_2)$ . Then there is a collection of  $|V(T_1)| - 1$  subsets  $X_1, \dots, X_{|V(T_1)|-1}$  of  $V(T_1)$  such that  $\text{Inv}(T_1; X_1, \dots, X_{|V(T_1)|-1}) = T_2$ .*

We are now ready to give the last preliminary result.

**Proposition 5.13.** *Let  $T$  be a tournament on  $2k+1$  vertices and  $X \subseteq V(T)$  with  $|X| \leq \frac{2}{3}k$  such that  $d_T^-(x) = d_T^+(x) = k$  for all  $x \in X$ . Then there exists a  $k$ -arc-strong tournament  $T'$  on  $V(T)$  such that all the edges in  $\text{UG}(T)\langle X \rangle \cup \delta_{\text{UG}(T)}(X)$  have the same orientation in  $T$  and  $T'$ .*

*Proof.* Let  $H$  be the mixed graph which is obtained from  $\text{UG}(T)$  by giving all edges in  $E(\text{UG}(T)\langle X \rangle) \cup \delta_{\text{UG}(T)}(X)$  the orientation they have in  $T$ . Now consider some  $S \subseteq V(T)$ . Let  $x \in S \cap X$ . Since  $d_H^+(x) = d_H^-(x)$ , we have

$$\begin{aligned} d_H^+(S \setminus \{x\}) - d_H^-(S \setminus \{x\}) &= (d_H^+(S) - d_H^+(x, V(H) \setminus (S \cup \{x\})) + d_H^+(S \setminus \{x\}, x)) \\ &\quad - (d_H^-(S) - d_H^-(x, V(H) \setminus (S \cup \{x\})) + d_H^-(S \setminus \{x\}, x)) \\ &= d_H^+(S) - d_H^-(S) \\ &\quad + d_H^+(S \setminus \{x\}, x) + d_H^+(V(H) \setminus (S \cup \{x\}), x) \\ &\quad - d_H^-(x, V(H) \setminus (S \cup \{x\})) - d_H^-(x, S \setminus \{x\}) \\ &= d_H^+(S) - d_H^-(S) + d_H^-(x) - d_H^+(x) \\ &= d_H^+(S) - d_H^-(S). \end{aligned}$$

Repeatedly applying this argument, we obtain  $d_H^+(S) - d_H^-(S) = d_H^+(S \setminus X) - d_H^-(S \setminus X) \leq d_H^+(S \setminus X) \leq |S \setminus X| \cdot |X|$  since every arc in  $H$  have an extremity in  $X$ . Symmetrically, we obtain  $d_H^+(S) - d_H^-(S) = d_H^-(V(H) \setminus S) - d_H^+(V(H) \setminus S) = d_H^-(V(H) \setminus (S \cup X)) - d_H^+(V(H) \setminus (S \cup X)) \leq |V(H) \setminus (S \cup X)| \cdot |X|$ .

As  $|X| \leq \frac{2}{3}k \leq \frac{1}{2}|V(H) \setminus X| = \frac{1}{2}(|S \setminus X| + |V(H) \setminus (S \cup X)|)$ , we obtain

$$\begin{aligned}
d_H^+(S) - d_H^-(S) &\leq \min\{|S \setminus X|, |V(H) \setminus (S \cup X)|\} \cdot |X| \\
&\leq \min\{|S \setminus X|, |V(H) \setminus (S \cup X)|\} \cdot \frac{1}{2}(|S \setminus X| + |V(H) \setminus (S \cup X)|) \\
&\leq \min\{|S \setminus X|, |V(H) \setminus (S \cup X)|\} \cdot \max\{|S \setminus X|, |V(H) \setminus (S \cup X)|\} \\
&= |S \setminus X| \cdot |V(H) \setminus (S \cup X)| \\
&= d_H(S).
\end{aligned}$$

Hence, by Proposition 2.4, there is an eulerian orientation  $T'$  of  $T$  such that all the edges in  $\text{UG}(T) \setminus X \cup \delta_{\text{UG}(T)}(X)$  have the same orientation in  $T$  and  $T'$ . As  $|V(T)| = 2k + 1$ , the tournament  $T'$  is  $k$ -arc-strong by Proposition 2.3.  $\square$

#### 5.4.2 Main proof for tournaments of size $2k + 1$

We here give the proof of Theorem 1.8 for tournaments of size  $2k + 1$ . More precisely, we prove the following statement.

**Theorem 5.14.** *For every  $\epsilon > 0$ , there is an integer  $k_0$  such that for every  $k \geq k_0$  and for every tournament  $T$  on  $2k + 1$  vertices, we have  $\text{sinv}'_k(T) \leq (\frac{4}{3} + \epsilon)k$ .*

*Proof.* Let  $\epsilon > 0$ . We may assume  $\epsilon \leq 1/3$ . We choose  $k_0$  large enough so that for all  $k \geq k_0$ , the two following inequalities hold:

- $\lceil \log^2 k \rceil + \lceil \frac{1}{4} \log k \rceil + \frac{2k}{3 \lceil \frac{1}{4} \log k \rceil} \leq \epsilon k$
- $10 \log k \leq k^{\frac{1}{8}}$ ,

In particular, the later inequality implies that  $k_0 > 2^{12}$ . Thus the following inequalities also hold for all  $k \geq k_0$ :  $\lceil \frac{1}{4} \log k \rceil \leq \frac{1}{3}k$ ;  $\lceil \frac{1}{4} \log k \rceil \leq \sqrt{k}$ ;  $k \geq 13$ ;  $3 \lceil \frac{1}{4} \log k \rceil \leq \log k$ ;  $16 \log k \lceil \frac{1}{4} \log k \rceil \leq \sqrt{k}$ .

Let  $T$  be a tournament on  $2k + 1$  vertices for some  $k \geq k_0$ . Let  $A \subseteq V(T)$  be an arbitrary subset of  $V(T)$  with  $|A| = k$  and  $B = V(T) \setminus A$ . For some  $b \in B$  and a tournament  $\tilde{T}$  on  $V(T)$ , we denote  $|k - d_{\tilde{T}}^-(b)|$  by  $\text{def}_{\tilde{T}}(b)$ . Let  $q = \lceil \frac{1}{4} \log k \rceil$  and  $q^* = \lceil \log^2 k \rceil$ . We now choose  $q^*$  sets  $X_1, \dots, X_{q^*} \subseteq V(T)$  independently and uniformly at random. Let  $T' = \text{Inv}(T; X_1, \dots, X_{q^*})$ . For some  $b \in B$ , we denote by  $\vec{b}$  the vector in  $\mathbb{F}_2^{q^*}$  whose  $i$ -th entry is 1 if  $b \in X_i$  and 0 otherwise for  $i \in [q^*]$ . Observe that by the choice of  $X_1, \dots, X_{q^*}$ , the vectors  $\{\vec{b} \mid b \in B\}$  follow a uniform distribution and are mutually independent.

We now define a list of possible events:

- $E_1$  : There is some  $b \in B$  with  $\text{def}_{T'}(b) \geq 2 \log k \sqrt{k}$ .
- $E_2$  : There are some  $b_1, \dots, b_q \in B$  such that  $\vec{b}_1, \dots, \vec{b}_q$  are linearly dependent.
- $E_3$  : There are some  $b_1, \dots, b_q \in B$  such that  $\vec{b}_1, \dots, \vec{b}_q$  are linearly independent and a sequence  $(\star_1, \dots, \star_q) \in \{+, -\}^q$  such that  $|\bigcap_{i=1}^q N_{T'}^{\star_i}(b_i) \cap A| \leq 5 \log k \sqrt{k}$ .

**Lemma 5.15.** *If none of  $E_1$ ,  $E_2$  and  $E_3$  occur, then  $\text{sinv}'_k(T) \leq (\frac{4}{3} + \epsilon)k$ .*

*Proof.* Let  $(B_1, \dots, B_t)$  be a maximal collection of disjoint subsets of  $B$  with the following properties:

- $|B_i| = q$  for  $i = 1, \dots, t$ ,
- $\text{def}_{T'}(b) \equiv \text{def}_{T'}(b') \pmod{2}$  for all  $b, b' \in B_i$  for  $i = 1, \dots, t$ ,
- $tq \leq \frac{2}{3}k$ .

**Claim 5.15.1.**  $|\bigcup_{i=1}^t B_i| \geq \frac{2}{3}k - q$ .

*Proof of claim.* Suppose otherwise. Observe that, as  $\lceil \frac{1}{4} \log k \rceil \leq \frac{1}{3}k$ , we have  $|B \setminus (\bigcup_{i=1}^t B_i)| = |B| - |\bigcup_{i=1}^t B_i| \geq (k+1) - (\frac{2}{3}k - q) \geq \frac{1}{3}k + q \geq 2q$ . By the Pigeonhole Principle, there exists a set  $B_0 \subseteq B \setminus (\bigcup_{i=1}^t B_i)$  with  $|B_0| = q$  such that  $\text{def}_{T'}(b) \equiv \text{def}_{T'}(b') \pmod{2}$  for all  $b, b' \in B_0$ . Hence  $B_0$  can be added to  $\mathcal{B}$  without violating any of the above conditions, a contradiction to the maximality of  $(B_1, \dots, B_t)$ .  $\diamond$

**Claim 5.15.2.** *Let  $i \in [t]$ . There is a set  $Y_i \subseteq B_i \cup A$  such that  $\text{def}_{\text{Inv}(T'; Y_i)}(b) = 0$  for all  $b \in B_i$ .*

*Proof of claim.* Let  $p = \max\{\text{def}_{\text{Inv}(T'; B_i)}(b) \mid b \in B_i\}$ . Observe that  $p \leq \max\{\text{def}_{T'}(b) + q - 1 \mid b \in B_i\}$ . Hence, as  $E_1$  does not occur and  $q = \lceil \frac{1}{4} \log k \rceil \leq \log k \sqrt{k}$ , we have  $p \leq 2 \log k \sqrt{k} + \log k \sqrt{k} = 3 \log k \sqrt{k}$ . We shall now build a sequence of sets  $Y_i^0, \dots, Y_i^p$  with the following properties:

- $Y_i^0 = B_i$ ,
- $Y_i^{j+1}$  is obtained from  $Y_i^j$  by adding one vertex from  $A \setminus Y_i^j$  for  $j = 0, \dots, p-1$ ,
- $\max\{\text{def}_{\text{Inv}(T'; Y_i^j)}(b) \mid b \in B_i\} = p - j$  for  $j = 0, \dots, p$ .

The statement then follows for  $Y_i = Y_i^p$ .

Suppose that we have already created the sets  $Y_i^0, \dots, Y_i^j$  for some  $j$  with  $0 \leq j \leq p-1$  and now want to create  $Y_i^{j+1}$ . Let  $(b_1, \dots, b_q)$  be an arbitrary ordering of the vertices in  $B_i$ . For every  $\mu \in [q]$ , if  $\max\{\text{def}_{\text{Inv}(T'; Y_i^j)}(b) \mid b \in B_i\} = d_{\text{Inv}(T'; Y_i^j)}^-(b_\mu) - k$ , let  $\star_\mu = -$ , if  $\max\{\text{def}_{\text{Inv}(T'; Y_i^j)}(b) \mid b \in B_i\} = k - d_{\text{Inv}(T'; Y_i^j)}^-(b_\mu)$ , let  $\star_\mu = +$ , otherwise choose  $\star_\mu \in \{+, -\}$  arbitrarily. As  $E_2$  does not occur,  $\vec{b}_1, \dots, \vec{b}_q$  are linearly independent. Moreover, as  $E_3$  does not occur,  $|\bigcap_{\mu=1}^q N_{T'}^{\star_\mu}(b_\mu) \cap A| \geq 5 \log k \sqrt{k}$ . As  $p \leq 3 \log k \sqrt{k}$ , and  $q \leq \log k \sqrt{k}$ , we have  $|Y_i^j| \leq p + q \leq 4 \log k \sqrt{k}$ . Thus  $(\bigcap_{\mu=1}^q N_{T'}^{\star_\mu}(b_\mu) \cap A) \setminus Y_i^j \neq \emptyset$ . We can hence choose some  $y \in (\bigcap_{\mu=1}^q N_{T'}^{\star_\mu}(b_\mu) \cap A) \setminus Y_i^j$  and define  $Y_i^{j+1} = Y_i^j \cup \{y\}$ . By definition,  $Y_i^{j+1}$  is obtained from  $Y_i^j$  by adding one vertex in  $A \setminus Y_i^j$ . In order to see that the last property holds, let  $\mu \in [q]$ . If  $\max\{\text{def}_{\text{Inv}(T'; Y_i^j)}(b) \mid b \in B_i\} = d_{\text{Inv}(T'; Y_i^j)}^-(b_\mu) - k$ , we have  $\star_\mu = -$  and hence  $y \in N_{\text{Inv}(T'; Y_i^j)}^-(b_\mu)$ . As  $d_{\text{Inv}(T'; Y_i^j)}^-(b_\mu) - k > 0$ , we obtain  $\text{def}_{\text{Inv}(T'; Y_i^{j+1})}(b_\mu) = d_{\text{Inv}(T'; Y_i^{j+1})}^-(b_\mu) - k = d_{\text{Inv}(T'; Y_i^j)}^-(b_\mu) - k - 1 = p - j - 1 = p - (j + 1)$ . Similarly, if  $\max\{\text{def}_{\text{Inv}(T'; Y_i^j)}(b) \mid b \in B_i\} = k - d_{\text{Inv}(T'; Y_i^j)}^-(b_\mu)$ , then  $\text{def}_{\text{Inv}(T'; Y_i^{j+1})}(b) = p - (j + 1)$ . Otherwise, let  $b' \in B$  with  $\text{def}_{\text{Inv}(T'; Y_i^j)}(b') = p - j$ . As  $\text{def}_{T'}(b_\mu) \equiv \text{def}_{T'}(b') \pmod{2}$ , we have  $\text{def}_{\text{Inv}(T'; Y_i^j)}(b_\mu) \equiv \text{def}_{T'}(b_\mu) + |Y_i^j| - 1 \equiv \text{def}_{T'}(b') + |Y_i^j| - 1 \equiv \text{def}_{\text{Inv}(T'; Y_i^j)}(b')$

$\pmod{2}$ , hence  $\text{def}_{\text{Inv}(T'; Y_i^j)}(b_\mu) \leq \text{def}_{\text{Inv}(T'; Y_i^j)}(b') - 2$ . This yields  $\text{def}_{\text{Inv}(T'; Y_i^{j+1})}(b_\mu) \leq \text{def}_{\text{Inv}(T'; Y_i^j)}(b_\mu) + 1 \leq \text{def}_{\text{Inv}(T'; Y_i^j)}(b') - 2 + 1 = p - (j + 1)$ .  $\diamond$

Let  $T'' = \text{Inv}(T'; (Y_1, \dots, Y_t))$ . Observe that by definition of the sets  $Y_i$ , we have  $\text{def}_{T''}(b) = 0$  for all  $b \in \bigcup_{i=1}^t B_i$ . As  $|\bigcup_{i=1}^t B_i| = qt \leq \frac{2}{3}k$ , we obtain by Proposition 5.13 that there is a  $k$ -arc-strong tournament  $T'''$  on  $V(T)$  such that all the edges in  $\text{UG}(T) \setminus (\bigcup_{i=1}^t B_i) \cup \delta_{\text{UG}(T)}(\bigcup_{i=1}^t B_i)$  have the same orientation in  $T'''$  and  $T''$ . Further, by Proposition 5.12, there is a collection of  $r = |V(T) \setminus \bigcup_{i=1}^t B_i| - 1$  subsets  $Z_1, \dots, Z_r$  of  $V(T) \setminus \bigcup_{i=1}^t B_i$  such that  $\text{Inv}(T''' \setminus V(T) \setminus \bigcup_{i=1}^t B_i; Z_1, \dots, Z_r) = T''' \setminus V(T) \setminus \bigcup_{i=1}^t B_i$ . We obtain that  $\text{Inv}(T; X_1, \dots, X_{q^*}, Y_1, \dots, Y_t, Z_1, \dots, Z_r) = \text{Inv}(T'; Y_1, \dots, Y_t, Z_1, \dots, Z_r) = \text{Inv}(T''; Z_1, \dots, Z_r) = T'''$ . As  $T'''$  is  $k$ -arc-strong, we have  $\text{sinv}'_k(T) \leq q^* + t + r$ . Now  $q^* = \lceil \log^2 k \rceil$ ,  $t \leq \frac{2k}{3q} = \frac{2k}{3 \lceil \frac{1}{4} \log k \rceil}$ , and  $r = |V(T)| - |\bigcup_{i=1}^t B_i| - 1 \leq 2k - (\frac{2}{3}k - q) = \frac{4}{3}k + q = \frac{4}{3}k + \lceil \frac{1}{4} \log k \rceil$  by Claim 5.15.1. Thus  $\text{sinv}'_k(T) \leq \frac{4}{3}k + \lceil \log^2 k \rceil + \lceil \frac{1}{4} \log k \rceil + \frac{2k}{3 \lceil \frac{1}{4} \log k \rceil} \leq \frac{4}{3}k + \epsilon k$ .  $\square$

In the following lemmas, we show that the probability for each of  $E_1, E_2$  and  $E_3$  is small.

**Lemma 5.16.**  $\Pr(E_1) < \frac{1}{3}$ .

*Proof.* Let  $b \in B$ . We first bound the probability that  $\vec{b} = \vec{0}$ . Due to the independent and uniform choice of  $(X_1, \dots, X_{q^*})$ , we have  $\Pr[\vec{b} = \vec{0}] = 2^{-q^*} \leq 2^{-2 \log k} = \frac{1}{k^2}$ .

Now suppose that  $\vec{b} \neq \vec{0}$ . Once  $\vec{b}$  revealed, for every  $v \in V(T)$ , let  $\Gamma_v$  be a random variable which is 1 if  $T'$  contains the arc  $bv$  and 0 otherwise and let  $\Gamma = \sum_{v \in V(T-b)} \Gamma_v$ . Observe that, due to the independent and uniform choice of  $(X_1, \dots, X_{q^*})$ , we have  $\Pr(\Gamma_v = 1) = \frac{1}{2}$  for all  $v \in V(T-b)$  and the  $\Gamma_v$  are independent. This yields that  $\Gamma \sim \text{Bin}(2k, \frac{1}{2})$ . Further, observe that  $\text{def}_{T'}(b) \leq 2 \log k \sqrt{k}$  if and only if  $k - 2 \log k \sqrt{k} \leq \Gamma \leq k + 2 \log k \sqrt{k}$ . By Chernoff's Bound (Proposition 2.6), we obtain

$$\begin{aligned} \Pr[\Gamma < k - 2 \log k \sqrt{k}] &\leq \Pr \left[ \Gamma < \left( 1 - \frac{2 \log k \sqrt{k}}{k} \right) k \right] \\ &\leq \exp \left( -\frac{4k \log^2 k}{2k^2} k \right) \\ &\leq \exp(-2 \log k) \\ &\leq \exp(-2 \ln k) \\ &= \frac{1}{k^2}. \end{aligned}$$

Similarly, since  $2k - \Gamma \sim \text{Bin}(2k, \frac{1}{2})$ , we obtain that  $\Pr[\Gamma > k + 2 \log k \sqrt{k}] \leq \frac{1}{k^2}$ . This yields that

$$\begin{aligned} \Pr[\text{def}_{T'}(b) > 2 \log k \sqrt{k}] &= \Pr[\vec{b} = \vec{0}] \cdot \Pr[\text{def}_{T'}(b) > 2 \log k \sqrt{k} \mid \vec{b} = \vec{0}] + \Pr[\vec{b} \neq \vec{0}] \cdot \Pr[\text{def}_{T'}(b) > 2 \log k \sqrt{k} \mid \vec{b} \neq \vec{0}] \\ &\leq \Pr[\vec{b} = \vec{0}] + \Pr[\Gamma < k - 2 \log k \sqrt{k} \mid \vec{b} \neq \vec{0}] + \Pr[\Gamma > k + 2 \log k \sqrt{k} \mid \vec{b} \neq \vec{0}] \\ &\leq \frac{3}{k^2}. \end{aligned}$$

As there are  $k + 1$  vertices contained in  $B$ , by the Union Bound (Proposition 2.5),  $\Pr(E_1) \leq (k + 1) \frac{3}{k^2} < \frac{1}{3}$ , as  $k \geq 10$ .  $\square$

**Lemma 5.17.**  $\Pr(E_2) < \frac{1}{3}$ .

*Proof.* Let  $b_1, \dots, b_q \in B$ . Recall that  $\vec{b}_1, \dots, \vec{b}_q$  are independent and uniformly distributed. For  $i \in [q]$ , observe that if  $\vec{b}_1, \dots, \vec{b}_i$  are linearly independent, they span a vector space containing  $2^i$  elements. Hence

$$\begin{aligned} \Pr[\vec{b}_1, \dots, \vec{b}_q \text{ linearly dependent}] &= \sum_{i=1}^q \Pr[\vec{b}_1, \dots, \vec{b}_i \text{ linearly dependent} \mid \vec{b}_1, \dots, \vec{b}_{i-1} \text{ linearly independent}] \\ &\quad \times \Pr[\vec{b}_1, \dots, \vec{b}_{i-1} \text{ linearly independent}] \\ &\leq \sum_{i=1}^q \Pr[\vec{b}_1, \dots, \vec{b}_i \text{ linearly dependent} \mid \vec{b}_1, \dots, \vec{b}_{i-1} \text{ linearly independent}] \\ &= \sum_{i=1}^q \frac{2^{i-1}}{2^{q^*}} \\ &\leq \frac{2^q}{2^{q^*}} \\ &\leq 2^{-2 \log(k)q} \quad \text{because } q^* \geq \log^2 k \geq 3 \log(k)q \geq 2 \log(k)q + q \\ &= k^{-2q}. \end{aligned}$$

Observe that there are  $\binom{k+1}{q}$  possibilities to choose  $b_1, \dots, b_q$ . Hence the Union Bound (Proposition 2.5) yields  $\Pr(E_2) \leq \binom{k+1}{q} k^{-2q} \leq (2k)^q k^{-2q} = \left(\frac{2}{k}\right)^q < \frac{1}{3}$  as  $q \geq 1$  and  $k \geq 7$ .  $\square$

**Lemma 5.18.**  $\Pr(E_3) < \frac{1}{3}$ .

*Proof.* Let  $b_1, \dots, b_q \in B$  such that  $\vec{b}_1, \dots, \vec{b}_q$  are linearly independent and let  $(\star_1, \dots, \star_q) \in \{+, -\}^q$ . Once  $\vec{b}_1, \dots, \vec{b}_q$  revealed, for every  $a \in A$ , let  $\Gamma_a$  be a random variable that is 1 if  $a \in \bigcap_{i=1}^q N_{T^i}^{\star_i}(b_i)$  and 0 otherwise and let  $\Gamma = \sum_{a \in A} \Gamma_a$ . Observe that, due to the independent and uniform choice of  $(X_1, \dots, X_{q^*})$  and by Proposition 5.11, we have that  $\Pr[\Gamma_a = 1] = \frac{1}{2^q}$  for all  $a \in A$  and that the  $\Gamma_a$  are independent. This yields that  $\Gamma \sim \text{Bin}(k, \frac{1}{2^q})$ . Further observe that  $\Gamma = |\bigcap_{i=1}^q N_{T^i}^{\star_i}(b_i) \cap A|$ . Now, since  $10 \log k \leq k^{\frac{1}{8}}$  and  $q = \lceil \frac{1}{4} \log k \rceil \leq \frac{3}{8} \log k$ , we obtain  $5 \log k \sqrt{k} \leq \frac{1}{2} k^{\frac{5}{8}} = \frac{1}{2} k 2^{-\frac{3}{8} \log k} \leq \frac{1}{2} k 2^{-q}$ . By Chernoff's Bound (Proposition 2.6), we obtain

$$\begin{aligned} \Pr[\Gamma < 5 \log k \sqrt{k}] &\leq \Pr\left[\Gamma < \left(1 - \frac{1}{2}\right) k 2^{-q}\right] \\ &\leq \exp\left(-\frac{1}{8} k 2^{-q}\right) \\ &\leq \exp\left(-\frac{1}{8} k^{\frac{1}{2}}\right) \\ &\leq \exp(-2q \log k) \quad \text{because } \sqrt{k} \geq 16 \log k \lceil \frac{1}{4} \log k \rceil \\ &\leq \exp(-2q \ln k) \\ &= k^{-2q}. \end{aligned}$$

Observe that there are  $\binom{k+1}{q}$  possibilities to choose  $b_1, \dots, b_q$  and for each of those, there are  $2^q$  possibilities to choose  $(\star_1, \dots, \star_q)$ . Hence the Union Bound (Proposition 2.5) yields  $\Pr(E_3) \leq \binom{k+1}{q} 2^q k^{-2q} \leq (2k)^q 2^q k^{-2q} = \left(\frac{4}{k}\right)^q < \frac{1}{3}$ , as  $q \geq 1$  and  $k \geq 13$ .  $\square$

We are now ready to conclude the proof of Theorem 5.14. By Lemmas 5.16, 5.17 and 5.18 and the Union Bound (Proposition 2.5), we obtain that, with positive probability, none of  $E_1$ ,  $E_2$  and  $E_3$  occur. Hence the statement follows from Lemma 5.15.  $\square$

### 5.4.3 Bigger tournaments

We now show how to extend this result to larger tournaments and derive Theorem 1.8 from Theorem 5.14.

**Theorem 1.8.**  $M'_k \leq \frac{4}{3}k + o(k)$ .

*Proof.* Let  $\epsilon > 0$ . We shall prove that there is an integer  $k^*$  such that for every  $k \geq k^*$  and for every tournament  $T$  on at least  $2k + 1$  vertices, we have  $\text{sinv}'_k(T) \leq \left(\frac{4}{3} + \epsilon\right)k$ .

By Theorem 5.14, there is an integer  $k_0$  such that for every  $k \geq k_0$  and for every tournament  $T$  on at least  $2k + 1$  vertices, we have  $\text{sinv}'_k(T) \leq \left(\frac{4}{3} + \frac{\epsilon}{2}\right)k$ . Further, let  $k_1$  be an integer such that  $\frac{\epsilon}{2}k \geq 6 \log k \sqrt{k} + 1$ ,  $\sqrt{k} \geq 3 \log k$  and  $\log k \geq 5$  hold for all  $k \geq k_1$ . Set  $k^* = \max\{k_0, k_1\}$ .

We now fix an integer  $k \geq k^*$  and prove that for every tournament  $T$  on  $n \geq 2k + 1$  vertices, we have  $\text{sinv}'_k(T) \leq \left(\frac{4}{3} + \frac{\epsilon}{2}\right)k + \min\{n - (2k + 1), 6 \log k \sqrt{k} + 1\}$  from which the statement follows immediately. Suppose that  $T$  is a tournament that does not satisfy this statement and whose number  $n$  of vertices is minimum with respect to this property.

**Case 1.**  $n \geq 4k - 2$ .

By Proposition 2.2, there is some  $v \in V(T)$  with  $\min\{d_T^+(v), d_T^-(v)\} \geq k$ . By the minimality of  $T$ , there is a collection  $\mathcal{X}$  of  $\left(\frac{4}{3} + \frac{\epsilon}{2}\right)k + \min\{(n - 1) - (2k + 1), 6 \log k \sqrt{k} + 1\}$  subsets of  $V(T - v)$  such that  $\text{Inv}(T - v; \mathcal{X})$  is  $k$ -arc-strong. By Proposition 2.1, we obtain that  $\text{Inv}(T; \mathcal{X})$  is  $k$ -arc-strong. As  $\left(\frac{4}{3} + \frac{\epsilon}{2}\right)k + \min\{(n - 1) - (2k + 1), 6 \log k \sqrt{k} + 1\} \leq \left(\frac{4}{3} + \frac{\epsilon}{2}\right)k + \min\{n - (2k + 1), 6 \log k \sqrt{k} + 1\}$ , we obtain a contradiction.

**Case 2.**  $2k + 1 + 6 \log \sqrt{k} \leq n \leq 4k - 2$ .

Let  $A$  be an arbitrary subset of  $V(T)$  of size  $2k + 1 + 6 \log k\sqrt{k}$  and let  $B = V(T) \setminus A$ . We now choose a subset  $X$  of  $A$  uniformly at random and let  $T' = \text{Inv}(T; B \cup X)$ . Now consider some  $b \in B$ . For every  $a \in A$ , let  $\Gamma_a$  be the probability that  $ab \in A(T')$  and observe that  $\Gamma = \sum_{a \in A} \Gamma_a$  is exactly  $|N_{T'}^-(b) \cap A|$ . Due to the uniform choice of  $X$ , we have  $\Pr[ab \in A(T)] = \frac{1}{2}$  and that the  $\Gamma_a$  are independent. This yields  $\Gamma \sim \text{Bin}(2k + 1 + 6 \log k\sqrt{k}, \frac{1}{2})$ . By Chernoff's Bound (Proposition 2.6), we obtain

$$\begin{aligned} \Pr[\Gamma < k] &\leq \Pr \left[ \Gamma < \left( 1 - \frac{\log k\sqrt{k}}{k} \right) \frac{1}{2} (2k + 1 + 6 \log k\sqrt{k}) \right] \\ &\leq \exp(-\log^2(k)/2) \\ &\leq \exp(-2 \log k) \quad \text{because } k \geq 16 \\ &\leq \exp(-2 \ln k) \\ &= \frac{1}{k^2}. \end{aligned}$$

Similarly, we obtain  $\Pr[|N_{T'}^+(b) \cap A| < k] \leq \frac{1}{k^2}$ . As  $B$  contains at most  $(4k - 3) - (2k + 1 + 2 \log k\sqrt{k}) \leq 2k$  elements, the Union Bound (Proposition 2.5) yields that the probability that there is at least one  $b \in B$  with  $\min\{|N_{T'}^-(b) \cap A|, |N_{T'}^+(b) \cap A|\} \leq k$  is at most  $\frac{2}{k^2} 2k = \frac{4}{k} < 1$ , as  $k \geq 5$ . Hence, with positive probability, we have  $\min\{|N_{T'}^-(b) \cap A|, |N_{T'}^+(b) \cap A|\} \geq k$  for all  $b \in B$ . Thus there is  $X_0$  such that every vertex of  $B$  has in- and out-degree at least  $k$  in  $T'_0 = \text{Inv}(T; B \cup X_0)$ . Further, by the minimality of  $T$ , there is a collection  $\mathcal{X}$  of  $\text{sinv}'_k(T_0 \langle A \rangle) \leq (\frac{4}{3} + \frac{\epsilon}{2})k + 6 \log k\sqrt{k}$  sets such that  $\text{Inv}(T'_0 \langle A \rangle)$  is  $k$ -arc-strong. Hence, by Proposition 2.1, we obtain that  $\text{Inv}(T; \{X_0\} \cup \mathcal{X})$  is  $k$ -arc-strong, a contradiction.

**Case 3.**  $2k + 2 \leq n \leq 2k + 6 \log k\sqrt{k}$ .

Let  $v \in V(T)$  be an arbitrary vertex. We can then find a set  $X \subseteq V(T)$  such that for  $T' = \text{Inv}(T; X)$ , we have  $\min\{d_{T'}^+(v), d_{T'}^-(v)\} \geq k$ . Further, by the minimality of  $T$ , there is a collection  $\mathcal{X}$  of  $\text{sinv}'_k(T) \leq (\frac{4}{3} + \frac{\epsilon}{2})k + n - (2k + 1) - 1$  sets such that  $\text{Inv}(T' \langle A \rangle; \mathcal{X})$  is 2-arc-strong. Hence by Proposition 2.1, we obtain that  $\text{Inv}(T; \{X\} \cup \mathcal{X})$  is  $k$ -arc-strong, a contradiction.  $\square$

## 6 Upper bounds on $m_k(n)$

In this section, we prove several results showing that tournaments on significantly more than  $2k$  vertices can be made  $k$ -strong by a small number of inversions. More precisely, in Subsection 6.1, we prove Theorem 1.10, in Subsection 6.2, we prove Theorem 1.13, in Subsection 6.3, we prove Theorem 1.15, and in Subsection 6.4, we prove Theorem 1.16.

### 6.1 First upper bounds on $m_k(n)$

In this subsection, we first establish that, for every fixed  $k$ , every tournament which is sufficiently large in comparison to  $k$  can be made  $k$ -strong by a single inversion. More precisely, we prove Theorem 1.10. While Theorem 1.10 is clearly weaker than Theorem 1.13, this result justifies some of the notation used later on, and may serve as a warm-up exercise of the more involved proof of Theorem 1.13.

*Proof of Theorem 1.10.* Let  $T$  be a tournament of order  $n \geq (2k - 1)2^{2k}$ .

It is easy and well-known that if  $D$  is an acyclic digraph,  $x$  a source in  $D$ , and  $D - x$  is contained (as a subdigraph) in every tournament of order  $n$ , then  $D$  is contained (as a subdigraph) in every tournament of order  $2n$ . An easy induction yields that  $T$  contains three sets  $A_1, A_2, A_3$  such that  $A_1 \Rightarrow (A_2 \cup A_3)$  and  $A_2 \Rightarrow A_3$  with  $|A_1| = |A_3| = k$  and  $|A_2| = 2k - 1$ . Set  $A = A_1 \cup A_2 \cup A_3$ . Let  $I$  be the set of vertices in  $V(T) \setminus A$  that have either less than  $k$  out-neighbours in  $A$  or less than  $k$  in-neighbours in  $A$ . Let  $X = A_1 \cup A_3 \cup I$ . Let us prove that  $T' = \text{Inv}(T; X)$  is  $k$ -strong.

In  $T'$ , we have  $A_1 \Rightarrow A_2 \Rightarrow A_3 \Rightarrow A_1$ . Since the three sets  $A_1$ ,  $A_2$ , and  $A_3$  have size at least  $k$ , the tournament  $T' \langle A \rangle$  is  $k$ -strong. Now consider a vertex  $v$  in  $V(T) \setminus A$ . If  $v \notin I$ , then no arcs incident to  $v$  have been reversed so its in- and out-degree have been unchanged and are at least  $k$  by definition of  $I$ . If  $v \in I$ , then all the arcs between  $v$  and  $A_1 \cup A_3$  have been reversed and those between  $v$  and  $A_2$  are unchanged. If  $v$  has less than  $k$  out-neighbours in  $A$  in  $T$ , then  $|N_{T'}^+(v) \cap A| \geq |N_T^-(v) \cap (A_1 \cup A_3)| \geq 2k - d_T^+(v) \geq k$ , and  $|N_{T'}^-(v) \cap A| \geq |N_T^-(v) \cap A_2| \geq 2k - 1 - d_T^+(v) \geq k$ . Similarly, if  $v$  has less than  $k$  in-neighbours in  $A$  in  $T$ , then  $|N_{T'}^+(v) \cap A| \geq k$  and  $|N_{T'}^-(v) \cap A| \geq k$ . Thus, by Lemma 2.1,  $T'$  is  $k$ -strong. Hence  $\text{sinv}_k(T) \leq 1$ .  $\square$

## 6.2 Linear upper bound on $N_k(1)$

In this subsection, we shall prove Theorem 1.13 which states that  $N_k(1) \leq 19k - 2$ . To prove it we need some preliminaries.

For a digraph  $D$ , let  $\sigma = (v_1, v_2, \dots, v_n)$  be an ordering of the vertices of  $D$ . An arc  $v_i v_j$  is **forward** (according to  $\sigma$ ) if  $i < j$  and **backward** (according to  $\sigma$ ) if  $j < i$ . A **median order** of  $D$  is an ordering of the vertices of  $D$  with the maximum number of forward arcs, or equivalently the minimum number of backward arcs.

Let us note basic well-known properties of median orders of tournaments (the feedback property in [HT00]).

**Lemma 6.1.** *Let  $T$  be a tournament and  $(v_1, v_2, \dots, v_n)$  a median order of  $T$ . Then, for any two indices  $i, j$  with  $1 \leq i < j \leq n$ :*

(M1)  $(v_i, v_{i+1}, \dots, v_j)$  is a median order of the induced subtournament  $T \langle \{v_i, v_{i+1}, \dots, v_j\} \rangle$ .

(M2) *the vertex  $v_i$  dominates at least half of the vertices  $v_{i+1}, v_{i+2}, \dots, v_j$ , and the vertex  $v_j$  is dominated by at least half of the vertices  $v_i, v_{i+1}, \dots, v_{j-1}$ . In particular, each vertex  $v_i$ ,  $1 \leq i < n$ , dominates its successor  $v_{i+1}$ .*

Let  $v$  be a vertex of a digraph  $D$ . We denote by  $R_D^+(v)$  (resp.  $R_D^-(v)$ ) the set of vertices which are **reachable** from  $v$  (resp. vertices that can reach  $v$ ) in  $D$ , that are the vertices  $w$  such that there is a directed  $(v, w)$ -path (resp.  $(w, v)$ -path) in  $D$ . Note that  $v \in R_D^+(v)$ .

**Lemma 6.2.** *Let  $T$  be a tournament with median order  $(v_1, v_2, \dots, v_n)$ . Let  $F$  be a subset of vertices such that  $v_1 \notin F$ . Then  $|R_{T-F}^+(v_1)| \geq n - 2|F|$ .*

*Proof.* We prove the result by induction on  $n + |F|$ , the result holding trivially by (M2) if  $|F| = 0$ .

If all the out-neighbours of  $v_1$  in  $T$  are in  $|F|$ , then by (M2),  $|N_T^-(v_1)| \leq |N_T^+(v_1)| \leq |F|$ . Hence  $n - 1 \leq 2|F|$ , and the result holds. Henceforth we may assume that  $v_1$  has an out-neighbour not in  $F$ . Let  $i_0$  be the smallest index of such a vertex. Let  $T_0 = T \langle \{v_1, \dots, v_{i_0-1}\} \rangle$ ,  $T_1 = T \langle \{v_{i_0}, \dots, v_n\} \rangle$ ,  $F_0 = F \cap V(T_0)$  and  $F_1 = F \cap V(T_1)$ . By (M1),  $(v_1, v_2, \dots, v_{i_0-1})$  is a median order of  $T_0$  and  $(v_{i_0}, \dots, v_n)$  is a median order of  $T_1$ . By definition of  $i_0$ , all out-neighbours of  $v_1$  in  $T_0$  are in  $F_0$ . Thus, as above, we have  $i_0 - 2 \leq 2|F_0|$ . By the induction hypothesis,  $|R_{T_1-F_1}^+(v_{i_0})| \geq n - i_0 + 1 - 2|F_1|$ . Now  $R_{T_1-F_1}^+(v_{i_0}) \cup \{v_1\} \subseteq R_{T-F}^+(v)$ . Hence

$$|R_{T-F}^+(v)| \geq |R_{T_1-F_1}^+(v_{i_0})| + 1 \geq n - i_0 + 2 - 2|F_1| \geq n - 2|F_0| - 2|F_1| = n - 2|F|.$$

$\square$

**Lemma 6.3.** *Let  $T$  be a tournament with median order  $(v_1, v_2, \dots, v_n)$ . For any  $i \in [n]$  and any  $F \subseteq V(T) \setminus \{v_i\}$ , we have  $|R_{T-F}^+(v_i)| \geq n + 1 - i - 2|F|$ , and  $R_{T-F}^-(v_i) \geq i - 2|F|$ .*

*Proof.* By (M1),  $(v_i, \dots, v_n)$  is a median order of  $T \langle \{v_i, \dots, v_n\} \rangle$  on which we can apply Lemma 6.2 to obtain the bound on  $R_{T-F}^+(v_i)$ . Symmetrically,  $(v_i, v_{i-1}, \dots, v_1)$  is a median order of the converse of  $T \langle \{v_1, \dots, v_i\} \rangle$ , and applying Lemma 6.2 to the converse of  $T$  yields the bound on  $R_{T-F}^-(v_i)$ .  $\square$

We are now ready to prove the main lemma of the proof of Theorem 1.13.



**Lemma 6.4.** *Let  $k$  be a positive integer, let  $T$  be a tournament on  $12k$  vertices and let  $(A, B)$  be a bipartition of  $V(T)$  such that  $|A| = |B| = 6k$ . Then there is a set  $X \subseteq V(T)$  with  $|X \cap A| = |X \cap B| = 2k$  such that for  $T' = \text{Inv}(T; X)$  and for any  $Y \subseteq V(T)$  with  $|Y| \leq k - 1$ , we have that  $T' - Y$  contains a directed path from  $a$  to  $B \setminus Y$  for every  $a \in A \setminus Y$ , and  $T' - Y$  contains a directed path from  $A \setminus Y$  to  $b$  for every  $b \in B \setminus Y$ .*

*Proof.* Let  $(a_1, \dots, a_{6k})$  be a median order of  $T\langle A \rangle$  and let  $(b_1, \dots, b_{6k})$  be a median order of  $T\langle B \rangle$ . Let  $A_0$  be the set of vertices in  $\{a_{4k+1}, \dots, a_{6k}\}$  which have less than  $k$  out-neighbours in  $B$  in  $T$ . Further, let  $A_1 = \{a_1, \dots, a_{2k-|A_0|}\}$ . Observe that  $|A_0 \cup A_1| = |A_0| + |A_1| = 2k$ . Similarly, let  $B_0$  be the set of vertices in  $\{b_1, \dots, b_{2k}\}$  that have less than  $k$  in-neighbours in  $A$  and let  $B_1 = \{b_{4k+|B_0|+1}, \dots, b_{6k}\}$ . Observe that  $|B_0 \cup B_1| = |B_0| + |B_1| = 2k$ . Let  $X = A_0 \cup A_1 \cup B_0 \cup B_1$  and let  $T' = \text{Inv}(T; X)$ . Consider any  $Y \subseteq V(T)$  with  $|Y| \leq k - 1$ . In order show that  $T'$  has the desired properties, by symmetry, it suffices to prove that  $T' - Y$  contains a directed path from  $a$  to  $B \setminus Y$  for every  $a \in A \setminus Y$ . Suppose by contradiction that this is not true. Then there is a largest integer  $i \in [6k]$  such that  $a_i \in A \setminus Y$  and  $T' - Y$  does not contain a directed path from  $a_i$  to  $B \setminus Y$ . We will distinguish several cases.

**Case 1.**  $i \in \{4k + 1, \dots, 6k\}$  and  $a_i \in A \setminus A_0$ .

In this case, by the choice of  $A_0$ , we have

$$\begin{aligned} |(N_{T'}^+(a_i) \cap B) \setminus Y| &\geq |N_{T'}^+(a_i) \cap B| - |Y| \\ &= |N_T^+(a_i) \cap B| - |Y| \\ &\geq k - (k - 1) \\ &= 1, \end{aligned}$$

so  $a_i$  has an out-neighbour in  $B \setminus Y$  in  $T' - Y$ , a contradiction.

**Case 2.**  $i \in \{4k + 1, \dots, 6k\}$  and  $a_i \in A_0$ .

In this case, by the choice of  $A_0$ , we have

$$\begin{aligned} |(N_{T'}^+(a_i) \cap B) \setminus Y| &\geq |N_{T'}^+(a_i) \cap B| - |Y| \\ &\geq |N_{T'}^+(a_i) \cap (B_0 \cup B_1)| - |Y| \\ &= |B_0 \cup B_1| - |N_{T'}^-(a_i) \cap (B_0 \cup B_1)| - |Y| \\ &= |B_0 \cup B_1| - |N_T^+(a_i) \cap (B_0 \cup B_1)| - |Y| \\ &\geq |B_0 \cup B_1| - |N_T^+(a_i) \cap B| - |Y| \\ &\geq 2k - (k - 1) - (k - 1) \\ &= 2, \end{aligned}$$

so  $a_i$  has an out-neighbour in  $B \setminus Y$  in  $T' - Y$ , a contradiction.

**Case 3.**  $i \in \{2k - |A_0| + 1, \dots, 4k\}$ .

As  $(a_1, \dots, a_{6k})$  is a median order of  $T\langle A \rangle$ , and by (M2) applied to  $T\langle A \rangle$ , we have

$$\begin{aligned} |(N_{T'}^+(a_i) \cap \{a_{i+1}, \dots, a_{6k}\}) \setminus Y| &\geq |N_{T'}^+(a_i) \cap \{a_{i+1}, \dots, a_{6k}\}| - |Y| \\ &\geq |N_T^+(a_i) \cap \{a_{i+1}, \dots, a_{6k}\}| - |Y| \\ &\geq \frac{1}{2}|\{a_{i+1}, \dots, a_{6k}\}| - |Y| \\ &\geq \frac{1}{2}(6k - i) - (k - 1) \\ &= 2k + 1 - \frac{i}{2} \\ &\geq 2k + 1 - 2k \\ &= 1. \end{aligned}$$

Hence there is some  $j > i$  such that  $a_j \in A \setminus Y$  and  $T' - Y$  contains the arc  $a_i a_j$ . By the maximality of  $i$ , there is a directed path from  $a_j$  to  $B \setminus Y$  in  $T' - Y$ . Hence  $T' - Y$  also contains a directed path from  $a_i$  to  $B \setminus Y$ , a contradiction.

**Case 4.**  $i \in \{1, \dots, 2k - |A_0|\}$ .

As  $(a_1, \dots, a_{6k})$  is a median order of  $T\langle A \rangle$  and by (M2) applied to  $T\langle A \rangle$ , we have

$$\begin{aligned}
|(N_{T'}^+(a_i) \cap \{a_{i+1}, \dots, a_{6k}\}) - Y| &\geq |N_{T'}^+(a_i) \cap \{a_{i+1}, \dots, a_{6k}\}| - |Y| \\
&\geq |N_T^+(a_i) \cap \{a_{i+1}, \dots, a_{6k}\}| - |(A_0 \cup A_1) \cap \{a_{i+1}, \dots, a_{6k}\}| - |Y| \\
&\geq \frac{1}{2}|\{a_{i+1}, \dots, a_{6k}\}| - |(A_0 \cup A_1) \cap \{a_{i+1}, \dots, a_{6k}\}| - |Y| \\
&\geq \frac{1}{2}(6k - i) - (2k - i) - (k - 1) \\
&= \frac{i}{2} + 1 \\
&\geq 1.
\end{aligned}$$

Hence there is some  $j > i$  such that  $a_j \in A \setminus Y$  and  $T' - Y$  contains the arc  $a_i a_j$ . By the maximality of  $i$ , there is a directed path from  $a_j$  to  $B \setminus Y$  in  $T' - Y$ . Hence  $T' - Y$  also contains a directed path from  $a_i$  to  $B \setminus Y$ , a contradiction.  $\square$

We are now ready to prove Theorem 1.13 which states  $N_k(1) \leq 19k - 2$ .

*Proof of Theorem 1.13.* Assume  $n \geq 19k - 2$ . Let  $(v_1, \dots, v_n)$  be a median order of  $T$ . Let  $B = \{v_1, \dots, v_{6k}\}$ ,  $A = \{v_{n-6k+1}, \dots, v_n\}$  and  $C = V(T) \setminus (A \cup B)$ . We have  $|C| \geq 7k - 2$ . By Lemma 6.4, applied to  $T\langle A \cup B \rangle$ , there is a subset  $X$  of  $A \cup B$  such that, for  $T' = \text{Inv}(T; X)$ , we have

- (i)  $|X \cap B| = |X \cap A| = 2k$ ;
- (ii) for any  $Y \subseteq V(T)$  with  $|Y| \leq k - 1$ ,  $T' - Y$  contains a directed path from  $a$  to  $B \setminus Y$  for every  $a \in A \setminus Y$ ;
- (iii) for any  $Y \subseteq V(T)$  with  $|Y| \leq k - 1$ ,  $T' - Y$  contains a directed path from  $A \setminus Y$  to  $b$  for every  $b \in B \setminus Y$ .

Let  $T_1 = \text{Inv}(T; X)$ . Let us now prove that  $T_1$  is  $k$ -strong, which implies  $\text{sinv}_k(T) \leq 1$ . Let  $F$  be a set of at most  $k - 1$  vertices of  $T_1$ . We first need the following intermediate result.

**Claim 6.4.1.** *Let  $b \in B \setminus F$  and  $a \in A \setminus F$ . Then  $T_1 - F$  contains a directed path from  $b$  to  $a$ .*

*Proof of claim.* By construction, there is some  $i \in \{1, \dots, 6k\}$  such that  $b = v_i$ . Let  $T' = T\langle \{v_i, \dots, v_{n-6k}\} \rangle$ . By (M2), every vertex in  $((X \cap B) \cup F) \cap \{v_i, \dots, v_{n-6k}\}$  and every vertex in  $C \setminus F$  is reachable from  $b$  in  $T'$ . On the other hand, there is obviously no vertex in  $((X \cap B) \cup F) \cap \{v_i, \dots, v_{n-6k}\}$  reachable from  $b$  in  $T' - ((X \cap B) \cup F)$ . By Lemma 6.3, this yields

$$\begin{aligned}
2|((X \cap B) \cup F) \cap \{v_i, \dots, v_{n-6k}\}| &\geq |R_{T'}^+(b)| - |R_{T' - ((X \cap B) \cup F)}^+(b)| \\
&= |R_{T'}^+(b) \cap B| - |R_{T' - ((X \cap B) \cup F)}^+(b) \cap B| \\
&\quad + |R_{T'}^+(b) \cap (C \cap F)| - |R_{T' - ((X \cap B) \cup F)}^+(b) \cap (C \cap F)| \\
&\quad + |R_{T'}^+(b) \cap (C \setminus F)| - |R_{T' - ((X \cap B) \cup F)}^+(b) \cap (C \setminus F)| \\
&\geq |((X \cap B) \cup F) \cap \{v_i, \dots, v_{n-6k}\}| \\
&\quad + |C \setminus F| - |R_{T' - ((X \cap B) \cup F)}^+(b) \cap (C \setminus F)|.
\end{aligned}$$

For every  $v \in C \setminus F$ , if  $v$  is reachable from  $b$  in  $T' - ((X \cap B) \cup F)$ , then  $v$  is clearly reachable from  $b$  in  $T_1 - F$ . Since  $|X \cap B| \leq 2k$  and  $|F| \leq k - 1$ , we obtain

$$\begin{aligned}
|R_{T_1-F}^+(b) \cap (C \setminus F)| &\geq |R_{T' - ((X \cap B) \cup F)}^+(b) \cap (C \setminus F)| \\
&\geq |C \setminus F| - |((X \cap B) \cup F) \cap \{v_i, \dots, v_{n-6k}\}| \\
&\geq |C \setminus F| - (|X \cap B| + |F|) \\
&\geq |C \setminus F| - (2k + (k - 1)) \\
&= |C \setminus F| - (3k - 1).
\end{aligned}$$

A similar argument shows that  $|R_{T_1-F}^-(a) \cap (C \setminus F)| \geq |C \setminus F| - (3k - 1)$ . As  $|C| \geq n - 12k \geq 7k - 2$ , we obtain

$$\begin{aligned}
|(R_{T_1-F}^+(b) \cap (C \setminus F)) \cap (R_{T_1-F}^-(a) \cap (C \setminus F))| &= |R_{T_1-F}^+(b) \cap (C \setminus F)| + |R_{T_1-F}^-(a) \cap (C \setminus F)| \\
&\quad - |(R_{T_1-F}^+(b) \cap (C \setminus F)) \cup (R_{T_1-F}^-(a) \cap (C \setminus F))| \\
&\geq |R_{T_1-F}^+(b) \cap (C \setminus F)| + |R_{T_1-F}^-(a) \cap (C \setminus F)| - |C \setminus F| \\
&\geq 2(|C \setminus F| - (3k - 1)) - |C \setminus F| \\
&= |C \setminus F| - 2(3k - 1) \\
&\geq |C| - |F| - (6k - 2) \\
&\geq (7k - 2) - (k - 1) - (6k - 2) \\
&\geq 1.
\end{aligned}$$

Hence there is a vertex  $v^* \in \{v_{6k+1}, \dots, v_{n-6k}\} \cap (R_{T_1-F}^+(b) \cap (C \setminus F)) \cap (R_{T_1-F}^-(a) \cap (C \setminus F))$ . By definition,  $T_1 - F$  contains a directed path from  $b$  to  $v^*$  and a directed path from  $v^*$  to  $a$ . Hence  $T_1 - F$  contains a directed path from  $b$  to  $a$ .  $\diamond$

We are now ready to show that  $T_1 - F$  is strong. Let  $x$  and  $y$  be two vertices in  $T_1 - F$ . It suffices to show that  $y$  is reachable from  $x$  in  $T_1 - F$ .

We first show that there is a path from  $x$  to  $B \setminus F$ . Clearly, we may suppose that  $x \in (A \cup C) \setminus F$ . Then, since  $|A| > 2|F|$ , Lemma 6.2 implies that there is a vertex  $x' \in A \setminus F$  reachable from  $x$  in  $T - F - B$ , and so in  $T_1 - F$ . By (iii), there is a directed path from  $x'$  to a vertex  $u$  in  $B \setminus F$  in  $T_1 - F$ . Hence there is a directed path  $P_x$  from  $x$  to  $u$  in  $T_1 - F$ . Similarly, by directional duality, in  $T_1 - F$ , there is a directed path  $P_y$  from a vertex  $w \in A \setminus F$  to  $y$ .

Finally, by Claim 6.4.1, there exists a path  $Q$  from  $u$  to  $w$  in  $T_1 - F$ . Then  $P_x Q P_y$  is a path from  $x$  to  $y$  in  $T_1 - F$ . This proves that  $T_1 - F$  is strong.  $\square$

Finally, we give the proof of Proposition 1.14, which states  $N_k(1) \geq 5k - 2$ .

*Proof of Proposition 1.14.* Let  $T$  be a tournament of order  $5k - 3$  whose vertex set has a partition  $(A, B, C)$  such that  $T\langle A \rangle$  and  $T\langle C \rangle$  are  $(k - 1)$ -diregular tournaments of order  $2k - 1$ , and  $A \Rightarrow B \cup C$  and  $B \Rightarrow C$ . We shall prove that  $\text{sinv}'_k(T) > 1$ .

Assume for a contradiction that there is a set  $X$  of vertices such that  $T' = \text{Inv}(T; X)$  is  $k$ -strong. Every vertex of  $A$  (resp.  $C$ ) has in-degree (resp. out-degree)  $k - 1$  in  $T$ , and so belongs to  $X$ . Thus  $A \cup C \subseteq X$ , and so  $C \Rightarrow A$  in  $T'$ . Hence  $T' - B$  is not strong. Since  $|B| = k - 1$ ,  $T'$  is not  $k$ -strong, a contradiction.  $\square$

### 6.3 Better upper bound on $N_k(3)$

This section is dedicated to the proof of Theorem 1.15. The structure is analogous to the proof of Theorem 1.13 in Section 6.2. First we show a result which is very similar to Lemma 6.4.

**Lemma 6.5.** *Let  $k$  be a positive integer,  $T$  a tournament on  $8k$  vertices and  $(A, B)$  a bipartition of  $V(T)$  such that  $|A| = |B| = 4k$ . There is a family  $\mathcal{X}$  of three subsets of  $V(T)$  such that the following hold with  $T' = \text{Inv}(T; \mathcal{X})$ .*

- (i)  $T'\langle A \rangle = T\langle A \rangle$  and  $T'\langle B \rangle = T\langle B \rangle$ ;

(ii) for any  $Y \subseteq V(T)$  with  $|Y| \leq k - 1$ ,  $T' - Y$  contains a directed path from  $a$  to  $B \setminus Y$  for every  $a \in A \setminus Y$ ;

(iii) for any  $Y \subseteq V(T)$  with  $|Y| \leq k - 1$ ,  $T' - Y$  contains a directed path from  $A \setminus Y$  to  $b$  for every  $b \in B \setminus Y$ .

*Proof.* Let  $(a_1, \dots, a_{4k})$  be a median order of  $T\langle A \rangle$  and let  $(b_1, \dots, b_{4k})$  be a median order of  $T\langle B \rangle$ . Let  $A_0$  be the set of vertices in  $\{a_{2k+1}, \dots, a_{4k}\}$  which have less than  $k$  out-neighbours in  $B$  in  $T$ . Further, let  $A_1 = \{a_1, \dots, a_{2k-|A_0|}\}$ . Observe that  $|A_0 \cup A_1| = |A_0| + |A_1| = 2k$ . Similarly, let  $B_0$  be the set of vertices in  $\{b_1, \dots, b_{2k}\}$  that have less than  $k$  in-neighbours in  $A$  and let  $B_1 = \{b_{2k+|B_0|+1}, \dots, b_{4k}\}$ . Observe that  $|B_0 \cup B_1| = |B_0| + |B_1| = 2k$ . Now let  $X_1 = A_0 \cup A_1$ ,  $X_2 = B_0 \cup B_1$ ,  $X_3 = X_1 \cup X_2$ , and  $\mathcal{X} = (X_i)_{i \in [3]}$ . Let  $T' = \text{Inv}(T; \mathcal{X})$ . Observe that  $T'$  is obtained from  $T$  by reversing the arcs between  $X_1$  and  $X_2$ . In particular, (i) holds.

We only show (ii) as (iii) follows symmetrically. Let  $Y \subseteq V(T)$  with  $|Y| \leq k - 1$ . Suppose for the sake of a contradiction that (ii) does not hold. There is a largest integer  $i \in [4k]$  such that  $a_i \in A \setminus Y$  and  $T' - Y$  does not contain a directed path from  $a_i$  to  $B \setminus Y$ . We will distinguish several cases.

**Case 1.**  $i \in \{2k + 1, \dots, 4k\}$  and  $a_i \in A \setminus A_0$ .

In this case, by the choice of  $A_0$ , we have

$$\begin{aligned} |(N_{T'}^+(a_i) \cap B) \setminus Y| &\geq |N_{T'}^+(a_i) \cap B| - |Y| \\ &= |N_T^+(a_i) \cap B| - |Y| \\ &\geq k - (k - 1) \\ &= 1, \end{aligned}$$

so  $a_i$  has an out-neighbour in  $B \setminus Y$  in  $T' - Y$ , a contradiction.

**Case 2.**  $i \in \{2k + 1, \dots, 4k\}$  and  $a_i \in A_0$ .

In this case, by the choice of  $A_0$ , we have

$$\begin{aligned} |(N_{T'}^+(a_i) \cap B) \setminus Y| &\geq |N_{T'}^+(a_i) \cap B| - |Y| \\ &\geq |N_{T'}^+(a_i) \cap (B_0 \cup B_1)| - |Y| \\ &= |B_0 \cup B_1| - |N_{T'}^-(a_i) \cap (B_0 \cup B_1)| - |Y| \\ &= |B_0 \cup B_1| - |N_T^+(a_i) \cap (B_0 \cup B_1)| - |Y| \\ &\geq |B_0 \cup B_1| - |N_{T'}^-(a_i) \cap B| - |Y| \\ &\geq 2k - (k - 1) - (k - 1) \\ &= 2, \end{aligned}$$

so  $a_i$  has an out-neighbour in  $B \setminus Y$  in  $T' - Y$ , a contradiction.

**Case 3.**  $i \in [2k]$ .

As  $(a_1, \dots, a_{4k})$  is a median order of  $T\langle A \rangle = T'\langle A \rangle$ , by (M2) we have

$$\begin{aligned} |(N_{T'}^+(a_i) \cap \{a_{i+1}, \dots, a_{4k}\}) \setminus Y| &\geq |N_{T'}^+(a_i) \cap \{a_{i+1}, \dots, a_{4k}\}| - |Y| \\ &= |N_T^+(a_i) \cap \{a_{i+1}, \dots, a_{4k}\}| - |Y| \\ &\geq \frac{1}{2}|\{a_{i+1}, \dots, a_{4k}\}| - |Y| \\ &\geq \frac{1}{2}(4k - i) - (k - 1) \\ &= k + 1 - \frac{i}{2} \\ &\geq k + 1 - k \\ &= 1. \end{aligned}$$

Hence there is some  $j > i$  such that  $a_j \in A \setminus Y$  and  $T' - Y$  contains the arc  $a_i a_j$ . By the maximality of  $i$ , there is a directed path from  $a_j$  to  $B \setminus Y$  in  $T' - Y$ . Hence  $T' - Y$  also contains a directed path from  $a_i$  to  $B \setminus Y$ , a contradiction.  $\square$

**Theorem 1.15.** *For any positive integer  $k$ , we have  $N'_k(3) \leq N_k(3) \leq 11k - 2$ .*

*Proof.* Assume  $n \geq 11k - 2$ . Let  $(v_1, \dots, v_n)$  be a median order of  $T$ . Let  $B = \{v_1, \dots, v_{4k}\}$ ,  $A = \{v_{n-4k+1}, \dots, v_n\}$  and  $C = V(T) \setminus (A \cup B)$ . We have  $|C| \geq 3k - 2$ . By Lemma 6.5 applied to  $T\langle A \cup B \rangle$ , there is a family  $\mathcal{X}$  of three subsets of  $A \cup B$  such that for  $T' = \text{Inv}(T; \mathcal{X})$  we have

- (i)  $T'\langle A \rangle = T\langle A \rangle$  and  $T'\langle B \rangle = T\langle B \rangle$ ;
- (ii) for any  $Y \subseteq V(T)$  with  $|Y| \leq k - 1$ ,  $T' - Y$  contains a directed path from  $a$  to  $B \setminus Y$  for every  $a \in A \setminus Y$ ;
- (iii) for any  $Y \subseteq V(T)$  with  $|Y| \leq k - 1$ ,  $T' - Y$  contains a directed path from  $A \setminus Y$  to  $b$  for every  $b \in B \setminus Y$ .

Let  $T_1 = \text{Inv}(T; \mathcal{X})$ . Let us now prove that  $T_1$  is  $k$ -strong, which implies  $\text{sinv}_k(T) \leq 3$ . Note that  $T\langle A \cup C \rangle$ , as well as  $T\langle B \cup C \rangle$  are unchanged by the inversions. Let  $F$  be a set of at most  $k - 1$  vertices of  $T_1$ . Let us show that  $T_1 - F$  is strong. Let  $x$  and  $y$  be two vertices in  $T_1 - F$ . It suffices to show that  $y$  is reachable from  $x$  in  $T_1 - F$ .

Let us first show that there is a vertex  $u \in B \setminus F$  that is reachable from  $x$  in  $T_1 - F$ . It is trivial if  $x \in B$ , so we may suppose that  $x \in A \cup C$ . Lemma 6.2 asserts that from any vertex of  $(A \cup C) \setminus F$  one can reach a vertex of  $A \setminus F$  in  $T_1 - F$ . Thus there exists a vertex  $x' \in A \setminus F$  reachable from  $x$  in  $T_1 - F$ . By (iii), there is a directed path from  $x'$  to a vertex of  $u$  in  $B \setminus F$  in  $T' - F$ . Hence there is a directed path  $P_x$  from  $x$  to  $u$  in  $T_1 - F$ . Similarly, by directional duality, in  $T_1 - F$ , there is a directed path  $P_y$  from a vertex  $w \in A \setminus F$  to  $y$ .

By Lemma 6.2,  $|R_{T\langle B \cup C \rangle - F}^+(u) \cap (C \setminus F)| \geq |C| - 2|F| = |C \setminus F| - |F| > |C \setminus F|/2$  because  $|C| \geq 3k - 2$ . Similarly, we obtain  $|R_{T\langle A \cup C \rangle - F}^-(w) \cap (C \setminus F)| > |C \setminus F|/2$ . Thus there is a vertex of  $C \setminus F$  in  $R_{T\langle B \cup C \rangle - F}^+(u) \cap R_{T\langle A \cup C \rangle - F}^-(w)$ , and so there exists a directed path  $Q$  from  $u$  to  $w$  in  $T_1 - F$ . Then  $P_x Q P_y$  is a directed path from  $x$  to  $y$  in  $T_1 - F$ .  $\square$

All the results of the previous subsections imply the following.

**Corollary 6.6.**  $m_k(n) \leq \begin{cases} 2k & \text{if } 2k + 1 \leq n < 11k - 2, \\ 3 & \text{if } 11k - 2 \leq n < 19k - 2, \\ 1 & \text{if } n \geq 19k - 2. \end{cases}$

## 6.4 Upper bounds for $k$ large.

In this section, we show that if a tournament has at least  $2k + 1 + \epsilon k$  vertices for some positive integer  $k$  and some  $\epsilon > 0$ , then it can be made  $k$ -strong by inverting a family of sets whose cardinality only depends on  $\epsilon$ . The proof consists in drawing this family uniformly at random, under the constraint that every vertex is contained in at least one of the sets.

To analyse this procedure we will need Chernoff's Bound (Proposition 2.6) as well as the two following technical lemmas.

**Lemma 6.7.** *Let  $\vec{u} \neq \vec{v} \in \mathbb{F}_2^t \setminus \{\vec{0}\}$  and  $x, y \in \mathbb{F}_2$  be fixed, and let  $\vec{w} \in \mathbb{F}_2^t \setminus \{\vec{0}\}$  be drawn uniformly at random. Then  $\Pr[\vec{u} \cdot \vec{w} = x, \vec{v} \cdot \vec{w} = y] \geq \frac{1}{4} - \frac{3}{4} \frac{1}{2^{t-1}}$ .*

*Proof.* As  $\vec{u} \neq \vec{v}$ , the mapping  $\mathbb{F}_2^t \rightarrow \mathbb{F}_2^2, \vec{w} \mapsto (\vec{u} \cdot \vec{w}, \vec{v} \cdot \vec{w})$  is surjective and linear. As a consequence, there are  $\frac{1}{4} 2^t$  vectors  $\vec{w} \in \mathbb{F}_2^t$  which satisfy  $\vec{u} \cdot \vec{w} = x$  and  $\vec{v} \cdot \vec{w} = y$ . Thus by possibly removing the solution  $\vec{w} = 0$ , we obtain  $\Pr[\vec{u} \cdot \vec{w} = x, \vec{v} \cdot \vec{w} = y] \geq \frac{2^{t-2}-1}{2^t-1} = \frac{1}{4} - \frac{3}{4} \frac{1}{2^{t-1}}$ .  $\square$

**Lemma 6.8.** *Let  $\epsilon > 0$ , let  $t \geq 16$  be an integer, and let  $k \geq \frac{8t}{\epsilon}$  be an integer. Let  $U, V \in (\mathbb{F}_2^t \setminus \{\vec{0}\})^{\lceil \epsilon k / 8 \rceil}$  be drawn uniformly at random and  $W \in \mathbb{F}_2^{\lceil \epsilon k / 8 \rceil \times \lceil \epsilon k / 8 \rceil}$  be fixed. Then  $\Pr[U^\top \cdot V = W] \leq 2^{-t\epsilon k / 128}$ .*

*Proof.* Note that since  $k \geq \frac{8t}{\epsilon}$ , we have  $\lceil \epsilon k/8 \rceil \geq t \geq t/2 + 1$ . First we bound the probability that  $\text{rk}(U) < t/2 + 1$ . If  $U$  has rank at most  $t/2$ , then there is a choice of  $\lfloor t/2 \rfloor$  columns of  $U$  such that all the other ones are in the linear span of these selected columns. Since the linear span of  $\lfloor t/2 \rfloor$  vectors has dimension at most  $t/2$ , and so size at most  $2^{t/2}$ , we deduce the following.

$$\begin{aligned}
\Pr[\text{rk}(U) \leq t/2] &\leq \binom{\lceil \epsilon k/8 \rceil}{\lfloor t/2 \rfloor} \left( \frac{2^{\lfloor t/2 \rfloor} - 1}{2^t - 1} \right)^{\lceil \epsilon k/8 \rceil - \lfloor t/2 \rfloor} \\
&\leq \binom{\lceil \epsilon k/8 \rceil}{\lfloor t/2 \rfloor} \left( \frac{2^{t/2} - 1}{2^t - 1} \right)^{\epsilon k/8 - t/2} \\
&\leq 2^{\epsilon k/8 + 1} \left( \frac{2^{t/2} - 1}{2^t - 1} \right)^{\epsilon k/16} && \text{because } \frac{t}{2} \leq \frac{\epsilon k}{16} \text{ since } k \geq \frac{8t}{\epsilon} \\
&\leq 2^{\epsilon k/4} (2^{t/2})^{-\epsilon k/16} \\
&\leq 2^{\epsilon kt/64} 2^{-\epsilon kt/32} && \text{because } t \geq 16 \\
&= 2^{-\epsilon kt/64}
\end{aligned}$$

Now we assume that  $\text{rk}(U) > t/2$ . Then for every column  $v$  of  $V$ ,  $v$  must be chosen in an affine space of dimension at most  $t - \lfloor t/2 \rfloor - 1 \leq t/2$ . It follows that

$$\begin{aligned}
\Pr[U^\top \cdot V = W \mid \text{rk}(U) > t/2] &\leq \left( \frac{2^{t/2} - 1}{2^t - 1} \right)^{\lceil \epsilon k/8 \rceil} \\
&\leq (2^{-t/2})^{\epsilon k/8} \\
&\leq 2^{-t\epsilon k/16}
\end{aligned}$$

Therefore

$$\begin{aligned}
\Pr[U^\top \cdot V = W] &\leq \Pr[\text{rk}(U) \leq t/2] + \Pr[U^\top \cdot V = W \mid \text{rk}(U) > t/2] \\
&\leq 2^{-t\epsilon k/64} + 2^{-t\epsilon k/16} \leq 2 \cdot 2^{-t\epsilon k/64}
\end{aligned}$$

We know that  $t\epsilon k \geq 8t^2 \geq 2 \cdot 64$ , thus  $2^{-t\epsilon k/64} \leq 1/4$ . As  $2x \leq \sqrt{x}$  for any  $x \in [0, 1/4]$ , we end with  $\Pr[U^\top \cdot V = W] \leq 2 \cdot 2^{-t\epsilon k/64} \leq 2^{-t\epsilon k/128}$ .  $\square$

For technical reasons, we prove the following seemingly weaker restatement of Theorem 1.16.

**Theorem 6.9.** *There exists a function  $f : \mathbb{R}_{>0} \rightarrow \mathbb{N}$  such that for every  $\epsilon > 0$  and every positive integer  $k$ , if  $T$  is an  $n$ -vertex tournament with  $n \geq 2k + 2\epsilon k + 2$ , then  $\text{sinv}_k(T) \leq f(\epsilon)$ .*

It is not difficult to see that Theorem 6.9 actually implies Theorem 1.16. Indeed, given a function  $f$  like in Theorem 6.9 at hand, define  $f' : \mathbb{R}_{>0} \rightarrow \mathbb{N}$  by  $f'(\epsilon) = \max\{\frac{4}{\epsilon}, f(\frac{\epsilon}{2})\}$ . Let  $T$  be a tournament with  $|V(T)| \geq 2k + 1 + \epsilon k$  for some positive integer  $k$ . If  $k \leq \frac{2}{\epsilon}$ , then Theorem 5.8 yields  $\text{sinv}_k(T) \leq 2k \leq \frac{4}{\epsilon} \leq f'(\epsilon)$ . Otherwise, we have  $|V(T)| \geq 2k + 1 + \epsilon k \geq 2k + 2 + \frac{\epsilon}{2}k$ , so  $\text{sinv}_k(T) \leq f(\frac{\epsilon}{2}) \leq f'(\epsilon)$  by Theorem 6.9.

*Proof.* Without loss of generality, we may assume  $\epsilon \leq \frac{1}{3}$ . Let  $C$  be a constant such that  $\text{sinv}_k(T) \leq 1$  if  $n \geq Ck$ , which exists by Theorem 1.13. Let  $t$  be the smallest integer such that

- $t \geq 16$ ,
- $t \geq \log(1 + \frac{48}{\epsilon})$ , and

- $t \geq \frac{128}{\epsilon}(2C + 2 + \epsilon/4) + 16$ .

Clearly,  $t$  is well defined and depends only on  $\epsilon$ . Let  $k_0(\epsilon)$  be the smallest integer such that for every  $k' \geq k_0(\epsilon)$

$$(2^t - 1) \exp\left(-\epsilon^2 \frac{(2 + \epsilon)k'}{24}\right) + 3(Ck')^2 \exp\left(-\epsilon^2 \frac{(2 + \epsilon)k'}{4096}\right) + 2^{-k'} < 1. \quad (1)$$

We now prove the statement for  $f(\epsilon) = \max\{t, 2k_0(\epsilon) - 2, \lceil \frac{16t}{\epsilon} \rceil - 2\}$ . If  $k < k_0(\epsilon)$ , then we conclude directly using Theorem 5.8 that  $\text{sinv}_k(T) \leq 2k \leq 2k_0(\epsilon) - 2 \leq f(\epsilon)$ . Similarly, if  $k \leq \frac{8t}{\epsilon} - 1$ , then we conclude by Theorem 5.8 that  $\text{sinv}_k(T) \leq \frac{16t}{\epsilon} - 2 \leq f(\epsilon)$ . Moreover, if  $n \geq Ck$ , we have  $\text{sinv}_k(T) \leq 1 \leq f(\epsilon)$ . Henceforth, we may assume  $k \geq \max\{k_0(\epsilon), \frac{8t}{\epsilon} - 1\}$  and  $n \leq Ck$ .

For every vertex  $u \in V(T)$ , we choose uniformly and independently at random a vector  $\vec{u} \in \mathbb{F}_2^t \setminus \{\vec{0}\}$ . For  $i \in [t]$ , let  $X_i = \{u \in V(T) \mid \vec{u}_i = 1\}$ . We will prove that with positive probability, the tournament  $T' = \text{Inv}(T; X_1, \dots, X_t)$  is  $k$ -strong. Note that for every arc  $uv \in A(T)$ , we have  $uv \in A(T')$  if and only if  $\vec{u} \cdot \vec{v} = 0 \pmod 2$ . For two disjoint subsets  $A$  and  $B$  of vertices of  $T'$ , a **directed**  $(A, B)$ -**matching** is a set of arcs with tails in  $A$ , heads in  $B$ , and without common tail or common head.

**Claim 6.9.1.** *If  $T'$  is not  $k$ -strong, then at least one of the following events occurs:*

$E_1$  : there is a vector  $z \in \mathbb{F}_2^t \setminus \{\vec{0}\}$  such that  $|\{v \in V(T) \mid \vec{v} \neq z\}| \leq k$ ,

$E_2$  : there are  $u, v \in V(T)$  with  $\vec{u} \neq \vec{v}$  such that  $\min\{|N_{T'}^+(u) \cap N_{T'}^-(v)|, |N_{T'}^+(u) \cap N_{T'}^+(v)|, |N_{T'}^-(u) \cap N_{T'}^-(v)|\} \leq (1 + \epsilon/4)\frac{k}{2}$ ,

$E_3$  : there are disjoint sets  $A, B \subseteq V(T')$  with  $|A|, |B| \geq (1 + \epsilon/4)\frac{k}{2}$  with no directed  $(A, B)$ -matching of size  $\frac{k}{2}$ .

*Proof of claim.* Assume that none of  $E_1, E_2$  and  $E_3$  holds. Suppose for a contradiction that there is a set  $X$  of at most  $k - 1$  vertices, and a partition  $(V_1, V_2)$  of  $V(T' - X)$  into nonempty sets such that  $V_2 \Rightarrow V_1$  in  $T' - X$ . Since  $E_1$  does not hold, there exist  $x, y \in V_1 \cup V_2$  with  $\vec{x} \neq \vec{y}$ . If both  $x$  and  $y$  are in  $V_1$  (resp.  $V_2$ ), consider  $v \in V_2$  (resp.  $u \in V_1$ ) and either  $\vec{x} \neq \vec{v}$  or  $\vec{y} \neq \vec{v}$  (resp.  $\vec{x} \neq \vec{u}$  or  $\vec{y} \neq \vec{u}$ ). If  $x \in V_1$  and  $y \in V_2$  we set  $u = x$  and  $v = y$ , and if  $y \in V_2$  and  $x \in V_1$  we set  $u = y$  and  $v = x$ . In all cases, there are  $u \in V_1$  and  $v \in V_2$  with  $\vec{u} \neq \vec{v}$ .

Now, as  $E_2$  does not hold, we have  $|N_{T'}^+(u) \cap N_{T'}^-(v)|, |N_{T'}^+(u) \cap N_{T'}^+(v)|, |N_{T'}^-(u) \cap N_{T'}^-(v)| \geq (1 + \epsilon/4)\frac{k}{2}$ . Finally, as  $E_3$  does not hold, there is a directed  $(N_{T'}^+(u) \cap N_{T'}^+(v), N_{T'}^-(u) \cap N_{T'}^-(v))$ -matching  $M$  of size  $k/2$  in  $T'$ . For every arc  $e = xy \in M$ , observe that  $P_e = uxyv$  is a directed  $(u, v)$ -path in  $T'$ . Furthermore, for every  $x \in N_{T'}^+(u) \cap N_{T'}^-(v)$ , observe that  $P_x = u xv$  is a directed  $(u, v)$ -path in  $T'$ . This yields a collection of at least  $k$  internally vertex-disjoint  $(u, v)$ -paths in  $T'$ , a contradiction since every  $(u, v)$ -path meets  $X$  which has size at most  $k - 1$ .  $\diamond$

We will show that with high probability none of the events  $E_1, E_2$  and  $E_3$  occurs.

**Claim 6.9.2.**  $\Pr(E_1) \leq (2^t - 1) \exp(-\epsilon^2 n/24)$

*Proof of claim.* If  $\vec{c} \in \mathbb{F}_2^t \setminus \{\vec{0}\}$  is fixed, then  $Y_{\vec{c}} = |\{u \in V(T) \mid \vec{u} \neq \vec{c}\}|$  is a random variable having a binomial law with parameters  $n$  and  $1 - \frac{1}{2^{t-1}}$ . As  $\epsilon \leq \frac{1}{3}$  and  $t \geq 2$ , we have  $k \leq \frac{1}{2}n \leq \frac{2}{3} \cdot \frac{5}{8}n \leq (1 - \frac{1}{2^{t-1}})(1 - \epsilon/2)n$ . By Chernoff's Bound (Proposition 2.6), and because  $t \geq 2$ , we have

$$\begin{aligned} \Pr[\exists \vec{c} \in \mathbb{F}_2^t \setminus \{\vec{0}\}, Y_{\vec{c}} < k] &\leq \sum_{\vec{c} \in \mathbb{F}_2^t \setminus \{\vec{0}\}} \Pr[Y_{\vec{c}} < k] \\ &\leq \sum_{\vec{c} \in \mathbb{F}_2^t \setminus \{\vec{0}\}} \Pr\left[Y_{\vec{c}} < \left(1 - \frac{\epsilon}{2}\right) \left(1 - \frac{1}{2^{t-1}}\right) n\right] \\ &\leq (2^t - 1) \exp\left(-(\epsilon/2)^2 \left(1 - \frac{1}{2^{t-1}}\right) n/3\right) \\ &\leq (2^t - 1) \exp(-\epsilon^2 n/24), \end{aligned}$$

as claimed.  $\diamond$

**Claim 6.9.3.**  $\Pr(E_2) \leq 3 \exp\left(-\epsilon^2 \frac{n-2}{4096}\right)$

*Proof of claim.* Let  $u, v$  be distinct vertices and let  $A = N_{T'}^+(u) \cap N_{T'}^-(v)$ ,  $B = N_{T'}^+(u) \cap N_{T'}^+(v)$  and  $C = N_{T'}^-(u) \cap N_{T'}^-(v)$ . Let  $X \in \{A, B, C\}$ . Once  $\vec{u}$  and  $\vec{v}$  revealed, for every vertex  $w \neq u, v$ , let  $Y_w$  be a random variable with  $Y_w = 1$  if  $w \in X$ ,  $Y_w = 0$  otherwise. By Lemma 6.7,  $Y_w$  is a random variable following a Bernoulli distribution whose parameter is at least  $\frac{1}{4}\left(1 - \frac{3}{2^t-1}\right)$ . Further, the  $Y_w$  are mutually independent.

We now define a random variable  $X_w$  for every  $w \in V(T) \setminus \{u, v\}$  as follows. If  $Y_w = 0$ , set  $X_w = 0$ , and otherwise set  $X_w = 1$  with probability  $\frac{\frac{1}{4}\left(1 - \frac{3}{2^t-1}\right)}{\Pr[Y_w=1]}$  and  $X_w = 0$  otherwise, where the latter random experiments are executed independently. Observe that, as the  $Y_w$  are mutually independent, so are the  $X_w$ . Moreover,  $\Pr[X_w = 1] = \frac{1}{4}\left(1 - \frac{3}{2^t-1}\right)$ .

We have  $(1 + \epsilon/4)k/2 \leq \frac{1+\epsilon/4}{2+\epsilon}(n-2)/2 = \frac{1}{2}\left(1 - \frac{\epsilon/4}{1+\epsilon/2}\right)(n-2)/2 \leq \frac{1}{4}(1 - \epsilon/8)(n-2)$  since  $n \geq (2 + \epsilon)k + 2$ . Moreover, as  $t \geq \log\left(\frac{48}{\epsilon} + 1\right)$  we have  $\frac{1}{4}(1 - \epsilon/8) \leq \frac{1}{4}(1 - \epsilon/16)^2 \leq \frac{1}{4}\left(1 - \frac{3}{2^t-1}\right)\left(1 - \frac{\epsilon}{16}\right) = \left(\frac{1}{4} - \frac{3}{4(2^t-1)}\right)\left(1 - \epsilon/16\right)$ . Hence  $(1 + \epsilon/4)k/2 \leq (1 - \epsilon/16)\left(\frac{1}{4} - \frac{3}{4(2^t-1)}\right)(n-2)$ , and by Chernoff's Bound (Proposition 2.6)

$$\begin{aligned} \Pr[|X| \leq (1 + \epsilon/4)k/2] &\leq \Pr\left[|X| \leq \left(1 - \frac{\epsilon}{16}\right)\left(\frac{1}{4} - \frac{3}{4(2^t-1)}\right)(n-2)\right] \\ &\leq \exp\left(-(\epsilon/16)^2\left(\frac{1}{4} - \frac{3}{4(2^t-1)}\right)\frac{n-2}{2}\right) \\ &\leq \exp\left(-\epsilon^2 \frac{n-2}{4096}\right) \end{aligned}$$

since  $t \geq 5$  implies  $\frac{3}{2^t-1} \leq 1/8$ . Hence, by the Union Bound (Proposition 2.5),  $\Pr[\min\{|A|, |B|, |C|\} \leq (1 + \epsilon/4)\frac{k}{2}] \leq \sum_{X \in \{A, B, C\}} \Pr[|X| \leq (1 + \epsilon/4)\frac{k}{2}] \leq 3 \exp\left(-\epsilon^2 \frac{n-2}{4096}\right)$ .  $\diamond$

For two disjoint sets of vertices  $X, Y$  in  $T'$ , we denote by  $\mu_{T'}(X, Y)$  the size of a largest directed  $(X, Y)$ -matching in  $T'$ .

**Claim 6.9.4.**  $\Pr(E_3) \leq 2^{-k}$ .

*Proof of claim.* Let  $A, B \subseteq V(T')$  be disjoint sets of  $\lceil (1 + \epsilon/4)\frac{k}{2} \rceil$  vertices. We shall prove that with high probability there is a directed  $(A, B)$ -matching in  $T'$  of size at least  $\frac{k}{2}$ .

Let  $M$  be a maximal directed  $(A, B)$ -matching and let  $Y_A$  (resp.  $Y_B$ ) the the set of vertices in  $A$  (resp. in  $B$ ) incident to no arc of  $M$ . Then  $Y_B \Rightarrow Y_A$  in  $T'$  since  $M$  is maximal. Moreover, if  $|M| \leq k/2$ , then  $|Y_A|, |Y_B| \geq (1 + \epsilon/4)\frac{k}{2} - \frac{k}{2} = \frac{\epsilon}{8}k$ .

For every such  $Y_A \subseteq A, Y_B \subseteq B$ , we identify  $Y_A$  and  $Y_B$  with the matrices whose columns are the  $\vec{u}$  for  $u \in Y_A$  (resp.  $u \in Y_B$ ). We also denote by  $T(Y_A, Y_B)$  the  $|A| \times |B|$  matrix whose cell  $(u, v)$  equal 1 if and only if  $uv \in A(T)$ . Then observe that  $Y_B \Rightarrow Y_A$  in  $T'$  if and only if  $Y_B^\top \cdot Y_A = T(Y_B, Y_A)$ . By these observations, we have

$$\begin{aligned} \Pr[\mu_{T'}(A, B) < k] &\leq \Pr[\exists Y_A \subseteq A, |Y_A| \geq \frac{\epsilon}{8}k, \exists Y_B \subseteq B, |Y_B| \geq \frac{\epsilon}{8}k, Y_B \Rightarrow Y_A \text{ in } T'] \\ &\leq \Pr[\exists Y_A \subseteq A, |Y_A| \geq \frac{\epsilon}{8}k, \exists Y_B \subseteq B, |Y_B| \geq \frac{\epsilon}{8}k, Y_B^\top \cdot Y_A = T(Y_B, Y_A)] \\ &\leq 2^{2\lceil(1+\epsilon/4)k/2\rceil} 2^{-\epsilon k/128} \quad \text{by Lemma 6.8 and the Union Bound} \\ &\leq 2^{(1+\epsilon/4)k+2} 2^{-\epsilon k/128} \\ &\leq 2^{-(2C+1)k} \end{aligned}$$

as  $t \geq \frac{128}{\epsilon}(2C + 2 + \epsilon/4) + 16$  and  $\epsilon k/64 \geq \frac{1}{8}$ . It follows from the Union Bound (Proposition 2.5) that  $\Pr(E_3) \leq 2^{2n} 2^{-(2C+1)k} \leq 2^{-k}$  using the fact that  $n \leq Ck$ .  $\diamond$



We can now conclude using Claims 6.9.1, 6.9.2, 6.9.3 and 6.9.4 and the Union Bound:

$$\begin{aligned}
\Pr[T' \text{ not } k\text{-strong}] &\leq \Pr(E_1) + \Pr(E_2) + \Pr(E_3) \\
&\leq (2^t - 1) \exp\left(-\epsilon^2 \frac{n}{24}\right) + 3n^2 \exp\left(-\epsilon^2 \frac{n-2}{4096}\right) + 2^{-k} \\
&\leq (2^t - 1) \exp\left(-\epsilon^2 \frac{(2+\epsilon)k}{24}\right) + 3(Ck)^2 \exp\left(-\epsilon^2 \frac{(2+\epsilon)k}{4096}\right) + 2^{-k} \\
&< 1,
\end{aligned}$$

by (1). This proves that there exist  $X_1, \dots, X_t \subseteq V(T)$  such that  $T' = \text{Inv}(T; X_1, \dots, X_t)$  is  $k$ -strong.  $\square$

## 7 Conclusion

In this paper, we investigated problems of the form: given a digraph, what is the minimum number of inversions needed such that the resulting digraph has a prescribed connectivity property?

For the first part, it is tempting to understand better the asymptotic behaviour of  $\text{sinv}_k(n)$ . In particular, it would be interesting to see if an analogue of Theorem 1.1 exists for  $\text{sinv}_k$ .

**Problem 7.1.** Let  $\text{sinv}_k(n) = \max\{\text{sinv}_k(D) \mid D \text{ } k\text{-strengthenable digraph of order } n\}$ . Find a good lower and upper bounds on  $\text{sinv}_k(n)$ .

For the algorithmic part, one may wonder whether polynomial-time constant-factor-approximation algorithms for computing  $\text{sinv}_k$  or  $\text{sinv}'_k$  exist. Actually, we believe that this is not the case and that Theorems 1.3 and 1.5 can be strengthened in the following way:

**Conjecture 7.2.** Unless  $P=NP$ , for any positive integer  $k$  and any constant  $\alpha$ , there is no  $\alpha$ -approximation algorithm for computing  $\text{sinv}_k(D)$  or  $\text{sinv}'_k(D)$  given an oriented graph  $D$ .

It would further be interesting to understand the complexity when restricting to tournaments. While the complexity for fixed  $k$  is resolved by Corollary 1.17, the following question remains open:

**Problem 7.3.** What is the complexity of computing  $\text{sinv}_k(T)$  (resp.  $\text{sinv}'_k(T)$ ) for a given tournament  $T$  if  $k$  is part of the input?

Next, we established several bounds on  $M_k$  and  $M'_k$ . There is still a significant gap between the logarithmic lower bounds and the linear upper bounds, so it would be good to improve these. A first question is the following.

**Problem 7.4.** Are  $M'_k$  and  $M_k$  sublinear functions of  $k$ ?

Further, we proved a collection of bounds for  $N_k(1)$  and  $N_k(3)$ . It would be interesting to have stronger bounds on  $N_k(i)$  for an integers  $i$ . In particular, a tight bound for  $N_k(1)$  would be satisfying.

Finally, Theorem 1.16 states that every tournament that is a constant factor bigger than  $2k$  can be made  $k$ -strong by a constant number of inversions. It would be interesting to know if this result can be strengthened in the following way:

**Problem 7.5.** Is there an integer  $t$  such that for every  $\epsilon > 0$ , there is an integer  $k_0$  with the property that for every  $k \geq k_0$ , if  $T$  is a tournament on at least  $(2 + \epsilon)k$  vertices, then we have  $\text{sinv}_k(T) \leq t$ ?

Proposition 1.14 shows that such an integer would have to satisfy  $t \geq 2$ , so  $t = 2$  is the first open case. The analogous statement for  $\text{sinv}'_k$  is also open.

## References

- [ACN08] Nir Ailon, Moses Charikar, and Alantha Newman. Aggregating inconsistent information: Ranking and clustering. *Journal of the ACM*, 55(5), 2008.
- [AHH<sup>+</sup>22] Guillaume Aubian, Frédéric Havet, Florian Hörsch, Felix Klingelhofer, Nicolas Nisse, Clément Rambaud, and Quentin Vermande. Problems, proofs, and disproofs on the inversion number, 2022.
- [Alo06] Noga Alon. Ranking tournaments. *SIAM Journal on Discrete Mathematics*, 20:137–142, 2006.
- [APS<sup>+</sup>22] Noga Alon, Emil Powierski, Michael Savery, Alex Scott, and Elizabeth Wilmer. Invertibility of digraphs and tournaments, 2022.
- [BBBP10] Houmem Belkhechine, Moncef Bouaziz, Imed Boudabbous, and Maurice Pouzet. Inversion dans les tournois. *Comptes Rendus Mathématique*, 348(13-14):703–707, 2010.
- [BJdSH22] Jørgen Bang-Jensen, Jonas Costa Ferreira da Silva, and Frédéric Havet. On the inversion number of oriented graphs. *Discrete Mathematics & Theoretical Computer Science*, 23(2):# 8, 2022.
- [BJG09] Jørgen Bang-Jensen and Gregory Z. Gutin. *Digraphs: theory, algorithms and applications*. Springer-Verlag, London, 2009.
- [BJHK23] Jørgen Bang-Jensen, Florian Hörsch, and Matthias Kriesell. Complexity of (arc)-connectivity problems involving arc-reversals or deorientations, 2023.
- [BJJY23] Jørgen Bang-Jensen, Kasper S. Johansen, and Anders Yeo. Making a tournament  $k$ -strong. *Journal of Graph Theory*, 103(1):5–11, 2023.
- [BJY04] Jørgen Bang-Jensen and Anders Yeo. Making a tournament  $k$ -arc-strong by reversing or deorienting arcs. *Discrete Applied Mathematics*, 136(2):161–171, 2004. The 1st Cologne-Twente Workshop on Graphs and Combinatorial Optimization.
- [Cam59] Paul Camion. Chemins et circuits hamiltoniens des graphes complets. *Comptes rendus hebdomadaires des séances de l’académie des sciences*, 249(21):2151–2152, 1959.
- [CTY07] Pierre Charbit, Stéphan Thomassé, and Anders Yeo. The minimum feedback arc set problem is NP-hard for tournaments. *Combinatorics, Probability and Computing*, 16:1–4, 2007.
- [DS05] Irit Dinur and Samuel Safra. On the hardness of approximating minimum vertex cover. *Annals of Mathematics, Second Series*, 162(1):439–485, 2005.
- [Dur20] Olivier Durand de Gevigney. On Frank’s conjecture on  $k$ -connected orientations. *Journal of Combinatorial Theory, Series B*, 141:105–114, 2020.
- [ENSS95] Guy Even, Joseph Naor, Baruch Schieber, and Madhu Sudan. Approximating minimum feedback sets and multi-cuts in directed graphs. In Egon Balas and Jens Clausen, editors, *Integer Programming and Combinatorial Optimization, 4th International IPCO Conference, Copenhagen, Denmark, May 29-31, 1995, Proceedings*, volume 920 of *Lecture Notes in Computer Science*, pages 14–28. Springer, 1995.
- [ET75] Shimon Even and R. Endre Tarjan. Network flow and testing graph connectivity. *SIAM Journal on Computing*, 4(4):507–518, 1975.
- [FF10] Lester R. Ford and Delbert R. Fulkerson. *Flows in Networks*. Princeton University Press, USA, 2010.
- [Gab91] Harold N. Gabow. A matroid approach to finding edge connectivity and packing arborescences. In *Proceedings of the Twenty-Third Annual ACM Symposium on Theory of Computing, STOC ’91*, page 112–122, 1991.

- [Gal80] Zvi Galil. Finding the vertex connectivity of graphs. *SIAM Journal on Computing*, 9(1):197–199, 1980.
- [HT00] Frédéric Havet and Stéphan Thomassé. Median orders of tournaments: A tool for the second neighborhood problem and sumner’s conjecture. *Journal of Graph Theory*, 35(4):244–256, 2000.
- [IWG] ANR Digraph Inversion Working Group. On the diameter of the inversion graph. unpublished.
- [Kan92] Vido Kann. *On the Approximability of NP-complete Optimization Problems*. PhD thesis, Department of Numerical Analysis and Computing Science, Royal Institute of Technology, Stockholm, 1992.
- [Kar72] Richard M. Karp. Reducibility among combinatorial problems. In *Complexity of Computer Computations*, The IBM Research Symposia Series, pages 85–103. Plenum Press, New York, 1972.
- [KMS07] Claire Kenyon-Mathieu and Warren Schudy. How to rank with few errors. In *Proceedings of the Thirty-Ninth Annual ACM Symposium on Theory of Computing*, STOC ’07, page 95–103, 2007.
- [KS06] Zoltán Király and Zoltán Szigeti. Simultaneous well-balanced orientations of graphs. *Journal of Combinatorial Theory, Series B*, 96(5):684–692, 2006.
- [McK90] Brendan D. McKay. The asymptotic numbers of regular tournaments, eulerian digraphs and eulerian oriented graphs. *Combinatorica*, 10(4):367–377, 1990.
- [MR02] Michael Molloy and Bruce Reed. *Graph colouring and the probabilistic method*, volume 23. Springer Science & Business Media, 2002.
- [MS89] Yishay Mansour and Baruch Schieber. Finding the edge connectivity of directed graphs. *Journal of Algorithms*, 10(1):76–85, 1989.
- [NW60] Crispin St J.A. Nash-Williams. On orientations, connectivity and odd-vertex-pairings in finite graphs. *Canadian Journal of Mathematics*, 12:555–567, 1960.
- [Rob39] Herbert E. Robbins. A theorem on graphs, with an application to a problem on traffic control. *American Mathematical Monthly*, 46(5):281–283, 1939.
- [Sch78] Thomas J. Schaefer. The complexity of satisfiability problems. In *Proceedings of the tenth annual ACM symposium on Theory of computing*, pages 216–226, 1978.
- [Spe74] Joel H. Spencer. Random regular tournaments. *Periodica Mathematica Hungarica*, 5(2):105–120, 1974.
- [Tho15] Carsten Thomassen. Strongly 2-connected orientations of graphs. *Journal of Combinatorial Theory, Series B*, 110:67–78, 2015.
- [TSKP22] Bhalchandra Thatte, Hamza Si Kaddour, and Maurice Pouzet. On the boolean dimension of a graph and other related parameters. *Discrete Mathematics & Theoretical Computer Science*, 23, 2022.
- [Zuc07] David Zuckerman. Linear degree extractors and the inapproximability of max clique and chromatic number. *Theory of Computing*, 3(6):103–128, 2007.