Invertibility of digraphs and tournaments

Noga Alon∗ Emil Powierski† Michael Savery†‡ Alex Scott†§ Elizabeth Wilmer¶

23 December 2022

Abstract

For an oriented graph $D$ and a set $X \subseteq V(D)$, the inversion of $X$ in $D$ is the digraph obtained by reversing the directions of the edges of $D$ with both endpoints in $X$. The inversion number of $D$, $\text{inv}(D)$, is the minimum number of inversions which can be applied in turn to $D$ to produce an acyclic digraph. Answering a recent question of Bang-Jensen, da Silva, and Havet, we show that for each fixed $k \in \mathbb{N}$ the problem of deciding whether a tournament $T$ has $\text{inv}(T) \leq k$ is solvable in time $O(|V(T)|^2)$. This exponent is optimal for all $k$. On the other hand, we build on their work to prove their conjecture that for $k \geq 1$ the problem of deciding whether a general oriented graph $D$ has $\text{inv}(D) \leq k$ is NP-complete. We also construct oriented graphs with inversion number equal to twice their cycle transversal number, confirming another conjecture of Bang-Jensen, da Silva, and Havet, and we provide a counterexample to their conjecture concerning the inversion number of so-called ‘dijoin’ digraphs while proving that it holds in certain cases. Finally, we asymptotically solve the natural extremal question in this setting, improving on previous bounds of Belkhechine, Bouaziz, Boudabbous, and Pouzet to show that the maximum inversion number of an $n$-vertex tournament is $(1 + o(1))n$.

1 Introduction

In this paper we only consider digraphs without loops, digons, or parallel edges, for which we use the terms digraph and oriented graph interchangeably. For such a digraph $D = (V, E)$ and a set $X \subseteq V$, the inversion of $X$ in $D$ is the digraph obtained from $D$ by reversing the direction of the edges with both endpoints in $X$; we say that we invert $X$ in $D$. Given a family of sets $X_1, \ldots, X_k \subseteq V$, we can invert $X_1$ in $D$, then $X_2$ in the resulting digraph, and so on, noting that the final digraph produced by these inversions is independent of the order in which we perform them. If inverting $X_1, \ldots, X_k$ in turn

∗Department of Mathematics, Princeton University, Princeton, NJ, USA. Research supported in part by NSF grant DMS-2154082 and BSF grant 2018267.
†Mathematical Institute, University of Oxford, Oxford OX2 6GG, UK.
‡Heilbronn Institute for Mathematical Research, Bristol, UK.
§Research supported by EPSRC grant EP/X013642/1.
¶Department of Mathematics, Oberlin College, Oberlin, OH, 44074, USA.

Email: nalon@math.princeton.edu, {powierski, savery, scott}@maths.ox.ac.uk, ewilmer@oberlin.edu.
transforms $D$ into an acyclic digraph, then we say that these sets form a decycling family of $D$. We will refer to a set $X \subseteq V$ which forms a decycling family by itself as a decycling set. The inversion number of $D$, denoted $\text{inv}(D)$, is defined to be the minimum size of a decycling family of $D$, and for $k \in \mathbb{N}_0$ we say that $D$ is $k$-invertible if $\text{inv}(D) \leq k$.

The study of inversions began in Houmem Belkhechine’s PhD thesis \cite{5} and continued in \cite{6,7,22}, in which many foundational results were established. The present work is inspired by a recent paper of Bang-Jensen, da Silva, and Havet \cite{4} which studied a wide range of questions about invertibility, with an emphasis on those of an algorithmic or extremal nature. They also posed a host of interesting conjectures and problems, some of which we answer in this paper.

1.1 The inversion number of $k$-joins

The cornerstone of many of the conjectures made by Bang-Jensen, da Silva, and Havet in \cite{4} is the following ‘dijoin conjecture’. For oriented graphs $L$ and $R$, the dijoin $L \rightarrow R$ from $L$ to $R$ is the oriented graph consisting of vertex-disjoint copies of $L$ and $R$, with an edge $\overrightarrow{uv}$ for all $u \in V(L)$ and $v \in V(R)$.

**Conjecture 1** (\cite{4}). For oriented graphs $L$ and $R$ we have $\text{inv}(L \rightarrow R) = \text{inv}(L) + \text{inv}(R)$.

Noting that the left-hand side is certainly at most the right-hand side for all $L$ and $R$ and that the conjecture holds trivially if $\text{inv}(L) = 0$ or $\text{inv}(R) = 0$, Bang-Jensen, da Silva, and Havet showed it to be true when $\text{inv}(L) + \text{inv}(R) \leq 3$, and when $\text{inv}(L) = \text{inv}(R) = 2$ and $L$ and $R$ are both strongly connected.

They also observed (see our Section 3) that the conjecture is equivalent to its restriction to tournaments. We disprove Conjecture 1 by exhibiting a tournament $R$ with $\text{inv}(R) = \text{inv}(\overrightarrow{C_3} \rightarrow R)$, where $\overrightarrow{C_3}$ is the directed cycle on three vertices.

**Theorem 1.** There exists a tournament $R$ with $\text{inv}(R) = \text{inv}(\overrightarrow{C_3} \rightarrow R) = 3$.

While Theorem 1 shows that the dijoin conjecture is false in general, we prove it in the case where $\text{inv}(L) = \text{inv}(R) = 2$.

**Theorem 2.** If $L$ and $R$ are digraphs with $\text{inv}(L) = \text{inv}(R) = 2$, then $\text{inv}(L \rightarrow R) = 4$.

The proof of Theorem 2 relies on the strongly connected case and our next result, which concerns the following generalisation of dijoins to arbitrarily many digraphs. For $k \in \mathbb{N}$ the $k$-join of digraphs $D_1, \ldots, D_k$, written $[D_1, \ldots, D_k]$, is the digraph consisting of vertex-disjoint copies of $D_1, \ldots, D_k$ with an additional edge $\overrightarrow{uv}$ whenever $u \in V(D_i)$, $v \in V(D_j)$ for $i < j$. We write $[D]_k = [D, \ldots, D]$ for the $k$-join of $k$ copies of the same oriented graph $D$. The following result can be viewed as a $k$-join analogue of the dijoin conjecture holding under certain conditions. It generalises a theorem of Pouzet, Kaddour, and Thatte \cite{22} which states that $\text{inv}(\overrightarrow{C_3}_k) = k$ for all $k$.

---

\footnote{The case where $\text{inv}(L) = 2$ and $\text{inv}(R) = 1$ is not explicitly mentioned in \cite{4}, but follows easily from the case where $\text{inv}(L) = 1$ and $\text{inv}(R) = 2$ by inverting $V(L \rightarrow R)$.}
Theorem 3. Let \( k \in \mathbb{N} \) and let \( D_1, \ldots, D_k \) be oriented graphs. Assume that \( \text{inv}(D_i) \leq 2 \) for all \( i \), with equality for at most one \( i \). Then

\[
\text{inv}([D_1, \ldots, D_k]) = \sum_{i=1}^{k} \text{inv}(D_i).
\]

We will use Theorem 3 to confirm another conjecture from [4] which was made based on the dijoin conjecture (see Theorem 5 below). Theorem 3 and, in turn, Theorem 2 follow from a characterisation of the decycling families of size \( k \) of arbitrary \( k \)-joins of oriented graphs with inversion number 1. We will need some further terminology to state this result: for a digraph \( D \), sets \( X_1, \ldots, X_k \subseteq V(D) \), and a vertex \( v \in V(D) \), we define the characteristic vector of \( v \) in \( X_1, \ldots, X_k \) to be \((I\{v \in X_i\} : i \in [k]) \in \mathbb{F}_2^k\), where \( I\{v \in X_i\} \) is the indicator function of the event \( v \in X_i \). For vectors \( u, v \in \mathbb{F}_2^k \) we write \( u \cdot v \) for the usual scalar product of \( u \) and \( v \) over \( \mathbb{F}_2 \). This is not a genuine inner product, but we say nevertheless that a collection \( u_1, \ldots, u_k \in \mathbb{F}_2^k \) is orthonormal if \( u_i \cdot u_i = 1 \) for all \( i \) and \( u_i \cdot u_j = 0 \) for all \( i \neq j \). Finally, we refer to the canonical copy of \( D_i \) in \( D = [D_1, \ldots, D_k] \) as the \( i \)th factor of \( D \). We are now ready to state our characterisation theorem, the case \( k = 2 \) of which was shown by Bang-Jensen, da Silva, and Havet [4]. Its proof is based on an approach used by Pouzet, Kaddour, and Thatte [22].

Theorem 4. Let \( D_1, \ldots, D_k \) be oriented graphs with \( \text{inv}(D_i) = 1 \) for all \( i \) and let \( \hat{D} = [D_1, \ldots, D_k] \) be their \( k \)-join. Then sets \( X_1, \ldots, X_k \subseteq V(\hat{D}) \) form a decycling family of \( \hat{D} \) if and only if there are orthonormal vectors \( u_1, \ldots, u_k \in \mathbb{F}_2^k \) and for each \( i \) a decycling set \( V_i \subseteq V(D_i) \) of the \( i \)th factor of \( \hat{D} \) such that the vertices in \( V_i \) have characteristic vector \( u_i \) for each \( i \) and all other vertices have characteristic vector \( 0 \).

In particular, any acyclic digraph obtained from \( \hat{D} \) by \( k \) inversions can also be obtained by inverting a decycling set for each factor in turn.

1.2 Computational complexity

One focus of Bang-Jensen, da Silva, and Havet’s paper [4] was on the computational complexity of deciding whether an oriented graph is \( k \)-invertible. More formally, they considered, for fixed \( k \in \mathbb{N}_0 \), the problem of \( k \)-INVERSION:

**INPUT:** an oriented graph \( D 

**PROBLEM:** \( \text{inv}(D) \leq k \)

A first observation is that 0-INVERSION is equivalent to checking whether a digraph \( D \) is acyclic, which is well known to be possible in time \( O(|V(D)|^2) \) (see Section 22.4 of [11]), so we need only consider \( k \geq 1 \).

Bang-Jensen, da Silva, and Havet [4] showed that 1-INVERSION is NP-complete using a reduction from MONOTONE 3-IN-1 SAT. Then, using the special cases of the dijoin
conjecture proved in that paper, they observed that for a digraph $D$ we have $\text{inv}(D \to D) = 2$ if and only if $\text{inv}(D) = 1$, from which it follows that $2\text{-INVERSION}$ is also NP-complete. They conjectured that NP-completeness extends to $k\text{-INVERSION}$ for all $k \geq 3$, noting that this would follow from a similar argument if the dijoin conjecture were true. Of course, the full dijoin conjecture is not required, and indeed it is easy to see that Theorem 3 is enough: it implies that $\text{inv}(D) = 1$ if and only if $\text{inv}(|D|_k) = k$, which reduces $1\text{-INVERSION}$ to $k\text{-INVERSION}$ and hence shows the following (see Section 4).

**Theorem 5.** $k\text{-INVERSION}$ is NP-complete for all $k \in \mathbb{N}$.

Bang-Jensen, da Silva, and Havet also considered the computational complexity of the same problem when the input is restricted to tournaments. For fixed $k \in \mathbb{N}$ the problem of $k\text{-TOURNAMENT-INVERSION}$ is:

**Input:** a tournament $T$.

**Problem:** $\text{inv}(T) \leq k$?

One way of analysing the complexity of this problem is to use $k$-inversion-critical tournaments: we say that a tournament $T$ is a $k$-inversion-critical tournament if $\text{inv}(T) = k$ but $\text{inv}(T - \{v\}) < k$ for all $v \in V(T)$, and denote by $\mathcal{IC}_k$ the set of $k$-inversion-critical tournaments. It is not difficult to see that a tournament has inversion number at most $k$ if and only if it contains no element of $\mathcal{IC}_{k+1} \cup \mathcal{IC}_{k+2}$ as a subtournament. Indeed, for any digraph $D$ and vertex $v \in V(D)$ with out-neighbourhood $A \subseteq V(D)$, adding $A$ and $A \cup \{v\}$ to a decycling family of $D - \{v\}$ gives a decycling family of $D$. We deduce that $\text{inv}(D) \leq \text{inv}(D - \{v\}) + 2$. Hence, if $\text{inv}(T) > k$, then there is a subtournament $T'$ of $T$ with $\text{inv}(T') \in \{k + 1, k + 2\}$ and this contains a member of $\mathcal{IC}_{k+1} \cup \mathcal{IC}_{k+2}$ as a subtournament.

Belkhechine, Bouaziz, Boudabbous, and Pouzet [7] showed that $\mathcal{IC}_k$ is finite for all $k \in \mathbb{N}$. Writing $m_k$ for the maximum number of vertices of an element of $\mathcal{IC}_k$, it follows that $k\text{-TOURNAMENT-INVERSION}$ can be solved in time $O(|V(T)|^{\max(m_{k+1}, m_{k+2})})$. Thus, in particular, $k\text{-TOURNAMENT-INVERSION}$ can be solved in polynomial time for any fixed $k$. Plainly $\mathcal{IC}_1 = \{C_3\}$, so $m_1 = 3$, and $\mathcal{IC}_2$ was explicitly described in [22], giving $m_2 = 6$. However, no upper bound on $m_k$ is known for $k \geq 3$, so for no $k \geq 1$ does the above give a concrete polynomial bound on the complexity of $k\text{-TOURNAMENT-INVERSION}$. Note also that this approach does not identify a decycling family of size $k$ given a $k$-invertible tournament, it can only confirm the existence of one.

Bang-Jensen, da Silva, and Havet [4] used an alternative approach to show that $1\text{-TOURNAMENT-INVERSION}$ can be solved in time $O(|V(T)|^3)$ while $2\text{-TOURNAMENT-INVERSION}$ can be solved in time $O(|V(T)|^6)$. The idea behind their algorithm for $1\text{-TOURNAMENT-INVERSION}$ is to check whether the tournament contains a vertex which can be made into a source, and for $2\text{-TOURNAMENT-INVERSION}$ they check whether it contains a pair of vertices which can be made into a source and a sink respectively. They went on to ask for the least real numbers $r_k$ such that $k\text{-TOURNAMENT-INVERSION}$ can
be solved in time $O(|V(T)|^r)$. We answer this question by showing that, perhaps surprisingly, for each fixed $k \in \mathbb{N}$ there is an algorithm solving $k$-TOURNAMENT-INVERSION in time $O(|V(T)|^2)$. Note that this is optimal, since any algorithm solving $k$-TOURNAMENT-INVERSION needs to inspect the orientation of every edge in the input tournament.

**Theorem 6.** For fixed $k \in \mathbb{N}$, $k$-TOURNAMENT-INVERSION can be solved for $n$-vertex tournaments in time $O(n^2)$. Moreover, if the input tournament is $k$-invertible, then our algorithm finds a decycling family of size at most $k$.

The implied constant in the running time of our algorithm is doubly exponential in $k$, so it is unlikely to be of practical use for large $k$.

### 1.3 Relation to other parameters

Bang-Jensen, da Silva, and Havet [4] also considered the relationship between the inversion number and other digraph parameters. Two well studied parameters of particular interest are the cycle transversal number and the cycle edge-transversal number, defined as follows. A cycle transversal (or feedback vertex set) in a digraph $D$ is a set of vertices of $D$ whose removal from $D$ leaves an acyclic digraph and the cycle transversal number of $D$, denoted $\tau(D)$, is the minimum size of a cycle transversal in $D$. Analogously, a cycle edge-transversal (or feedback arc set) in $D$ is a set of edges of $D$ whose removal leaves an acyclic digraph and the cycle edge-transversal number of $D$, $\tau'(D)$, is the minimum size of a cycle edge-transversal in $D$. Note that the inequality $\tau(D) \leq 2 \tau'(D)$ always holds, since the endpoints of the edges in a cycle edge-transversal of $D$ form a cycle transversal of $D$.

Bang-Jensen, da Silva, and Havet [4] made the following observations concerning the relationships between $\text{inv}(D)$, $\tau(D)$, and $\tau'(D)$ for a digraph $D$. Firstly, we have $\text{inv}(D) \leq \tau'(D)$. This follows from the fact that if $F \subseteq E(D)$ is a cycle edge-transversal of $D$, then since $(V(D), E(D) \setminus F)$ is acyclic, there is a labelling $v_1, \ldots, v_n$ of $V(D)$ such that $\overrightarrow{v_j v_i} \notin E(D) \setminus F$ if $i < j$. Applying the family of inversions $(\{v_i, v_j\} : i < j, v_j v_i \in F)$ transforms $D$ into an acyclic digraph and hence $\text{inv}(D) \leq \tau'(D)$ as claimed. They also observed that this inequality is tight for all values of $\tau'(D)$ as exhibited by $[\overrightarrow{C_3}]_k$, which clearly has cycle edge-transversal number $k$, and as mentioned above was shown in [22] to have inversion number $k$.

Turning to $\tau(D)$, the inequality $\text{inv}(D) \leq 2 \tau(D)$ was obtained in [4] as follows. After observing that $\tau(D) = 0$ implies $\text{inv}(D) = 0$, we may assume that $\tau(D) \geq 1$. Let $S \subseteq V(D)$ be a cycle transversal in $D$ of size $\tau(D)$ and pick $v \in S$. Write $A$ for the out-neighbourhood of $v$ and consider inverting $A$ and then $A \cup \{v\}$ in $D$. It is not difficult to see that the digraph produced has cycle transversal number $\tau(D) - 1$, with $S \setminus \{v\}$ a cycle transversal. It follows by induction that $\text{inv}(D) \leq 2 \tau(D)$.

Bang-Jensen, da Silva, and Havet conjectured that this inequality is tight for all values of $\tau(D)$. Indeed, they considered the graph $V_5$ obtained by adding a vertex $v$ and edges $\overrightarrow{1v}, \overrightarrow{2v}, \overrightarrow{3v}, \overrightarrow{4v}$ to the (transitive) tournament on vertex set $\{1, 2, 3, 4\}$ with edges $\overrightarrow{ij}$ for $i < j$, which can easily be seen to have $\tau(V_5) = 1$ and $\text{inv}(V_5) = 2$. They noted that
if the dijoin conjecture holds, then $\tau([V_5]_k) = k$ and $\text{inv}([V_5]_k) = 2k$ for all $k$ (in fact, since $V_5$ is strongly connected, the case $k = 2$ follows from the special cases for which they proved the dijoin conjecture). We construct digraphs with a similar character to $V_5$ which confirm their conjecture.

**Theorem 7.** For all $k \in \mathbb{N}$ there exists an oriented graph $D$ with $\text{inv}(D) = 2\tau(D) = 2k$.

### 1.4 The extremal problem

Finally, we consider $\text{inv}(n)$, defined for each $n \in \mathbb{N}$ as the maximum inversion number of an oriented graph (or, equivalently, a tournament) on $n$ vertices. Belkhechine, Bouaziz, Boudabbous, and Pouzet [7] were the first to study this parameter, obtaining bounds of the form

$$\frac{n}{2} - \log(n) + O(1) \leq \text{inv}(n) \leq n + O(1).$$

Their lower and upper bounds follow from counting and inductive arguments respectively (see Section 7 for details), and they conjectured that $\text{inv}(n) \geq \left\lfloor \frac{n-1}{2} \right\rfloor$ for all $n$. Bounds of the form above previously remained the best known, with Bang-Jensen, da Silva, and Havet [4] noting that the $O(1)$ term in the upper bound can be improved very slightly.

Using a random construction, we show that $\text{inv}(n) = (1 + o(1))n$.

**Theorem 8.** For sufficiently large $n$ we have

$$\text{inv}(n) \geq n - \sqrt{2n \log(n)}.$$ 

Moreover, a uniformly random labelled $n$-vertex tournament has at least this inversion number with probability tending to 1.

In Section 7 we also show that $\text{inv}(n) \leq n - \log(n + 1)$.

### 1.5 Outline of the paper

The remainder of the paper is organised as follows. In Section 2 we introduce some further notation, definitions, and preliminary observations which will be useful in the rest of the paper. In the very short Section 3 we prove Theorem 1, constructing a counterexample to the dijoin conjecture. Our results on the inversion number of $k$-joins, Theorems 2, 3, and 4, are proved in Section 4, along with Theorem 5. Section 5 concerns the complexity of $k$-TOURNAMENT-INVERSION and contains the proof of Theorem 6. We give the proof of Theorem 7 in Section 6. In Section 7 we discuss the existing bounds on $\text{inv}(n)$ before proving Theorem 8 and giving an improved upper bound. Finally, in Section 8 we restate some conjectures and questions from previous papers which remain open and pose some new ones of our own.

**Note added.** Almost simultaneously with the release of this paper, Aubian, Havet, Hörsch, Klingelhoefer, Nisse, Rambaud, and Vermande announced independent work [2]

\[\text{All logarithms in this paper are taken base } 2.\]
on some of the problems we address here. Specifically, they prove a stronger version of Theorem 1 (in fact, they prove that a strong version of our Conjecture 9 holds provided at least one of \( \ell \) and \( r \) is odd and at least 3) and they show upper and lower bounds on \( \text{inv}(n) \) of forms similar to those we give in Section 1.4.

## 2 Notation and preliminaries

In this section we detail some of the definitions, observations, and notation to be used in the rest of the paper. As noted above, all digraphs will be oriented graphs, that is, loopless directed graphs with at most one edge between each pair of vertices. An *acyclic digraph* is a digraph with no directed cycles. In the case where the digraph is a tournament, we use the term *transitive* instead of acyclic. Note that for each \( n \in \mathbb{N} \) there is a unique unlabelled transitive tournament on \( n \) vertices. To a transitive tournament \( T \) we associate the total order \( < \) on \( V(T) \) where \( u < v \) for all \( u, v \in V(T) \) such that \( uv \in E(T) \). We write \( [n] \) for the set \( \{1, 2, \ldots, n\} \). For a digraph \( D \) and a set \( S \subseteq V(D) \) we write \( D - S \) for the digraph produced by deleting the vertices in \( S \) from \( D \). We now give the following key definitions.

**Definition 1.** Recall that for a digraph \( D \), sets \( X_1, \ldots, X_k \subseteq V(D) \), and a vertex \( v \in V(D) \), the *characteristic vector* of \( v \) in \( X_1, \ldots, X_k \) is \( (I_{\{v \in X_i\}} : i \in [k]) \in \mathbb{F}_2^k \), where \( I_{\{v \in X_i\}} \) is the indicator function of the event \( v \in X_i \). Define an equivalence relation \( \sim \) on \( V(D) \) by setting \( u \sim v \) if \( u \) and \( v \) have the same characteristic vector in \( X_1, \ldots, X_k \). The *atoms* of \( X_1, \ldots, X_k \) in \( D \) are the equivalence classes of this relation.

Note that, equivalently, the atoms of \( X_1, \ldots, X_k \) in \( D \) are the atoms of the set algebra on \( V(D) \) generated by \( X_1, \ldots, X_k \), and that there are at most \( 2^k \) atoms for given \( D \) and \( X_1, \ldots, X_k \). The next observation will be useful throughout the paper.

**Observation 1.** Let \( D \) be a digraph and suppose that \( u, v \in V(D) \) are joined by an edge in \( D \). Let \( X_1, \ldots, X_k \subseteq V(D) \). Write \( u, v \in \mathbb{F}_2^k \) for the characteristic vectors of \( u \) and \( v \) in \( X_1, \ldots, X_k \) respectively. Then the edge between \( u \) and \( v \) undergoes a net change in orientation when \( X_1, \ldots, X_k \) are inverted in \( D \) if and only if \( u \cdot v = 1 \).

This follows from the fact that \( u \cdot v \) is the parity of the number of \( X_1, \ldots, X_k \) which contain both \( u \) and \( v \). An obvious implication of Observation 1 is that given \( D \) and \( X_1, \ldots, X_k \), for every pair of (not necessarily distinct) atoms \( A \) and \( B \), either all edges \( \{ab : a \in A, b \in B\} \) undergo a net orientation change when \( X_1, \ldots, X_k \) are inverted, or none of them do. In particular, for every vertex \( v \) and atom \( A \), either all edges \( \{va : a \in A\} \) change orientation or none of them do.

Finally, we note some simple observations which will be used freely in what follows.

(i) If \( D' \) is a subdigraph of an oriented graph \( D \), then \( \text{inv}(D') \leq \text{inv}(D) \).

(ii) For every oriented graph \( D \) and every non-negative integer \( k \leq \text{inv}(D) \), there exists a spanning subdigraph of \( D \) with inversion number \( k \).

(iii) If \( X_1, \ldots, X_k \) is a decycling family of an oriented graph \( D \), then \( D \) can be extended to a tournament \( T \) for which \( X_1, \ldots, X_k \) is still a decycling family. In particular \( \text{inv}(T) = \text{inv}(D) \).
For (ii), delete edges of \( D \) one by one, noting that the inversion number drops by at most 1 at each step. For (iii), invert the decycling family in \( D \), extend the resulting acyclic digraph to a transitive tournament, then invert the decycling family again.

3 A counterexample to the dijoin conjecture

In this short section we give a counterexample to the dijoin conjecture of Bang-Jensen, da Silva, and Havet [4], that is, the conjecture that \( \text{inv}(L \to R) = \text{inv}(L) + \text{inv}(R) \) for all oriented graphs \( L \) and \( R \). As noted in the introduction, this conjecture is equivalent to its restriction to tournaments. Indeed, suppose that \( L \) and \( R \) are digraphs with \( \text{inv}(L \to R) < \text{inv}(L) + \text{inv}(R) \). Extend \( L \to R \) to a tournament of the same inversion number and observe that this tournament is \( L' \to R' \) for some tournaments \( L' \) and \( R' \) extending \( L \) and \( R \) respectively. These clearly satisfy \( \text{inv}(L') \geq \text{inv}(L) \) and \( \text{inv}(R') \geq \text{inv}(R) \), so we have tournaments \( L' \) and \( R' \) with \( \text{inv}(L' \to R') < \text{inv}(L) + \text{inv}(R) \leq \text{inv}(L') + \text{inv}(R') \).

**Proof of Theorem 1.** Let \( L \) be a copy of \( C_3^* \). Suppose that \( R \) is a tournament with \( \text{inv}(R) = 3 \) for which there exist \( A, B, C \subseteq V(R) \) such that \( A \cup B, A \cup C \) and \( B \cup C \) form a decycling family of \( R \). Then for distinct vertices \( u, v \in V(L) \) the sets \( A \cup B \cup \{u, v\}, A \cup C \cup \{u, v\} \) and \( B \cup C \cup \{u, v\} \) form a decycling family of \( L \to R \), demonstrating that \( \text{inv}(L \to R) = 3 < 4 = \text{inv}(L) + \text{inv}(R) \).

One way to construct such an \( R \) is as follows: let \( R \) be the tournament with vertex set [9], let \( A = \{1, 3\} \), \( B = \{4, 6\} \), and \( C = \{7, 9\} \), and let the edge \( ij \) be directed backwards (that is, from \( j \) to \( i \) when \( i < j \)) if and only if \( i \) and \( j \) are both in \( A \cup B \cup C \), but not both in \( A \), \( B \), or \( C \). It is clear that inverting \( A \cup B \), \( A \cup C \) and \( B \cup C \) transforms \( R \) into a transitive tournament, and a computer check shows that \( \text{inv}(R) = 3 \), as required. \( \square \)

4 Decycling families of \( k \)-joins

In this section we prove Theorem 4, which characterises the decycling families of size \( k \) of \( k \)-joins of digraphs each with inversion number 1. We will then deduce Theorem 3 from this characterisation, and use Theorem 3 to obtain Theorems 2 and 5. The bulk of the work in our proof of Theorem 4 is put towards proving Lemma 9, which deals with the case \( \hat{D} = [C_3^*]_k \).

**Lemma 9.** Let \( k \in \mathbb{N} \), let \( \hat{D} = [C_3^*]_k \), and let \( X_1, \ldots, X_k \subseteq V(\hat{D}) \) be a decycling family of \( \hat{D} \). Then there exist orthonormal vectors \( u_1, \ldots, u_k \in \mathbb{F}_2^k \) such that in the \( i \)th factor of \( [C_3^*]_k \), one vertex has characteristic vector \( 0 \) and the other two have characteristic vector \( u_i \).

We will use the setup that Pouzet, Kaddour, and Thatte [22] introduced in their proof that \( \text{inv}(\hat{D}) = k \). The first part of our argument is essentially a reformulation of theirs, but we include it for completeness and to build intuition.
Proof of Lemma 9. Let \( T \) be the transitive tournament obtained by inverting the sets \( X_1, \ldots, X_k, \) in \( \hat{D} \), and let \( \leq \) be the total order on \( V(\hat{D}) \) associated to \( T \). Note that for all \( i \), after inverting \( X_1, \ldots, X_k \) the \( i \)th factor has one vertex that has out-edges to the other two vertices in the factor and exactly one of these edges has undergone a net reversal. Thus we can label the vertices in the \( i \)th factor as \( u_i, v_i, w_i \) where \( u_i v_i w_i \) is a directed 3-cycle in \( \hat{D} \), and the edge between \( u_i \) and \( w_i \) undergoes a net reversal under \( X_1, \ldots, X_k \) while the edge between \( u_i \) and \( v_i \) does not. In particular, we will use throughout that \( u_i < v_i, w_i \) and that, by Observation 1, \( u_i \cdot v_i = 0 \) and \( u_i \cdot w_i = 1 \) where \( u_i, v_i, w_i \in \mathbb{F}_2^k \) are the respective characteristic vectors of \( u_i, v_i, w_i \) in \( X_1, \ldots, X_k \). We have the following claim, originally proved in [22].

Claim 1 ([22]). The vectors \( u_1, \ldots, u_k \in \mathbb{F}_2^k \) are linearly independent.

Proof. The statement is equivalent to the claim that for all non-empty \( I \subseteq [k] \) we have \( \sum_{i \in I} u_i \neq 0 \). Fix such an \( I \) and note that it is sufficient to show that there exists some \( x \in \mathbb{F}_2^k \) such that \( (\sum_{i \in I} u_i) \cdot x \neq 0 \). Let \( m \in I \) be such that \( u_m \leq u_i \) for all \( i \in I \setminus \{m\} \). Note that \( u_m < v_m, w_m \), so by the transitivity of \( T \) we have \( u_i < v_m, w_m \) for all \( i \in I \).

It is straightforward to deduce from this that for all \( i \in I \setminus \{m\} \), the orientations of the edges \( u_i v_m \) and \( u_i w_m \) are either both unchanged after \( X_1, \ldots, X_k \) are inverted, or both reversed. By Observation 1, in other words we have \( u_i \cdot v_m = u_i \cdot w_m = 0 \) while \( u_m \cdot v_m = 1 \), so it follows by linearity of the dot product that \( (\sum_{i \in I} u_i) \cdot v_m = (\sum_{i \in I} u_i) \cdot w_m = 1 \). One of these two dot products is thus non-zero, and we deduce that \( \sum_{i \in I} u_i \neq 0 \), as required.

We now build on Claim 1 as follows.

Claim 2. Let \( \ell \in [k] \) and suppose that the vectors \( u_i, v_i, w_i \) for \( \ell \leq i \leq k \) all lie in a subspace \( V \) of \( \mathbb{F}_2^k \) of dimension \( k - \ell + 1 \). Then \( u_1, \ldots, u_k \) are orthonormal, and for all \( \ell \leq i \leq k \) we have \( u_i = w_i \) and \( v_i = 0 \).

Proof. We will prove the claim by reverse induction on \( \ell \). In the \( \ell = k \) case the claim follows easily from the fact that \( u_k \cdot w_k = 1 \) while \( u_k \cdot v_k = 0 \). Thus, let \( \ell \leq k - 1 \) and write \( \{\ell, \ell + 1, \ldots, k\} \) for \( \{\ell, \ell + 1, \ldots, k\} \). Let \( \ell \) be the \( \leq \)-minimal vertex among \( v_\ell, \ldots, v_k, w_\ell, \ldots, w_k \). Write \( z \in V \subseteq \mathbb{F}_2^k \) for the characteristic vector of \( z \) in \( X_1, \ldots, X_k \) and let \( t \in \{\ell, k\} \) be such that \( z \in \{v_t, w_t\} \). By Claim 1, the vectors \( u_\ell, \ldots, u_k \) form a basis of \( V \) so there exists \( I \subseteq [\ell, k] \) such that \( z + \sum_{i \in I} u_i = 0 \).

Let \( I \subseteq [\ell, k] \) with \( I \neq \emptyset \) and \( I \neq \{\ell, k\} \), and let \( m \in I \) be such that \( u_i < u_m \) for all \( i \in I \setminus \{m\} \). If \( m \neq \ell \), then we have \( z < v_m, w_m \), so \( z \cdot v_m = z \cdot w_m \) by Observation 1. As in the proof of Claim 1, we have \( u_i \cdot v_m = u_i \cdot w_m = 1 \) for all \( i \in I \setminus \{m\} \), but \( u_m \cdot v_m \neq u_m \cdot w_m \), so \( (z + \sum_{i \in I} u_i) \cdot v_m = (z + \sum_{i \in I} u_i) \cdot w_m \), and hence \( z + \sum_{i \in I} u_i \neq 0 \). If \( m = \ell \), then let \( j \in I \setminus \{\ell\} \) and note that \( z < v_j, w_j \) by the minimality of \( z \). Consequently, \( z \cdot v_j = z \cdot w_j \).

Moreover, since \( u_m = u_\ell < z \), we have \( u_m < v_j, w_j \). From this it follows that \( u_i < v_j, w_j \) for all \( i \in I \). Thus, \( u_i \cdot v_j = u_i \cdot w_j \) for all \( i \in I \setminus \{\ell\} \), while \( u_\ell \cdot v_j \neq u_\ell \cdot w_j \). Hence, similarly to above, we have \( (z + \sum_{i \in I} u_i) \cdot v_j = (z + \sum_{i \in I} u_i) \cdot w_j \), so \( z + \sum_{i \in I} u_i = 0 \).

The remaining cases are \( I = \emptyset \) and \( I = \{\ell\} \), so we have \( z \in \{0, u_\ell\} \). Suppose that \( z = u_\ell \). If \( z = v_t \), then we have \( v_t = u_t \), so \( v_t \cdot w_t = u_t \cdot w_t = 1 \), i.e. the edge between
we find that the cycle that there are orthonormal
we have from the lemma. In the proof, we will use the easy
induction hypothesis 

\[ u \]

\[ v \]

\[ x \]

clinging family of

\[ \widehat u \]

induced on the vertex set of the

\[ i \]

\[ V \]

so if

\[ u \]

\[ v \]

\[ x \]

and by the established properties of the

\[ u \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

Since

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]

\[ w \]
We have shown that all vertices in the $i$th factor of $\hat{D}$ have characteristic vector either $u$ or 0 in $X_1, \ldots, X_k$. The effect on this copy of $D_i$ of inverting these sets in $\hat{D}$ is therefore the same as inverting the set of vertices with characteristic vector $u$. The latter is therefore a decycling set for the $i$th factor of $\hat{D}$. This completes the proof of the theorem.

Theorem 3 now follows easily.

Proof of Theorem 3. It is clear that the left-hand side of equation (1) is at most the right-hand side. For the reverse inequality, let $\hat{D} = [D_1, \ldots, D_k]$ and note that we may assume that none of the $D_i$ have inversion number 0. Indeed, if $\text{inv}(D_i) = 0$ for some $i \geq 2$, then view $\hat{D}$ as the $(k-1)$-join $[D_1, \ldots, D_{i-2}, D_{i-1} \rightarrow D_i, D_{i+1}, \ldots, D_k]$ and, since $\text{inv}(D_{i-1} \rightarrow D_i) = \text{inv}(D_{i-1})$, the result follows by induction on $k$. The case where $i = 1$ can be handled similarly.

Thus, consider the case where $\text{inv}(D_i) = 1$ for all $i$ and suppose for a contradiction that $X_1, \ldots, X_k$ is a decycling family of $\hat{D}$ with $X_k = \emptyset$. By Theorem 4 there exist $k$ orthonormal, and hence linearly independent, vectors in $\mathbb{F}_2^k$ each of which occurs as the characteristic vector of some vertex of $\hat{D}$ in $X_1, \ldots, X_k$. This contradicts the fact that all such vectors have a 0 in their final coordinate. Hence, in this case, $\text{inv}(\hat{D}) = k$.

It remains to check the case where $\text{inv}(D_j) = 2$ for some $j$ and $\text{inv}(D_i) = 1$ for all $i \neq j$. Start by deleting edges from the $j$th factor of $\hat{D}$ until the digraph induced on those vertices has inversion number 1. We can now apply Theorem 4 to the resulting subdigraph of $\hat{D}$ to deduce that there is a vector $u_j \in \mathbb{F}_2^k$ such that all the vertices in the $j$th factor of $\hat{D}$ have characteristic vector either 0 or $u_j$ in $X_1, \ldots, X_k$. Inverting $X_1, \ldots, X_k$ in $\hat{D}$ therefore has the same effect on the $j$th factor as inverting the set of vertices with characteristic vector $u_j$. It follows that this set of vertices is a decycling set for $D_j$, contradicting $\text{inv}(D_j) = 2$.

As mentioned in the introduction, it follows from Theorem 3 that for any digraph $D$ we have $\text{inv}(D) = 1$ if and only if $\text{inv}([D]_k) = k$, which in turn implies Theorem 5. Indeed, Theorem 3 directly gives $\text{inv}([D]_k) = k$ in the case $\text{inv}(D) = 1$, and if $\text{inv}(D) = 0$ then clearly $\text{inv}([D]_k) = 0$. If $\text{inv}(D) > 1$, then there are subdigraphs $D'$ and $D''$ of $D$ with $\text{inv}(D') = 1$ and $\text{inv}(D'') = 2$. The $k$-join $D' \rightarrow [D']_{k-1}$, which has inversion number $k + 1$ by Theorem 3, is a subdigraph of $[D]_k$ and thus $\text{inv}([D]_k) \geq k + 1$ as required.

Finally, we deduce Theorem 2 from Theorem 3. We will use the fact, shown in [4], that if $L$ and $R$ are strongly connected digraphs with $\text{inv}(L), \text{inv}(R) \geq 2$, then $\text{inv}(L \rightarrow R) \geq 4$.

Proof of Theorem 2. Let $L$ and $R$ be digraphs with $\text{inv}(L) = \text{inv}(R) = 2$. It is immediate that $\text{inv}(L \rightarrow R) \leq 4$, so it is sufficient to prove the lower bound. For this, extend $L \rightarrow R$ to a tournament $T$ of the same inversion number and let the tournaments to which $L$ and $R$ are extended be $L'$ and $R'$ respectively. Note that $\text{inv}(L'), \text{inv}(R') \geq 2$ and $T$ is $L' \rightarrow R'$.
Every tournament can be written as the $k$-join of its strongly connected components, so let $L'$ be $[L_1, \ldots, L_{k_1}]$ and $R'$ be $[R_1, \ldots, R_{k_2}]$ for some $k_1, k_2 \in \mathbb{N}$ and strongly connected tournaments $L_1, \ldots, L_{k_1}, R_1, \ldots, R_{k_2}$. Since $\text{inv}(L') \geq 2$, either there is some $L_i$ with $\text{inv}(L_i) \geq 2$, or there are $i < j$ such that $\text{inv}(L_i) = \text{inv}(L_j) = 1$. An analogous condition holds for $R'$. If there are $i$ and $j$ such that $\text{inv}(L_i), \text{inv}(R_j) \geq 2$, then since $T$ contains $L_i \rightarrow R_j$, we have $\text{inv}(T) \geq \text{inv}(L_i \rightarrow R_j) \geq 4$ by the above result of [4]. Otherwise, either there exist $i < j$ such that $\text{inv}(L_i) = \text{inv}(L_j) = 1$, in which case $\text{inv}(T) \geq \text{inv}([L_i, L_j, R]) = 4$ by Theorem 3, or there exist $i < j$ with $\text{inv}(R_i) = \text{inv}(R_j) = 1$, in which case it follows similarly that $\text{inv}(T) \geq 4$. 

\section{Complexity of $k$-TOURNAMENT-INVERSION}

In this section we prove Theorem 6 by constructing for each fixed $k \in \mathbb{N}$ an algorithm which solves $k$-TOURNAMENT-INVERSION in time $O(|V(T)|^2)$. The proof is based on the following lemma which roughly says that for constant $k$, given an $n$-vertex tournament $T_0$ and a decycling family of $T_0$ of constant size, in time linear in $n$ we can find a decycling family of $T_0$ if one exists. Throughout this section we store a total order $< \in S = \{s_1, \ldots, s_m\}$ as the tuple $(s_1, \ldots, s_m)$ where $s_1 < \cdots < s_m$.

\begin{lemma}
Fix $k, s \in \mathbb{N}$. There is an algorithm which solves the following problem for $n$-vertex tournaments in time $O(n)$:

\begin{center}
\textbf{Inputs:}
\begin{itemize}
  \item a tournament $T_0$;
  \item a decycling family $X_1, \ldots, X_s$ of $T_0$ (transforming $T_0$ into $T$, say);
  \item the order on $V(T_0)$ associated to $T$.
\end{itemize}
\textbf{Outputs:}
\begin{itemize}
  \item Either
    \begin{itemize}
      \item that $T_0$ is not $k$-invertible;
    \end{itemize}
  \item or
    \begin{itemize}
      \item a decycling family $Y_1, \ldots, Y_k$ of $T_0$ (transforming $T_0$ into $T'$, say);
      \item the order on $V(T_0)$ associated to $T'$.
    \end{itemize}
\end{itemize}
\end{center}

Our algorithm works by exploring what happens if, starting from $T$, we invert $X_1, \ldots, X_s$ and $k$ further sets $Y_1, \ldots, Y_k$ to obtain a tournament $T_Y$, where $Y = (Y_1, \ldots, Y_k)$. Since $T_Y$ is the tournament obtained by inverting $Y_1, \ldots, Y_k$ in $T_0$, these $k$ sets are a decycling family of $T_0$ if and only if $T_Y$ is transitive. If we were to examine each possibility individually there would be too many for this exploration process to be tractable. However, the fact that we are starting from a transitive tournament $T$ makes it possible to identify cycles in the final tournament $T_Y$ without fully specifying the sets $Y_1, \ldots, Y_k$. This means there are far fewer cases to consider, indeed few enough that the exploration process is linear in $n$ for fixed $k$ and $s$. 

12
Proof of Lemma 10. Fix \( k, s \in \mathbb{N} \) and let \( T_0, X_1, \ldots, X_s, \) and \( T \) be as in the statement of the lemma. Let \( n = |V(T_0)| \) and label the vertices of \( T_0 \) as \( u_1, \ldots, u_n \) in \( T \)-increasing order. With notation as above, we wish to investigate for which \( Y \) the tournament \( T_Y \) is transitive. For each \( Y \) we write \( u_i \in \mathbb{F}_2^{s+k} \) for the characteristic vector of \( u_i \) in \( X_1, \ldots, X_s, Y_1, \ldots, Y_k \) (suppressing the dependence on \( Y \) in the notation) and then let \( u = (u_1, \ldots, u_n) \). There is a bijective correspondence between \( Y \) and \( u \) and it will be more convenient to work with the latter, so let \( T_u = T_Y \) and write \( \mathcal{U} \) for the set of all possible \( u \). Our first aim is to determine in linear time whether there exists \( u \in \mathcal{U} \) such that \( T_u \) is transitive, and to identify such a \( u \) if so.

The tournament \( T_u \) is transitive exactly when it contains no cyclic triples. It is straightforward to use Observation 1 to show that this is equivalent to the condition that there are no \( a < b < c \) in \([n]\) such that \( u_a \cdot u_b = u_b \cdot u_c \) but \( u_a \cdot u_b \neq u_a \cdot u_c \). We describe the triple \((u_a, u_b, u_c)\) as bad if this occurs. Thus, \( T_u \) is transitive if and only if \( \mathcal{B}(u) = \{(u_a, u_b, u_c) : a < b < c\} \) contains no bad triples, and \( T_0 \) is \( k \)-invertible if and only if \( \mathcal{B} = \{\mathcal{B}(u) : u \in \mathcal{U}\} \) contains a set which is free of bad triples. Our algorithm will construct this set \( \mathcal{B} \) and check whether any of its elements are free of bad triples. If one of these sets is free of bad triples, then we need to be able to output a corresponding decycling family of \( T_0 \), so for each \( B \in \mathcal{B} \) we will also record some \( u \in \mathcal{U} \) for which \( B = \mathcal{B}(u) \).

We will now explain how the above can be achieved in linear time. First note that we may assume that \( n \geq 4 \). Let \( \mathcal{U}' \) be the set of all possible vectors \( u' = (u_1, \ldots, u_{n-1}) \) of characteristic vectors of \( u_1, \ldots, u_{n-1} \) in \( X_1, \ldots, X_s, Y_1, \ldots, Y_k \). For \( u' \in \mathcal{U}' \), let \( \mathcal{B}'(u') = \{(u_a, u_b, u_c) : 1 \leq a < b < c \leq n - 1\} \) and let \( \mathcal{B}' = \{\mathcal{B}'(u') : u' \in \mathcal{U}'\} \). We may assume inductively that there is a constant \( C \) depending only on \( k \) and \( s \) such that in time \( C \cdot (n - 1) \) we can construct \( \mathcal{B}' \) and associate to each \( B' \in \mathcal{B}' \) some \( u' \in \mathcal{U}' \) such that \( B' = \mathcal{B}'(u') \). For the induction step, we need to show that we can use this to obtain in time \( C \) the set \( \mathcal{B} \) and for each \( B \in \mathcal{B} \) some \( u \in \mathcal{U} \) such that \( B = \mathcal{B}(u) \).

The key observation is that there are only \( 2^{s+k} \) possible characteristic vectors for each of \( u_1, \ldots, u_n \), so the number of triples of characteristic vectors is at most \( 2^{3(s+k)} \) and the sizes of \( \mathcal{B} \) and \( \mathcal{B}' \) are at most \( 2^{2(s+k)} \). In particular, there are only boundedly many pairs \((B', u_n)\) where \( B' \in \mathcal{B}' \) and \( u_n \) is a possible characteristic vector for \( u_n \). For each such pair, we can construct in bounded time the set \( S(B', u_n) \) consisting of all triples in \( B' \), and all triples of the form \((v_i, v_j, u_n)\) for \((v_i, v_j, v_3) \in B' \) and \( 1 \leq i < j \leq 3 \). It is not hard to see that \( \mathcal{B} \) equals the set of all sets \( S(B', u_n) \) and that each \( S(B', u_n) \) can be associated with the \( u \in \mathcal{U} \) formed by appending \( u_n \) to the \( u' \in \mathcal{U}' \) associated with \( B' \). Indeed, given \( B' \) and \( u_n \) and defining \( u \) as in the previous sentence, since \( n \geq 4 \), we have \( S(B', u_n) = \mathcal{B}(u) \). For the other direction, given \( u = (u_1, \ldots, u_n) \in \mathcal{U} \) and letting \( u' = (u_1, \ldots, u_{n-1}) \), we have \( \mathcal{B}(u) = S(B'(u'), u_n) \).

We can construct this set in bounded time and then forget about all but one of the elements of \( \mathcal{U} \) associated to each \( B \in \mathcal{B} \). Thus, in constant time we have obtained \( \mathcal{B} \) and for each \( B \in \mathcal{B} \) some \( u \in \mathcal{U} \) such that \( B = \mathcal{B}(u) \), and the induction continues.

Once we have constructed \( \mathcal{B} \) in linear time, since it has bounded size we can check whether any of its members is free of bad triples in bounded time. If not, then \( T_0 \) is
not \(k\)-invertible. If so, then pick \(B \in \mathcal{B}\) with no bad triples and use the \(u\) associated to it to construct a decycling family \(Y_1, \ldots, Y_k\) of \(T_0\). Let \(T'\) be the transitive tournament obtained by inverting these sets in \(T_0\).

It remains to show that we can obtain the order on \(V(T_0)\) associated to \(T'\) in linear time. Inverting the sets \(X_1, \ldots, X_s, Y_1, \ldots, Y_k\) transforms \(T\) into \(T'\), and we have the characteristic vector of each vertex in these sets as well as the order on the vertices associated to \(T\). We can therefore in linear time obtain the atoms of these \(s+k\) inversions and the order on each atom inherited from the order on \(T\). If the edges within a given atom undergo a net reversal under the inversions, then reverse the order on that atom. The \(T'\)-minimal vertex is now the minimal vertex of one of the atoms under their current orderings. There are at most \(2^{s+k}\) atoms so we can identify the \(T'\)-minimal vertex in constant time. After deleting this vertex from its atom, the second smallest vertex according to \(T'\) is one of the new minimal vertices of the atoms so can be found in constant time again. Continuing in this way we can obtain the full ordering in linear time, as required.

Note that the implicit constant in the running time given by this proof is doubly exponential in \(s+k\). We are now ready to prove Theorem 6.

**Proof of Theorem 6.** Fix \(k \geq 1\). We will induct on \(n\) to define an algorithm solving the following problem for \(n\)-vertex tournaments in time \(C_k \cdot n^2\) for some constant \(C_k\):

**INPUT:**
- a tournament \(T_0\).

**OUTPUTS:**
- Either
  - that \(T_0\) is not \(k\)-invertible;
  - a decycling family \(Y_1, \ldots, Y_k\) of \(T_0\) (transforming \(T_0\) into \(T\), say);
- the order on \(V(T_0)\) associated to \(T\).

In particular, this algorithm solves \(k\)-TOURNAMENT-INVERSION.

Fix \(n \geq 2\) and assume that we have defined such an algorithm for all smaller tournaments. Let \(T_0\) be an \(n\)-vertex tournament and pick some \(v \in V(T_0)\). Applying the induction hypothesis, in time \(C_k \cdot (n-1)^2\) we either find that \(T_0 - \{v\}\) is not \(k\)-invertible or we obtain a decycling family \(X_1, \ldots, X_k\) of \(T_0 - \{v\}\) and the order on \(V(T_0) \setminus \{v\}\) associated to the transitive tournament obtained by inverting these sets in \(T_0\). In the former case, it follows that \(T_0\) is also not \(k\)-invertible and we can output that fact. In the latter case, let \(A\) be the out-neighbourhood of \(v\) in \(T_0\), and define \(X_{k+1} = A \cup \{v\}\) and \(X_{k+2} = A\). Then \(X_1, \ldots, X_{k+2}\) is a decycling family of \(T_0\), and we can obtain the order associated to the resulting transitive tournament by adding \(v\) to the previous order as the maximal element. By Lemma 10 we can now, in linear time, either find that \(T_0\) is not \(k\)-invertible or obtain a decycling family \(Y_1, \ldots, Y_k\) of \(T_0\) of size \(k\) and the order associated to the transitive tournament obtained by inverting these sets in \(T_0\). As required,
this algorithm runs in time \( C_k \cdot (n - 1)^2 + O(n) \), which is at most \( C_k \cdot n^2 \) if \( C_k \) is large enough.

\[ \square \]

6 Cycle transversals

In this section we will prove Theorem 7, constructing for each \( k \in \mathbb{N} \) a digraph \( D \) with \( \tau(D) = k \) and \( \text{inv}(D) = 2k \). We will use the so-called Eventown theorem, proved independently by Berlekamp [8] and Graver [15].

**Theorem 11** (Eventown [8], [15]). Let \( n \in \mathbb{N} \) and let \( \mathcal{F} \subseteq \mathcal{P}([n]) \) be a family of subsets of \([n]\) such that \(|F_1 \cap F_2|\) is even for all \( F_1, F_2 \in \mathcal{F} \). Then \(|\mathcal{F}| \leq 2^{n/2} \).

For a digraph \( D \) and vertices \( u, v, w \in V(D) \), we will say that \( u \) and \( v \) differ on \( w \) if either \( uvw \), \( wuv \in E(D) \) or \( vwu \), \( wu \in E(D) \). We are now ready to prove the theorem.

**Proof of Theorem 7.** Fix \( k \in \mathbb{N} \) and let \( n \in \mathbb{N} \) be large and divisible by \( 2^k \). We will define a digraph \( D \) on vertex set \( \{u_0, \ldots, u_{k-1}, v_0, \ldots, v_{n-1}\} \) and then show that it satisfies the conditions of the theorem. Start by including all directed edges \( \overrightarrow{uv}_j \) for \( i < j \), so that the subdigraph of \( D \) induced on \( \{v_0, \ldots, v_{n-1}\} \) is a transitive tournament. For \( i \in \{0, \ldots, k-1\} \) and \( j \in \{0, \ldots, n-1\} \), add the edge \( \overrightarrow{uw}_i \) if in the binary expansion of \( j \), the digit in the \( 2^i \) place is a 0, and add the edge \( \overrightarrow{vw}_i \) otherwise. For ease of exposition we will not include any edges among the \( u_i \) (though including any combination of such edges would still give a valid construction), so this completes the definition of \( D \). As noted above, the removal of the vertices \( u_0, \ldots, u_{k-1} \) from \( D \) leaves an acyclic digraph, so \( \tau(D) \leq k \).

It remains to show that \( \text{inv}(D) \geq 2k \), as then \( \text{inv}(D) = 2\tau(D) = 2k \) follows from \( \text{inv}(D) \leq 2\tau(D) \). Suppose for a contradiction that \( X_1, \ldots, X_{2k-1} \subseteq V(D) \) form a decaying family of \( D \) and let \( D' \) be the acyclic digraph obtained by inverting these sets in \( D \). Consider the characteristic vectors of \( v_0, \ldots, v_{n-1} \) in \( X_1, \ldots, X_{2k-1} \), which we will denote by \( \mathbf{v}_0, \ldots, \mathbf{v}_{n-1} \in \mathbb{F}_2^{2k-1} \) respectively. Let \( K = 2^k \). By the pigeonhole principle, if \( n \) is large enough, then there exist distinct \( i, i' \in \{0, \ldots, n/K - 1\} \) such that

\[
(\mathbf{v}_{iK}, \mathbf{v}_{iK+1}, \ldots, \mathbf{v}_{(i+1)K-1}) = (\mathbf{v}_{i'K}, \mathbf{v}_{i'K+1}, \ldots, \mathbf{v}_{(i'+1)K-1}).
\]

We may assume that \( i = 0 \) and \( i' = 1 \).

We will show that \( \mathbf{v}_0, \ldots, \mathbf{v}_{K-1} \) are pairwise distinct and that \( \mathbf{v}_i \cdot \mathbf{v}_j \) is constant as \( i, j \in \{0, \ldots, K - 1\} \) vary. We claim that these conditions force a contradiction. Indeed, in the case where \( \mathbf{v}_i \cdot \mathbf{v}_j = 0 \) for all \( i, j \), we have that the \( \mathbf{v}_i \) are indicator vectors of pairwise distinct subsets of \([2k-1]\) which each have even size, and each pair of which have even intersection. By Eventown, every such collection has at most \( 2^{(2k-1)/2} < 2^k = K \) members, giving the required contradiction. On the other hand, if \( \mathbf{v}_i \cdot \mathbf{v}_j = 1 \) for all \( i, j \), then consider the ‘complement’ vectors \( \mathbf{w}_0, \ldots, \mathbf{w}_{K-1} \), which have 1’s where the \( \mathbf{v}_i \) have 0’s and 0’s where the \( \mathbf{v}_i \) have 1’s. It is straightforward to use the fact that the vectors have odd length to show that these \( \mathbf{w}_i \) are pairwise distinct and satisfy \( \mathbf{w}_i \cdot \mathbf{w}_j = 0 \) for all \( i, j \), from which we can derive a contradiction as above.
We continue by showing that the vectors \( \mathbf{v}_0, \ldots, \mathbf{v}_{K-1} \) are pairwise distinct, which is equivalent to showing that each of \( \mathbf{v}_0, \ldots, \mathbf{v}_{K-1} \) is in a different atom. Suppose for a contradiction that \( \mathbf{v}_i \) and \( \mathbf{v}_j \) are in the same atom for some \( i < j \in \{0, \ldots, K-1\} \), and note that \( \mathbf{v}_{K+i} \) and \( \mathbf{v}_{K+j} \) are in this atom too by assumption. By the construction of \( D \) there is some \( \ell \in \{0, \ldots, k-1\} \) such that \( \mathbf{v}_i \) and \( \mathbf{v}_j \) differ on \( u_\ell \) in \( D \). Since they are in the same atom as each other, they also differ on \( u_\ell \) in \( D' \). If \( \overrightarrow{v_i u_\ell}, \overrightarrow{u \ell v_j} \in E(D') \), then to avoid a cyclic triple in \( D' \) we have \( \overrightarrow{v_i v_j} \in E(D') \). This means that the edges within \( \mathbf{v}_i \) and \( \mathbf{v}_j \)'s atom have the same orientations in \( D \) as they do in \( D' \), so in particular we have \( \overrightarrow{v_j v_{K+i}} \in E(D') \). Moreover \( \mathbf{v}_i \) and \( \mathbf{v}_{K+i} \) are in the same atom and are either both in-neighbours of \( u_\ell \) in \( D \) or both out-neighbours, so since \( \overrightarrow{v_i u_\ell} \in E(D') \) we also have \( \overrightarrow{v_{K+i} u_\ell} \in E(D') \). Hence, the cycle \( \overrightarrow{v_j v_{K+i} u_\ell} \) appears in \( D' \). Similarly if \( \overrightarrow{v_j u_\ell}, \overrightarrow{u_\ell v_i} \in E(D') \), then we have \( \overrightarrow{v_j v_i} \in E(D') \). In this case the edges within the atom of \( \mathbf{v}_i \) and \( \mathbf{v}_j \) switch orientation between \( D \) and \( D' \), so the cycle \( \overrightarrow{v_j u_\ell v_{K+i}} \) appears in \( D' \). In both cases we have the desired contradiction, and we deduce that the vertices \( \mathbf{v}_0, \ldots, \mathbf{v}_{K-1} \) are all in different atoms.

It remains to show that \( \mathbf{v}_i \cdot \mathbf{v}_j \) is constant as \( i, j \in \{0, \ldots, K-1\} \) vary. Suppose for a contradiction that this is not the case, then there exists \( i \in \{0, \ldots, K-1\} \) such that \( \mathbf{v}_i \cdot \mathbf{v}_j \) is not constant as \( j \in \{0, \ldots, K-1\} \) varies. For such \( i \) we can pick \( j \in \{0, \ldots, K-1\} \) such that \( \mathbf{v}_i \cdot \mathbf{v}_i \neq \mathbf{v}_i \cdot \mathbf{v}_j \). Now if \( \mathbf{v}_i \cdot \mathbf{v}_i = 0 \), then \( \mathbf{v}_i \cdot \mathbf{v}_j = 1 \) so by Observation 1, \( D' \) contains the cycle \( \overrightarrow{v_{K+i} v_j v_i} \) if \( i < j \) or the cycle \( \overrightarrow{v_{K+i} v_{K+j} v_i} \) if \( i > j \). Similarly if \( \mathbf{v}_i \cdot \mathbf{v}_i = 1 \), then \( D' \) contains one of the cycles \( \overrightarrow{v_i v_j v_{K+i}} \) and \( \overrightarrow{v_i v_{K+j} v_{K+i}} \). We have a contradiction in all cases, so the value of \( \mathbf{v}_i \cdot \mathbf{v}_j \) is constant as \( i, j \in \{0, \ldots, K-1\} \) vary, as required. \( \square \)

7 Bounds on \( \text{inv}(n) \)

7.1 Lower bounds

In this section we discuss the previous known lower bound on \( \text{inv}(n) \) and give the proof of Theorem 8. As noted in the introduction, Belkhechine, Bouaziz, Boudabbous, and Pouzet [7] used a counting argument to lower bound \( \text{inv}(n) \). They observed that since there are \( n! \) labelled transitive tournaments on \( n \) vertices, there are at most \( n! \cdot 2^{n(k-1)} \) labelled \( (k-1) \)-invertible tournaments on \( n \) vertices. There are a total of \( 2^{n(n-1)/2} \) labelled \( n \)-vertex tournaments, so for any \( k \) such that \( 2^{n(n-1)/2} > n! \cdot 2^{n(k-1)} \) we have \( \text{inv}(n) \geq k \). Taking logarithms base 2 and rearranging, this condition becomes \( k < (n - 1)/2 - \log(n!)/n \), so we have

\[
\text{inv}(n) \geq \left\lceil \frac{n - 1}{2} - \frac{\log(n!)}{n} \right\rceil \geq \left[ \frac{n - 1}{2} - \log(n) \right],
\]

where for the final inequality we used \( n! \leq n^n \). Lower bounds on \( \text{inv}(n) \) of this form were the best known (disregarding very slight tightenings of the argument).

The proof of Theorem 8 uses the following lemma which gives a bound on the probability that a random symmetric binary matrix has at most a certain rank. In fact, these probabilities are known exactly [19], but we will use a simpler bound which is essentially tight for our purposes and for which we include a short proof.
Lemma 12. The probability that a uniformly random $n \times n$ symmetric matrix over $\mathbb{F}_2$ has rank at most $n-s$ (over $\mathbb{F}_2$) is at most $2^{s \log(n)-(\binom{s}{2})}$.

Proof. Construct the random matrix in $n$ steps, in the $i$th step choosing the first $i$ entries of the $i$th row of the matrix (and also, by symmetry, the $i$th column). For each $i \in [n]$, let $M_i$ be the random symmetric $i \times i$ matrix obtained after step $i$.

Note that for each $i$ the nullity increases by at most 1 between $M_i$ and $M_{i+1}$. It follows that if the nullity of $M_n$ is at least $s$, then for all $1 \leq j \leq s-1$ we can define $k_j$ to be the smallest $i$ such that the nullity of $M_i$ is $j+1$, and we have $2 \leq k_1 < k_2 < \cdots < k_{s-1} \leq n$. For each $j$, the ranks of $M_{k_j-1}$ and $M_{k_j}$ are equal, so the first $k_j-1$ entries of the $k_j$th row of $M_{k_j}$ lie in the $(k_j-1-j)$-dimensional row space of $M_{k_j-1}$, which happens with probability $2^{-j}$. There are $\binom{s-1}{j}$ ways to choose $k_1, \ldots, k_{s-1}$ as above, so the probability that $M_n$ has rank at most $n-s$ is at most

$$\left(\frac{n}{s-1}\right)^{s-1} \prod_{j=1}^{s-1} 2^{-j} \leq 2^{s \log(n)-(\binom{s}{2})},$$

as required. \hfill \Box

We are now ready to prove Theorem 8.

Proof of Theorem 8. Let $T$ be a uniformly random tournament on vertex set $[n]$ and let $M_T = (m_{ab})$ be the $n \times n$ matrix over $\mathbb{F}_2$ defined as follows. For $a < b$, let $m_{ab}$ be 0 if $ab$ is an edge of $T$ and 1 otherwise, then define $m_{ba} = m_{ab}$, and finally choose each diagonal entry uniformly at random. Note that the $\binom{n}{2}$ entries of $M_T$ above the diagonal determine $T$, and the other entries are defined such that $M_T$ is a uniformly random symmetric binary matrix.

Let $s = \sqrt{2n \log(n)}$ and write $k = n-s$. Suppose that $\text{inv}(T) \leq k$ and let $X_1, \ldots, X_k$ be a decycling family of $T$. For each $X_i$, let $M_i$ be the $n \times n$ binary matrix whose $(a, b)$ entry is 1 if and only if $a, b \in X_i$. Observe that, working over $\mathbb{F}_2$, we have $\text{rank}(M_i) \leq 1$ for all $i$, and thus $\text{rank}(\sum_i M_i) \leq k$. By construction, $M_T + \sum_i M_i$ is a matrix whose entries above the diagonal correspond to a transitive tournament on $[n]$ (its diagonal entries can be anything). Let $\mathcal{M}$ be the set of binary matrices corresponding in this manner to a transitive tournament on $[n]$, and note that $|\mathcal{M}| = n!2^n$.

Putting all of this together, we have that if $\text{inv}(T) \leq k$, then there exists $M \in \mathcal{M}$ such that $\text{rank}(M_T + M) \leq k$. For each fixed $M$, we have that $M_T + M$ is a uniformly random symmetric binary matrix and hence has rank at most $k$ with probability at most $2^{s \log(n)-(\binom{s}{2})}$ by Lemma 12. Taking a union bound over all $M \in \mathcal{M}$ we obtain

$$\mathbb{P}(\text{inv}(T) \leq k) \leq n!2^n 2^{s \log(n)-(\binom{s}{2})}.$$ 

Since $n! = O(\sqrt{n}(n/e)^n)$, the right-hand side is $O(2^f(n))$ where

$$f(n) = \frac{\log(n)}{2} + n \log(n) - n \log(e) + n + s \log(n) - \binom{s}{2}$$

$$= -n(\log(e) - 1) + o(n),$$

and thus $\mathbb{P}(\text{inv}(T) \leq k) \to 0$ as $n \to \infty$ as desired. \hfill \Box
7.2 Upper bounds

The only approach which has been used to prove upper bounds on $\text{inv}(n)$, introduced in [7], is to ‘solve’ one vertex at a time, as follows. Given a tournament $T$, pick a vertex $v$ and invert the set consisting of $v$ and its out-neighbourhood. In the resulting tournament $T_1$, $v$ is a sink. Using a further $\text{inv}(n-1)$ inversions we can transform $T_1 - \{v\}$ into a transitive tournament, so $\text{inv}(n) \leq \text{inv}(n-1) + 1$ for all $n \geq 2$. The authors of [7] observed that $\text{inv}(4) = 1$, so $\text{inv}(n) \leq n - 3$ for $n \geq 4$. For $n \geq 6$ this was improved by 1 in [4] using the fact that $\text{inv}(6) = 2$ (which they attribute to [6] and which we have verified by a computer check). We introduce a slightly different approach to prove the following.

**Proposition 13.** For all $n \in \mathbb{N}$,

$$\text{inv}(n) \leq \left\lfloor \frac{n-1}{2} \right\rfloor + \text{inv}\left(\left\lceil \frac{n-1}{2} \right\rceil\right).$$

**Proof.** Let $n \in \mathbb{N}$ and let $T$ be an $n$-vertex tournament. Pick $v \in V(T)$ and write $A$ and $B$ for the in- and out-neighbourhoods of $v$ respectively. We may assume that $|A| \geq \lceil (n-1)/2 \rceil$ (the case where $B$ is the larger of the two is similar). By ‘solving’ each vertex in $B$ one after another, we can find at most $|B|$ inversions which transform $T$ into a tournament $T'$ such that the subtournament of $T'$ induced on $B \cup \{v\}$ is transitive (with $v$ as the minimal element) and every edge of $T'$ between $A$ and $B \cup \{v\}$ is oriented away from $A$. With a further $\text{inv}(A) \leq \text{inv}(|A|)$ inversions we can transform $T'$ into a transitive tournament. Thus, $\text{inv}(T) \leq |B| + \text{inv}(|A|)$.

We have $\text{inv}(k) \leq \text{inv}(k-1)+1$ for all $k \in \mathbb{N}$ and we can apply this repeatedly to obtain $\text{inv}(|A|) \leq \text{inv}\left(\lceil (n-1)/2 \rceil\right) + |A| - \left\lfloor (n-1)/2 \right\rfloor$. Using the fact that $|A| + |B| = n - 1$, this yields

$$\text{inv}(T) \leq |B| + \text{inv}\left(\left\lceil \frac{n-1}{2} \right\rceil\right) + |A| - \left\lfloor \frac{n-1}{2} \right\rfloor = \left\lfloor \frac{n-1}{2} \right\rfloor + \text{inv}\left(\left\lceil \frac{n-1}{2} \right\rceil\right),$$

and the claim follows. \hfill \Box

We can use this result to improve (for large $n$) the upper bound on $\text{inv}(n)$.

**Corollary 14.** For all $n \in \mathbb{N}_0$, $\text{inv}(n) \leq n - \log(n + 1)$.

**Proof.** We prove the statement by induction on $n$, with the case $n = 0$ clear. If $n \geq 1$ and the claim holds for all smaller values, then we have

$$\text{inv}(n) \leq \left\lfloor \frac{n-1}{2} \right\rfloor + \text{inv}\left(\left\lceil \frac{n-1}{2} \right\rceil\right),$$

$$\leq \left\lfloor \frac{n-1}{2} \right\rfloor + \left\lceil \frac{n-1}{2} \right\rceil - \log \left(\left\lceil \frac{n-1}{2} \right\rceil + 1\right),$$

$$\leq n - 1 - \log \left(\frac{n+1}{2}\right),$$

$$= n - \log(n + 1).$$

\hfill \Box
8 Conclusion

In this paper we have answered several of the questions posed in [4]. We have shown that their ‘dijoin conjecture’, that \( \text{inv}(L \rightarrow R) = \text{inv}(L) + \text{inv}(R) \), is false in general, but have verified it in the case where \( \text{inv}(L) = \text{inv}(R) = 2 \) and have also shown that a \( k \)-join analogue holds under certain conditions. In addition, we have confirmed their related conjectures that \( k \)-INVERSION is NP-complete for all \( k \geq 1 \), and that the inequality \( \text{inv}(D) \leq 2\tau(D) \) is tight. We have answered their question concerning the minimal \( r_k \) such that \( k \)-TOURNAMENT-INVERSION can be solved in time \( O(|V(T)|^{r_k}) \), showing that \( r_k = 2 \) for all \( k \), and have improved the lower bound on \( \text{inv}(n) \) to show that \( \text{inv}(n) = (1+o(1))n \).

There are, however, still many interesting open problems in this area. Before discussing some of them, we will touch on two operations similar to inversion.

8.1 Similar operations

We first consider an operation on permutations which is used by molecular biologists as a model for genetic mutations, and could loosely be seen as a vertex analogue of inversions in tournaments. Given a permutation \( \pi = (\pi_1 \pi_2 \ldots \pi_n) \) of \([n]\), for \( 1 \leq i < j \leq n \), the reversal of the interval \([i,j]\) is the permutation obtained by reversing the order of \( \pi_i, \ldots, \pi_j \) in \( \pi \). The reversal distance, \( d(\pi) \), of a permutation \( \pi \) is the minimum number of reversals required to transform \( \pi \) into the identity permutation. For a survey of reversals and the reversal distance (and many other combinatorial models of genome rearrangements) see [13]. We highlight two results of particular relevance to our work. Firstly, Caprara [9] showed that the problem of sorting by reversals, that is, determining whether \( d(\pi) \leq k \) for an input permutation \( \pi \) and input \( k \in \mathbb{N} \), is NP-complete. Secondly, the natural extremal problem was solved by Bafna and Pevzner [3] who proved that for a permutation \( \pi \) of \([n]\), we have \( d(\pi) \leq n - 1 \) with equality if and only if \( \pi \in \{\gamma_n, \gamma_n^{-1}\} \) for an explicit \( \gamma_n \).

Inversions in digraphs can also be thought of as generalisations of edge reversals, i.e., the operations which reverse the orientation of a single edge. It is not difficult to see (using an argument from Section 1.3) that the minimum number of such operations required to transform a digraph \( D \) into an acyclic digraph is equal to \( \tau'(D) \), the cycle edge-transversal number of \( D \). Determining this quantity is the famous feedback arc set problem, which has been widely studied (see [17] for an overview). In particular the problem of determining for inputs \( D \) and \( k \) whether \( \tau'(D) \leq k \) was one of the first shown to be NP-complete [16]. Moreover, the problem remains NP-complete when the input is restricted to tournaments [1, 10]. It was shown by Spencer [23, 24] that the maximum cycle edge-transversal number of an \( n \)-vertex tournament is \( \frac{1}{2}n(n-1) - \Theta(n^{3/2}) \) and that a random labelled \( n \)-vertex tournament has this cycle edge-transversal number with probability tending to 1. Bounds of this form remain the best known (see also [12, 21]).
8.2 Open problems

Our Theorems 5 and 6 are concerned with the problem of fixing \( k \in \mathbb{N} \), then deciding whether an input digraph \( D \) or tournament \( T \) is \( k \)-invertible. In both the settings discussed in Section 8.1 the analogous problem is trivially polynomial time solvable for all \( k \), but, as we have noted, if we instead take \( k \) as input then these problems become NP-complete. Bang-Jensen, da Silva, and Havet [4] conjectured that, analogously, the problem which takes inputs of \( k \in \mathbb{N} \) and a tournament \( T \), then asks whether \( \text{inv}(T) \leq k \), is NP-complete. Note that Theorem 6 does not make any progress towards disproving this because the implied constant in the \( O(n^2) \) running time grows very quickly with \( k \).

Conjecture 2 ([4]). The problem of deciding whether \( \text{inv}(T) \leq k \) for inputs of \( k \in \mathbb{N} \) and a tournament \( T \) is NP-complete.

As discussed in the introduction, the set \( \mathcal{IC}_k \) of \( k \)-inversion-critical tournaments was shown to be finite for all \( k \) in [7]. They explicitly described \( \mathcal{IC}_1 \) and \( \mathcal{IC}_2 \), for the latter using results of Gallai [14] (see [20] for an English translation) and Latka [18], but for \( k \geq 3 \) very little is known about these sets. In particular, it would be interesting to determine \( m_k \), the maximum number of vertices in a tournament in \( \mathcal{IC}_k \), for \( k \geq 3 \).

Question 3 ([4]). What is the value of \( m_k \) for \( k \geq 3 \)?

Finding the minimum possible size of a \( k \)-inversion-critical tournament is equivalent to the problem of determining \( \text{inv}(n) \). The best known bounds on \( \text{inv}(n) \) for large \( n \) are now

\[
 n - \sqrt{2n \log(n)} \leq \text{inv}(n) \leq n - \log(n + 1),
\]

and it would be interesting to tighten these further.

Question 4. What is the asymptotic behaviour of \( n - \text{inv}(n) \)?

In light of our improved lower bound on \( \text{inv}(n) \), the lack of an explicit construction for a tournament of large inversion number is even more apparent: no \( n \)-vertex construction with inversion number more than about \( n/3 \) (as given by the \( (n/3) \)-join \([\overline{C}_3]_{n/3}\)) is known.

Problem 5. Construct \( n \)-vertex tournaments with inversion number closer to \( \text{inv}(n) \).

Belkhechine, Bouaziz, Boudabbous, and Pouzet ([6]; see [4]) defined for each \( n \in \mathbb{N} \) a tournament \( Q_n \) on vertex set \([n]\) in which for \( i < j \) the edge \( ij \) is oriented towards \( j \), except if \( j = i + 1 \), in which case it is oriented towards \( i \), and conjectured that these graphs satisfy \( \text{inv}(Q_n) = \left\lfloor \frac{n-1}{2} \right\rfloor \).

Conjecture 6 ([6]). For all \( n \in \mathbb{N} \) we have \( \text{inv}(Q_n) = \left\lfloor \frac{n-1}{2} \right\rfloor \).

The conjecture is known to hold for \( n \leq 8 \) [4, 7], and it is certainly true that \( \text{inv}(Q_n) \leq \left\lfloor \frac{n-1}{2} \right\rfloor \) for all \( n \) since the sets

\[
 \{2, 3\}, \{4, 5\}, \{6, 7\}, \ldots, \{2 \left\lfloor (n-1)/2 \right\rfloor , 2 \left\lfloor (n-1)/2 \right\rfloor + 1\}
\]
form a decycling family of $Q_n$.

Defining the inversion distance, $\text{inv}(T, T')$, between two labelled tournaments $T$ and $T'$ on the same vertex set to be the minimum number of inversions required to transform $T$ into $T'$, we remark that the matrix rank techniques developed in Section 7.1 can be used to show that the maximum inversion distance between two $n$-vertex tournaments is exactly $n - 1$. Moreover, combining these ideas with Lemma 12 gives an upper bound of $2\binom{n}{2} + (\frac{n}{2} + \log(n))$ on the number of labelled tournaments within inversion distance $n - s$ of a given labelled tournament.

It is natural in this context to study the random walk $W$ on the space of labelled tournaments on $[n]$ where each step in the walk consists of picking a uniform random subset of $[n]$ and inverting that set in the current tournament. In particular, we ask the following.

**Question 7.** What is the mixing time of $W$? Does it satisfy the cutoff phenomenon?

Returning to the dijoin conjecture, Theorem 2 completes the work of Bang-Jensen, da Silva, and Havet in showing that the conjecture holds when $\text{inv}(L), \text{inv}(R) \leq 2$. We have also shown (Theorem 3) a $k$-join analogue of the dijoin conjecture for collections of 2-invertible digraphs $D_1, \ldots, D_k$ at most one of which has inversion number 2. We conjecture that this final condition can be removed.

**Conjecture 8.** Let $k \in \mathbb{N}$ and let $D_1, \ldots, D_k$ be oriented graphs with $\text{inv}(D_i) \leq 2$ for all $i$. Then

$$\text{inv}([D_1, \ldots, D_k]) = \sum_{i=1}^{k} \text{inv}(D_i).$$

On the other hand, Theorem 1 gives a counterexample to the dijoin conjecture where $\text{inv}(L) = 1$ and $\text{inv}(R) = 3$. From this, we can obtain counterexamples with $\text{inv}(L) = k$ and $\text{inv}(R) = 3$ for any $k \in \mathbb{N}$: let $L = [C^*_3]_k$ and let $R$ be as in the proof of Theorem 1. The tournaments obtained from these by inverting the whole vertex set give counterexamples in which $\text{inv}(L) = 3$ and $\text{inv}(R) = k$. We conjecture that here 3 can be replaced with any larger integer, or in other words that the only values of $\text{inv}(L)$ and $\text{inv}(R)$ for which the dijoin conjecture always holds are those where $\text{inv}(L), \text{inv}(R) \leq 2$ or where one of $\text{inv}(L)$ or $\text{inv}(R)$ is 0.

**Conjecture 9.** For all $\ell, r \in \mathbb{N}$ with $\ell \geq 3$ or $r \geq 3$ there exist oriented graphs $L$ and $R$ with $\text{inv}(L) = \ell$ and $\text{inv}(R) = r$, but $\text{inv}(L \rightarrow R) < \ell + r$.

This conjecture is equivalent to the claim that for all $r \geq 3$ there exists a tournament $R$ with $\text{inv}(R) = \text{inv}(C_3^* \rightarrow R) = r$. To see that this follows from the conjecture, note that for $r \geq 3$, if $\text{inv}(L) = 1$ and $\text{inv}(R) = \text{inv}(L \rightarrow R) = r$, then we can extend $L \rightarrow R$ to a tournament $T = L' \rightarrow R'$ with inversion number $r$. Clearly $\text{inv}(R') = r$ and $\text{inv}(L') \geq 1$, so $L'$ contains a copy of $C_3^*$. Thus, $r = \text{inv}(R') \leq \text{inv}(C_3^* \rightarrow R') \leq \text{inv}(T) = r$, as required. The converse follows from the arguments of the previous paragraph.

Finally, we noted in Section 1.2 that $\text{inv}(D) \leq \text{inv}(D - \{v\}) + 2$ for all digraphs $D$ and vertices $v \in V(D)$. It is certainly the case that this inequality is tight for some
Indeed, a reformulation of Theorem 7 yields the stronger statement that for all $k \in \mathbb{N}$ there exists a digraph $D$ and a set $S \subseteq V(D)$ with $|S| = k$ such that for all $T \subseteq S$ we have $\text{inv}(D - T) = \text{inv}(D) - 2|T|$. We conjecture, however, that the inequality $\text{inv}(D) \leq \text{inv}(D - \{v\}) + 2$ cannot be tight for all vertices $v$ in a given digraph $D$.

**Conjecture 10.** Let $D$ be a digraph with at least one vertex. Then there exists $v \in V(D)$ such that $\text{inv}(D - \{v\}) \geq \text{inv}(D) - 1$.

**References**


