Invertibility of digraphs and tournaments

Noga Alon[*]   Emil Powierski[†]   Michael Savery[†‡]
Alex Scott[†§]   Elizabeth Wilmer[¶]

8 September 2023

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 orientations of the edges of D with both endpoints in X. The *inversion number of D*, $\mathrm{inv}(D)$, 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 $k \in \mathbb{N}$ and tournament T, the problem of deciding whether $\mathrm{inv}(T) \leq k$ is solvable in time $O_k(|V(T)|^2)$, which is tight for all k. In particular, the problem is fixed-parameter tractable when parameterised by 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 $\mathrm{inv}(D) \leq k$ 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 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 $\mathrm{inv}(D)$, 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 $\mathrm{inv}(D) \leq k$.

The study of inversions began in Houmem Belkhechine's PhD thesis [5] and continued in [6, 7, 26], in which many foundational results were established. The present work is inspired by a recent paper of Bang-Jensen, da Silva, and Havet [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 [4] is the following 'dijoin conjecture'. For oriented graphs L and R, the *dijoin $L \to R$ 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 ([4])**.** *For oriented graphs L and R we have* $\mathrm{inv}(L \to R) = \mathrm{inv}(L) + \mathrm{inv}(R)$.

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 $\mathrm{inv}(L) = 0$ or $\mathrm{inv}(R) = 0$, Bang-Jensen, da Silva, and Havet showed it to be true when $\mathrm{inv}(L) + \mathrm{inv}(R) \leq 3$, and when $\mathrm{inv}(L) = \mathrm{inv}(R) = 2$ and L and R are both strongly connected.[1] 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 $\mathrm{inv}(R) = \mathrm{inv}(\overrightarrow{C_3} \to R)$, where $\overrightarrow{C_3}$ is the directed cycle on three vertices.

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

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

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

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 [26] which states that $\mathrm{inv}([\overrightarrow{C_3}]_k) = k$ for all k.

---

[1]The case where $\mathrm{inv}(L) = 2$ and $\mathrm{inv}(R) = 1$ is not explicitly mentioned in [4], but follows easily from the case where $\mathrm{inv}(L) = 1$ and $\mathrm{inv}(R) = 2$ by inverting $V(L \to R)$.

Theorem 3. *Let $k \in \mathbb{N}$ and let $D_1, \ldots, D_k$ be oriented graphs. Assume that $\mathrm{inv}(D_i) \leq 2$ for all i, with equality for at most one i. Then*

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

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 $\mathbf{u}, \mathbf{v} \in \mathbb{F}_2^k$ we write $\mathbf{u} \cdot \mathbf{v}$ for the usual scalar product of $\mathbf{u}$ and $\mathbf{v}$ over $\mathbb{F}_2$. This is not a genuine inner product, but we say nevertheless that a collection $\mathbf{u}_1, \ldots, \mathbf{u}_\ell \in \mathbb{F}_2^k$ is *orthonormal* if $\mathbf{u}_i \cdot \mathbf{u}_i = 1$ for all i and $\mathbf{u}_i \cdot \mathbf{u}_j = 0$ for all $i \neq j$. Finally, we refer to the canonical copy of D_i in $D = [D_1, \ldots, D_k]$ as the *ith 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 [26].

Theorem 4. *Let $D_1, \ldots, D_k$ be oriented graphs with $\mathrm{inv}(D_i) = 1$ for all i and let $\widehat{D} = [D_1, \ldots, D_k]$ be their k-join. Then sets $X_1, \ldots X_k \subseteq V(\widehat{D})$ form a decycling family of $\widehat{D}$ if and only if there are orthonormal vectors $\mathbf{u}_1, \ldots, \mathbf{u}_k \in \mathbb{F}_2^k$ and for each i a decycling set $V_i \subseteq V(D_i)$ of the ith factor of $\widehat{D}$ such that, for each i, the vertices in V_i have characteristic vector $\mathbf{u}_i$ (in $X_1 \ldots, X_k$), and all other vertices have characteristic vector $\mathbf{0}$ (in $X_1 \ldots, X_k$).*

In particular, any acyclic digraph obtained from $\widehat{D}$ 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: $\mathrm{inv}(D) \leq k$?

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 [12, p. 612]), 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 $\mathrm{inv}(D \to D) = 2$ if and only if $\mathrm{inv}(D) = 1$, from which it follows that 2-INVERSION is also NP-complete. They conjectured that NP-completeness extends to k-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 $\mathrm{inv}(D) = 1$ if and only if $\mathrm{inv}([D]_k) = k$, which reduces 1-INVERSION to k-INVERSION and hence shows the following (see Section 4).

Theorem 5. k-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-TOURNAMENT-INVERSION is:

> INPUT: a tournament T.
> PROBLEM: $\mathrm{inv}(T) \leq k$?

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 $\mathrm{inv}(T) = k$ but $\mathrm{inv}(T - \{v\}) < k$ 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 $\mathrm{inv}(D) \leq \mathrm{inv}(D - \{v\}) + 2$. Hence, if $\mathrm{inv}(T) > k$, then by arbitrarily deleting vertices from T one by one, we can obtain a subtournament T' of T with $\mathrm{inv}(T') \in \{k+1, k+2\}$. This T' 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-TOURNAMENT-INVERSION can be solved in time $O(|V(T)|^{\max(m_{k+1}, m_{k+2})})$. Thus, in particular, k-TOURNAMENT-INVERSION can be solved in polynomial time for any fixed k. Plainly $\mathcal{IC}_1 = \{\overrightarrow{C_3}\}$, so $m_1 = 3$, and $\mathcal{IC}_2$ was explicitly described in [26], 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-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-TOURNAMENT-INVERSION can be solved in time $O(|V(T)|^3)$ while 2-TOURNAMENT-INVERSION can be solved in time $O(|V(T)|^6)$. The idea behind their algorithm for 1-TOURNAMENT-INVERSION is to check whether the tournament contains a vertex which can be made into a source, and for 2-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-TOURNAMENT-INVERSION can

be solved in time $O(|V(T)|^{r_k})$. 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 $O(|V(T)|^2)$. In the language of complexity theory, this means that the likely NP-hard problem of determining whether $\mathrm{inv}(T) \leq k$ for inputs k and T (see Conjecture 2) is fixed-parameter tractable when parameterised by k.[2]

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.*

Note that the exponent of n in this running time is optimal, since any algorithm solving k-TOURNAMENT-INVERSION needs to inspect the orientation of every edge in the input tournament. However, 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 $\mathrm{inv}(D)$, $\tau(D)$, and $\tau'(D)$ for a digraph D. Firstly, we have $\mathrm{inv}(D) \leq \tau'(D)$. 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 $(\{v_i, v_j\} \colon i < j, \overrightarrow{v_j v_i} \in F)$ transforms D into an acyclic digraph and hence $\mathrm{inv}(D) \leq \tau'(D)$ 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 [26] to have inversion number k.

Turning to $\tau(D)$, the inequality $\mathrm{inv}(D) \leq 2\tau(D)$ was obtained in [4] as follows. After observing that $\tau(D) = 0$ implies $\mathrm{inv}(D) = 0$, 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 $v \in S$. Then observe that $D - \{v\}$ has cycle transversal number $\tau(D) - 1$, with $S \setminus \{v\}$ a cycle transversal. Moreover, as noted in Section 1.2 we have $\mathrm{inv}(D) \leq \mathrm{inv}(D - \{v\}) + 2$, from which it follows by induction that $\mathrm{inv}(D) \leq 2\tau(D)$.

---

[2]See [14] for the definition of fixed-parameter tractability and an exposition of the surrounding theory.

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 $\overrightarrow{v1}, \overrightarrow{2v}, \overrightarrow{v3}, \overrightarrow{4v}$ 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 $\mathrm{inv}(V_5) = 2$. They noted that if the dijoin conjecture holds, then $\tau([V_5]_k) = k$ and $\mathrm{inv}([V_5]_k) = 2k$ 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 $\mathrm{inv}(D) = 2\tau(D) = 2k$.*

1.4 The extremal problem

Finally, we consider $\mathrm{inv}(n)$, 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[3]

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

Their lower and upper bounds follow from counting and inductive arguments respectively (see Section 7 for details), and they conjectured that $\mathrm{inv}(n) \geq \lfloor \frac{n-1}{2} \rfloor$ 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 $\mathrm{inv}(n) = (1 + o(1))n$.

Theorem 8. *For sufficiently large n we have*

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

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 $\mathrm{inv}(n) \leq n - \log(n + 1)$.

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, Theorem 2, Theorem 3, and Theorem 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 $\mathrm{inv}(n)$ before proving Theorem 8 and giving an improved upper

---

[3] All logarithms in this paper are taken base 2.

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 initial release of this paper, Aubian, Havet, Hörsch, Klingelhoefer, Nisse, Rambaud, and Vermande announced independent work [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 10 holds provided at least one of ℓ and r is odd and at least 3) and they show upper and lower bounds on $\mathrm{inv}(n)$ 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 $\overrightarrow{uv} \in E(T)$. 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 $\mathbf{u}, \mathbf{v} \in \mathbb{F}_2^k$ 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 $\mathbf{u} \cdot \mathbf{v} = 1$.*

This follows from the fact that $\mathbf{u} \cdot \mathbf{v}$ 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 $\operatorname{inv}(D') \leq \operatorname{inv}(D)$.

(ii) For every oriented graph D and every non-negative integer $k \leq \operatorname{inv}(D)$, 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 $\operatorname{inv}(T) = \operatorname{inv}(D)$.

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 $\operatorname{inv}(L \to R) = \operatorname{inv}(L) + \operatorname{inv}(R)$ 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 $\operatorname{inv}(L \to R) < \operatorname{inv}(L) + \operatorname{inv}(R)$. 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 $\operatorname{inv}(L') \geq \operatorname{inv}(L)$ and $\operatorname{inv}(R') \geq \operatorname{inv}(R)$, so we have tournaments L' and R' with $\operatorname{inv}(L' \to R') < \operatorname{inv}(L) + \operatorname{inv}(R) \leq \operatorname{inv}(L') + \operatorname{inv}(R')$.

Proof of Theorem 1. Let L be a copy of $\vec{C_3}$. Suppose that R is a tournament with $\operatorname{inv}(R) = 3$ for which there exist disjoint $A, B, C \subseteq V(R)$ 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

$$\operatorname{inv}(L \to R) = 3 < 4 = \operatorname{inv}(L) + \operatorname{inv}(R).$$

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 $\operatorname{inv}(R) = 3$, as required. □

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 Theorem 2 and Theorem 5. The bulk of the work in our proof of Theorem 4 is put towards proving Lemma 9, which deals with the case $\widehat{D} = [\vec{C_3}]_k$.

Lemma 9. *Let $k \in \mathbb{N}$, let $\widehat{D} = [\vec{C_3}]_k$, and let $X_1, \ldots, X_k \subseteq V(\widehat{D})$ be a decycling family of $\widehat{D}$. Then there exist orthonormal vectors $\mathbf{u}_1, \ldots, \mathbf{u}_k \in \mathbb{F}_2^k$ such that in the ith factor*

of $[\overrightarrow{C_3}]_k$, one vertex has characteristic vector $\mathbf{0}$ and the other two have characteristic vector $\mathbf{u}_i$.

We will use the setup that Pouzet, Kaddour, and Thatte [26] introduced in their proof that $\mathrm{inv}(\widehat{D}) = k$. 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 $\widehat{D}$, and let $<$ be the total order on $V(\widehat{D})$ associated to T. Note that for all i, after inverting $X_1, \ldots, X_k$ the ith 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 ith factor as u_i, v_i, w_i where $\overrightarrow{u_i v_i w_i}$ is a directed 3-cycle in $\widehat{D}$, 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, $\mathbf{u}_i \cdot \mathbf{v}_i = 0$ and $\mathbf{u}_i \cdot \mathbf{w}_i = 1$ where $\mathbf{u}_i, \mathbf{v}_i, \mathbf{w}_i \in \mathbb{F}_2^k$ 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 [26].

Claim 1 ([26])**.** *The vectors* $\mathbf{u}_1, \ldots, \mathbf{u}_k \in \mathbb{F}_2^k$ *are linearly independent.*

Proof. The statement is equivalent to the claim that for all non-empty $I \subseteq [k]$ we have $\sum_{i \in I} \mathbf{u}_i \neq \mathbf{0}$. Fix such an I and note that it is sufficient to show that there exists some $\mathbf{x} \in \mathbb{F}_2^k$ such that $(\sum_{i \in I} \mathbf{u}_i) \cdot \mathbf{x} \neq 0$. Let $m \in I$ be such that $u_i < u_m$ 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 $\mathbf{u}_i \cdot \mathbf{v}_m = \mathbf{u}_i \cdot \mathbf{w}_m$ for all $i \in I \setminus \{m\}$. On the other hand we have $\mathbf{u}_m \cdot \mathbf{v}_m = 0$ while $\mathbf{u}_m \cdot \mathbf{w}_m = 1$, so it follows by linearity of the dot product that $(\sum_{i \in I} \mathbf{u}_i) \cdot \mathbf{v}_m \neq (\sum_{i \in I} \mathbf{u}_i) \cdot \mathbf{w}_m$. One of these two dot products is thus non-zero, and we deduce that $\sum_{i \in I} \mathbf{u}_i \neq \mathbf{0}$, as required. $\square$

We now build on Claim 1 as follows.

Claim 2. *Let* $\ell \in [k]$ *and suppose that the vectors* $\mathbf{u}_i, \mathbf{v}_i, \mathbf{w}_i$ *for* $\ell \leq i \leq k$ *all lie in a subspace* V *of* $\mathbb{F}_2^k$ *of dimension* $k - \ell + 1$. *Then* $\mathbf{u}_\ell, \ldots, \mathbf{u}_k$ *are orthonormal, and for all* $\ell \leq i \leq k$ *we have* $\mathbf{u}_i = \mathbf{w}_i$ *and* $\mathbf{v}_i = \mathbf{0}$.

Proof. We will prove the claim by reverse induction on ℓ. In the $\ell = k$ case the claim follows easily from the fact that $\mathbf{u}_k \cdot \mathbf{w}_k = 1$ while $\mathbf{u}_k \cdot \mathbf{v}_k = 0$. Thus, let $\ell \leq k-1$ and write $[\ell, k]$ for $\{\ell, \ell+1, \ldots, k\}$. Let z be the $<$-minimal vertex among $v_\ell, \ldots, v_k, w_\ell, \ldots, w_k$. Write $\mathbf{z} \in V \subseteq \mathbb{F}_2^k$ 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 $\mathbf{u}_\ell, \ldots, \mathbf{u}_k$ form a basis of V so there exists $I \subseteq [\ell, k]$ such that $\mathbf{z} + \sum_{i \in I} \mathbf{u}_i = \mathbf{0}$.

First suppose that $I \notin \{\emptyset, \{t\}\}$ and let $m \in I$ be such that $u_i < u_m$ for all $i \in I \setminus \{m\}$. If $m \neq t$, then we have $z < v_m, w_m$, so $\mathbf{z} \cdot \mathbf{v}_m = \mathbf{z} \cdot \mathbf{w}_m$ by Observation 1. As in the proof of Claim 1, we have $\mathbf{u}_i \cdot \mathbf{v}_m = \mathbf{u}_i \cdot \mathbf{w}_m$ for all $i \in I \setminus \{m\}$, but $\mathbf{u}_m \cdot \mathbf{v}_m \neq \mathbf{u}_m \cdot \mathbf{w}_m$, so

9

$(\mathbf{z} + \sum_{i \in I} \mathbf{u}_i) \cdot \mathbf{v}_m \neq (\mathbf{z} + \sum_{i \in I} \mathbf{u}_i) \cdot \mathbf{w}_m$, and hence $\mathbf{z} + \sum_{i \in I} \mathbf{u}_i \neq \mathbf{0}$. If $m = t$, then let $j \in I \setminus \{t\}$ and note that $z < v_j, w_j$ by the minimality of z. Consequently, $\mathbf{z} \cdot \mathbf{v}_j = \mathbf{z} \cdot \mathbf{w}_j$. Moreover, since $u_m = u_t < z$, 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, $\mathbf{u}_i \cdot \mathbf{v}_j = \mathbf{u}_i \cdot \mathbf{w}_j$ for all $i \in I \setminus \{j\}$, while $\mathbf{u}_j \cdot \mathbf{v}_j \neq \mathbf{u}_j \cdot \mathbf{w}_j$. Hence, similarly to above, we have $(\mathbf{z} + \sum_{i \in I} \mathbf{u}_i) \cdot \mathbf{v}_j \neq (\mathbf{z} + \sum_{i \in I} \mathbf{u}_i) \cdot \mathbf{w}_j$, so $\mathbf{z} + \sum_{i \in I} \mathbf{u}_i \neq \mathbf{0}$.

The remaining cases are $I = \emptyset$ and $I = \{t\}$, so we have $\mathbf{z} \in \{\mathbf{0}, \mathbf{u}_t\}$. Suppose that $\mathbf{z} = \mathbf{u}_t$. If $z = v_t$, then we have $\mathbf{v_t} = \mathbf{z} = \mathbf{u_t}$, so $\mathbf{v_t} \cdot \mathbf{w_t} = \mathbf{u_t} \cdot \mathbf{w_t} = 1$, i.e. the edge between v_t and w_t undergoes a net reversal under $X_1, \ldots, X_k$. This would imply that $w_t < v_t = z$, which contradicts the minimality of z. Similarly, if $z = w_t$, then since the edge between u_t and v_t is not inverted, neither is the edge between w_t and v_t, so $v_t < w_t = z$, another contradiction. Therefore $\mathbf{z} = \mathbf{0}$. This means no edges incident to z are reversed when $X_1, \ldots, X_k$ are inverted so by the minimality of z we have $z = v_\ell$.

We have shown that $\mathbf{v}_\ell = \mathbf{0}$, so the only vertex among the u_i, v_i, w_i with $i \geq \ell$ which precedes v_ℓ in $<$ is u_ℓ. It follows that u_ℓ is the least element among the u_i, v_i, w_i with $i \geq \ell$. Hence, by Observation 1 we have $\mathbf{u}_\ell \cdot \mathbf{u}_i = \mathbf{u}_\ell \cdot \mathbf{v}_i = \mathbf{u}_\ell \cdot \mathbf{w}_i = 0$ for all $i \geq \ell + 1$, so if V' is the subspace of V spanned by the $\mathbf{u}_i, \mathbf{v}_i, \mathbf{w}_i$ with $i \geq \ell + 1$, then $\mathbf{u}_\ell \cdot \mathbf{x} = 0$ for all $\mathbf{x} \in V'$. We have $\mathbf{u}_\ell \cdot \mathbf{w}_\ell = 1$, so V' is a proper subspace of V, but $\mathbf{u}_{\ell+1}, \ldots, \mathbf{u}_k \in V'$ are linearly independent, so we deduce that V' has dimension $k - \ell$. Therefore by the induction hypothesis $\mathbf{u}_{\ell+1}, \ldots, \mathbf{u}_k$ are orthonormal, and we have $\mathbf{u}_i = \mathbf{w}_i$ and $\mathbf{v}_i = \mathbf{0}$ for $i \geq \ell + 1$.

To complete the induction step it remains to show that $\mathbf{u}_\ell = \mathbf{w}_\ell$ and $\mathbf{u}_\ell \cdot \mathbf{u}_\ell = 1$. The latter follows from the fact that $\mathbf{u}_\ell, \ldots, \mathbf{u}_k$ is a basis for V with $\mathbf{u}_\ell \cdot \mathbf{u}_i = 0$ for all $i \geq \ell + 1$, but $\mathbf{w}_\ell \in V$ has $\mathbf{u}_\ell \cdot \mathbf{w}_\ell = 1$. For the former, note that $\mathbf{w}_\ell = \sum_{i \in I} \mathbf{u}_i$ for some $I \subseteq [\ell, k]$ and by the established properties of the $\mathbf{u}_i$ this set I contains exactly those i for which $\mathbf{u}_i \cdot \mathbf{w}_\ell = 1$. Thus, we certainly have $\ell \in I$. Suppose that $\mathbf{u}_i \cdot \mathbf{w}_\ell = 1$ for some $i \geq \ell + 1$. Since $\mathbf{w}_i = \mathbf{u}_i$ and $\mathbf{v}_i = \mathbf{0}$, by Observation 1 we find that the cycle $\overrightarrow{w_\ell v_i w_i}$ appears in T, which is a contradiction. Hence $I = \{\ell\}$ and $\mathbf{u}_\ell = \mathbf{w}_\ell$, as required. $\qquad\square$

The lemma now follows from the $\ell = 1$ case of Claim 2. $\qquad\square$

We will now deduce Theorem 4 from the lemma. In the proof, we will use the easy fact that every family of orthonormal vectors in $\mathbb{F}_2^k$ is linearly independent.

Proof of Theorem 4. The sufficiency of the given conditions for $X_1, \ldots, X_k$ to be a decycling family of $\widehat{D}$ is straightforward to verify using Observation 1. This observation also allows the 'in particular' part of the theorem statement to be easily deduced from the preceding part. It remains to prove that the given conditions are necessary.

Given a decycling family $X_1, \ldots, X_k$ of $\widehat{D}$, extend $\widehat{D}$ to a tournament T for which $X_1, \ldots, X_k$ is still a decycling family. For each i, let T_i be the subtournament of T induced on the vertex set of the ith factor of $\widehat{D}$. Since D_i contains a directed cycle, so does T_i, and hence the latter contains a copy of $\overrightarrow{C_3}$. We can thus find a copy of $[\overrightarrow{C_3}]_k$ in T whose ith factor is contained in T_i. It follows by Lemma 9 that there are orthonormal vectors $\mathbf{u}_1, \ldots, \mathbf{u}_k \in \mathbb{F}_2^k$ and for each i a triangle $\overrightarrow{u_i v_i w_i}$ in T_i such that u_i and w_i have characteristic vector $\mathbf{u}_i$ and v_i has characteristic vector $\mathbf{0}$ in $X_1, \ldots, X_k$.

We next show that for all i, all vertices in T_i have characteristic vector either $\mathbf{u}_i$ or $\mathbf{0}$ in $X_1, \ldots, X_k$. Let $z \in V(T_i)$ and let $\mathbf{z}$ be its characteristic vector. Since $\mathbf{u}_1, \ldots, \mathbf{u}_k$ form a basis of $\mathbb{F}_2^k$, there exists $J \subseteq [k]$ such that $\mathbf{z} = \sum_{j \in J} \mathbf{u}_j$. If there exists $\ell \in J \setminus \{i\}$, then $\mathbf{z} \cdot \mathbf{u}_\ell = \mathbf{u}_\ell \cdot \mathbf{u}_\ell = 1$ and hence the directions of the edge between z and u_ℓ and the edge between z and w_ℓ are reversed under $X_1, \ldots, X_k$. If $\ell < i$, then the cycle $\overrightarrow{u_\ell v_\ell z}$ appears in T and if $i < \ell$, then the cycle $\overrightarrow{z v_\ell w_\ell}$ appears in T. We have a contradiction in both cases, so $J = \emptyset$ or $J = \{i\}$ as desired.

We have shown that all vertices in the ith factor of $\widehat{D}$ have characteristic vector either $\mathbf{u}_i$ or $\mathbf{0}$ in $X_1, \ldots, X_k$. The effect on this copy of D_i of inverting these sets in $\widehat{D}$ is therefore the same as inverting the set of vertices with characteristic vector $\mathbf{u}_i$, which we call V_i. The latter is therefore a decycling set for the ith factor of $\widehat{D}$. This completes the proof of the theorem. $\qquad\square$

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 $\widehat{D} = [D_1, \ldots, D_k]$ and note that we may assume that none of the D_i have inversion number 0. Indeed, if $\mathrm{inv}(D_i) = 0$ for some $i \geq 2$, then view $\widehat{D}$ as the $(k-1)$-join $[D_1, \ldots, D_{i-2}, D_{i-1} \to D_i, D_{i+1}, \ldots, D_k]$ and, since $\mathrm{inv}(D_{i-1} \to D_i) = \mathrm{inv}(D_{i-1})$, the result follows by induction on k. The case where $i = 1$ can be handled similarly.

Thus, consider the case where $\mathrm{inv}(D_i) = 1$ for all i and suppose for a contradiction that $X_1, \ldots, X_k$ is a decycling family of $\widehat{D}$ 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 $\widehat{D}$ 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, $\mathrm{inv}(\widehat{D}) = k$.

It remains to check the case where $\mathrm{inv}(D_j) = 2$ for some j and $\mathrm{inv}(D_i) = 1$ for all $i \neq j$. Start by letting D_j' be a spanning subdigraph of the jth factor of $\widehat{D}$ with $\mathrm{inv}(D_j') = 1$, then define $\widehat{D}'$ to be the digraph obtained by replacing the jth factor of $\widehat{D}$ by D_j'. Assume for a contradiction that $X_1, \ldots, X_k$ is a decycling family of $\widehat{D}$, in which case it is also a decycling family of $\widehat{D}'$. Theorem 4 thus yields a vector $\mathbf{u}_j \in \mathbb{F}_2^k$ with $\mathbf{u}_j \cdot \mathbf{u}_j = 1$ such that all the vertices in the jth factor of $\widehat{D}'$ (and hence also the jth factor of $\widehat{D}$) have characteristic vector either $\mathbf{0}$ or $\mathbf{u}_j$ in $X_1, \ldots, X_k$. Inverting $X_1, \ldots, X_k$ in $\widehat{D}$ therefore has the same effect on its jth factor as inverting the set of vertices with characteristic vector $\mathbf{u}_j$. It follows that this set of vertices is a decycling set for D_j, contradicting $\mathrm{inv}(D_j) = 2$. $\qquad\square$

As mentioned in the introduction, it follows from Theorem 3 that for any digraph D we have $\mathrm{inv}(D) = 1$ if and only if $\mathrm{inv}([D]_k) = k$, which in turn implies Theorem 5 (which states that k-INVERSION is NP-complete for all $k \in \mathbb{N}$). Indeed, Theorem 3 directly gives $\mathrm{inv}([D]_k) = k$ in the case $\mathrm{inv}(D) = 1$, and if $\mathrm{inv}(D) = 0$ then clearly $\mathrm{inv}([D]_k) = 0$. If $\mathrm{inv}(D) > 1$, then there are subdigraphs D' and D'' of D with $\mathrm{inv}(D') = 1$ and

$\text{inv}(D'') = 2$. The k-join $D'' \to [D']_{k-1}$, 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 (which states that $\text{inv}(L \to R) = 4$ for all digraphs L and R with inversion number 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 $\text{inv}(L \to R) \geq 4$.

Proof of Theorem 2. Let L and R be digraphs with $\text{inv}(L) = \text{inv}(R) = 2$. It is immediate that $\text{inv}(L \to R) \leq 4$, so it is sufficient to prove the lower bound. For this, extend $L \to R$ 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 $L' \to R'$.

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 $L_i \to R_j$, we have $\text{inv}(T) \geq \text{inv}(L_i \to R_j) \geq 4$ 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 $\text{inv}(T) \geq 4$. $\qquad\square$

5 Complexity of k-TOURNAMENT-INVERSION

In this section we prove Theorem 6 by constructing, for each fixed $k \in \mathbb{N}$, an algorithm solving k-TOURNAMENT-INVERSION in time $O(|V(T)|^2)$. Our proof uses a technique known as iterative compression; see [14] for a description of this method and other applications of it. The most involved part of our proof concerns the 'compression step' of the algorithm. This step is handled by 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 of size k if one exists. Throughout this section, we represent a total order $<$ on a finite set $S = \{s_1, \ldots, s_m\}$ by the tuple $(s_1, \ldots, s_m)$ where $s_1 < \cdots < s_m$.

Lemma 10. *Fix $k, s \in \mathbb{N}$. There is an algorithm which solves the following problem for n-vertex tournaments in time $O(n)$:*

INPUTS:
- a tournament T_0;
- a decycling family $X_1, \ldots, X_s$ of T_0 (transforming T_0 into T, say);
- the order on $V(T_0)$ associated to T.

OUTPUTS:
EITHER
- that T_0 is not k-invertible;

OR
- 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'.

We now use iterative compression to prove Theorem 6 before returning to Lemma 10.

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.

OUPUTS:
EITHER
- that T_0 is not k-invertible;

OR
- 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. $\square$

It is left to prove Lemma 10. To this end, we describe an algorithm which explores what happens if, starting from T, we invert $X_1, \ldots, X_s$ and k further sets $Y_1, \ldots, Y_k$ to

13

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.

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 $\mathbf{u}_i \in \mathbb{F}_2^{s+k}$ 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 $\mathbf{u} = (\mathbf{u}_1, \ldots, \mathbf{u}_n)$. There is a bijective correspondence between Y and $\mathbf{u}$ and it will be more convenient to work with the latter, so let $T_\mathbf{u} = T_Y$ and write $\mathcal{U}$ for the set of all possible $\mathbf{u}$. Our first aim is to determine in linear time whether there exists $\mathbf{u} \in \mathcal{U}$ such that $T_\mathbf{u}$ is transitive, and to identify such a $\mathbf{u}$ if so.

The tournament $T_\mathbf{u}$ 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 $\mathbf{u}_a \cdot \mathbf{u}_b = \mathbf{u}_b \cdot \mathbf{u}_c$ but $\mathbf{u}_a \cdot \mathbf{u}_b \neq \mathbf{u}_a \cdot \mathbf{u}_c$. We describe the triple $(\mathbf{u}_a, \mathbf{u}_b, \mathbf{u}_c)$ as *bad* if this occurs. Thus, $T_\mathbf{u}$ is transitive if and only if $B(\mathbf{u}) = \{(\mathbf{u}_a, \mathbf{u}_b, \mathbf{u}_c) : a < b < c\}$ contains no bad triples, and T_0 is k-invertible if and only if $\mathcal{B} = \{B(\mathbf{u}) : \mathbf{u} \in \mathcal{U}\}$ 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 $\mathbf{u} \in \mathcal{U}$ for which $B = B(\mathbf{u})$.

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 $\mathbf{u}' = (\mathbf{u}_1, \ldots, \mathbf{u}_{n-1})$ of characteristic vectors of $u_1, \ldots, u_{n-1}$ in $X_1, \ldots, X_s, Y_1, \ldots, Y_k$. For $\mathbf{u}' \in \mathcal{U}'$, let $B'(\mathbf{u}') = \{(\mathbf{u}_a, \mathbf{u}_b, \mathbf{u}_c) : 1 \leq a < b < c \leq n - 1\}$ and let $\mathcal{B}' = \{B'(\mathbf{u}') : \mathbf{u}' \in \mathcal{U}'\}$. 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 $\mathbf{u}' \in \mathcal{U}'$ such that $B' = B'(\mathbf{u}')$. 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 $\mathbf{u} \in \mathcal{U}$ such that $B = B(\mathbf{u})$.

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 $2^{2^{3(s+k)}}$. In particular, there are only boundedly many pairs $(B', \mathbf{u}_n)$ where $B' \in \mathcal{B}'$ and $\mathbf{u}_n$ is a possible characteristic vector for u_n. For each such pair, we can construct in bounded time the set $S(B', \mathbf{u}_n)$ consisting of all triples in B', and all triples of the form $(\mathbf{v}_i, \mathbf{v}_j, \mathbf{u}_n)$ for $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3) \in B'$ and $1 \leq i < j \leq 3$. It is not hard to see that $\mathcal{B}$ equals the set of all sets $S(B', \mathbf{u}_n)$ and that each $S(B', \mathbf{u}_n)$ can be associated with the $\mathbf{u} \in \mathcal{U}$ formed by appending $\mathbf{u}_n$ to the $\mathbf{u}' \in \mathcal{U}'$ associated with B'. Indeed, given B' and $\mathbf{u}_n$ and defining $\mathbf{u}$ as in the previous sentence, since $n \geq 4$, we

have $S(B', \mathbf{u}_n) = B(\mathbf{u})$. For the other direction, given $\mathbf{u} = (\mathbf{u_1}, \ldots, \mathbf{u}_n) \in \mathcal{U}$ and letting $\mathbf{u}' = (\mathbf{u_1}, \ldots, \mathbf{u}_{n-1})$, we have $B(\mathbf{u}) = S(B'(\mathbf{u}'), \mathbf{u}_n)$.

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 $\mathbf{u} \in \mathcal{U}$ such that $B = B(\mathbf{u})$, 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 $\mathbf{u}$ 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 for each atom A the restriction to A of the order associated to T. By reversing the order on each atom whenever the edges within it undergo a net reversal under the inversions, we obtain the order on that atom associated to T'. 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. $\qquad\square$

Note that the implicit constant in the running time given by this proof is doubly exponential in $s + k$.

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 [17].

Theorem 11 (Eventown [8], [17]). *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 $|\mathcal{F}| \le 2^{\lfloor n/2 \rfloor}$.*

For a digraph D and vertices $u, v, w \in V(D)$, we will say that u and v *differ* on w if either $\overrightarrow{uw}, \overrightarrow{wv} \in E(D)$ or $\overrightarrow{vw}, \overrightarrow{wu} \in E(D)$. 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 $\overrightarrow{v_i v_j}$ 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 $\overrightarrow{u_i v_j}$ if in the binary expansion of j, the digit in the 2^i place is a 0, and add the edge $\overrightarrow{v_j u_i}$ 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 $\mathrm{inv}(D) \geq 2k$, as then $\mathrm{inv}(D) = 2\tau(D) = 2k$ follows from $\mathrm{inv}(D) \leq 2\tau(D)$. Suppose for a contradiction that $X_1, \ldots, X_{2k-1} \subseteq V(D)$ form a decycling 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 $v_0, \ldots, v_{K-1}$ is in a different atom. Suppose for a contradiction that v_i and v_j are in the same atom for some $i < j$ in $\{0, \ldots, K-1\}$, and note that v_{K+i} and 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 v_i and v_j differ on u_ℓ in D. Since they are in the same atom as each other, they also differ on u_ℓ 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 v_i and 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 v_i and v_{K+i} are in the same atom and are either both in-neighbours of u_ℓ 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 v_i and 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 $v_0, \ldots, v_{K-1}$ 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. $\quad\square$

7   Bounds on $\mathrm{inv}(n)$

7.1   Lower bounds

In this section we discuss the previous known lower bound on $\mathrm{inv}(n)$ 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 $\mathrm{inv}(n)$. 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 $\mathrm{inv}(n) \geq k$. Taking logarithms base 2 and rearranging, this condition becomes $k < (n-1)/2 - \log(n!)/n$, so we have

$$\mathrm{inv}(n) \geq \left\lfloor \frac{n-1}{2} - \frac{\log(n!)}{n} \right\rfloor \geq \left\lfloor \frac{n-1}{2} - \log(n) \right\rfloor,$$

where for the final inequality we used $n! \leq n^n$. Lower bounds on $\mathrm{inv}(n)$ 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 [23], 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 ith step choosing the first i entries of the ith row of the matrix (and also, by symmetry, the ith 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_jth 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 $\binom{n}{s-1}$ 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

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

as required. $\quad\square$

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 $\overrightarrow{ab}$ 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 $s = \left\lfloor \sqrt{2n \log(n)} \right\rfloor$ and write $k = n - s$. Suppose that $\mathrm{inv}(T) \le k$ 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 $\mathrm{rank}(M_i) \le 1$ for all i, and thus $\mathrm{rank}(\sum_i M_i) \le k$. 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 $\mathrm{inv}(T) \le k$, then there exists $M \in \mathcal{M}$ such that $\mathrm{rank}(M_T + M) \le k$. 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

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

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

$$\begin{aligned}
f(n) &= \frac{\log(n)}{2} + n \log(n) - n \log(e) + n + s \log(n) - \binom{s}{2} \\
&= -n(\log(e) - 1) + o(n),
\end{aligned}$$

and thus $\mathbb{P}(\mathrm{inv}(T) \le k) \to 0$ as $n \to \infty$ as desired. $\qquad\square$

7.2 Upper bounds

The only approach which has been used to prove upper bounds on $\mathrm{inv}(n)$, 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 $\mathrm{inv}(n - 1)$ inversions we can transform $T_1 - \{v\}$ into a transitive tournament, so $\mathrm{inv}(n) \le \mathrm{inv}(n - 1) + 1$ for all $n \ge 2$. The authors of [7] observed that $\mathrm{inv}(4) = 1$, so $\mathrm{inv}(n) \le n - 3$ for $n \ge 4$. For $n \ge 6$ this was improved by 1 in [4] using the fact that $\mathrm{inv}(6) = 2$ (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 $n \in \mathbb{N}$,*

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

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 $\mathrm{inv}(A) \leq \mathrm{inv}(|A|)$ inversions we can transform T' into a transitive tournament. Thus, $\mathrm{inv}(T) \leq |B| + \mathrm{inv}(|A|)$.

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

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

and the claim follows. $\qquad\square$

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

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

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

$$\begin{aligned}
\mathrm{inv}(n) &\leq \left\lfloor \frac{n-1}{2} \right\rfloor + \mathrm{inv}\left(\left\lceil \frac{n-1}{2} \right\rceil\right), \\
&\leq \left\lfloor \frac{n-1}{2} \right\rfloor + \left\lceil \frac{n-1}{2} \right\rceil - \log\left(\left\lceil \frac{n-1}{2} \right\rceil + 1\right), \\
&\leq n - 1 - \log\left(\frac{n+1}{2}\right), \\
&= n - \log(n+1).
\end{aligned}$$

$\qquad\square$

8 Conclusion

In this paper we have answered several of the questions posed in [4]. We have shown that their 'dijoin conjecture', that $\mathrm{inv}(L \to R) = \mathrm{inv}(L) + \mathrm{inv}(R)$, is false in general, but have verified it in the case where $\mathrm{inv}(L) = \mathrm{inv}(R) = 2$ 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 $\mathrm{inv}(D) \leq 2\tau(D)$ 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 $\mathrm{inv}(n)$ to show that $\mathrm{inv}(n) = (1+o(1))n$. 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 π. The *reversal distance*, $d(\pi)$, of a permutation π is the minimum number of reversals required to transform π into the identity permutation. For a survey of reversals and the reversal distance (and many other combinatorial models of genome rearrangements) see [15]. We highlight some results of particular relevance to our work. With regards to computational complexity, Caprara [9] showed that the problem of SORTING BY REVERSALS, that is, determining whether $d(\pi) \leq k$ for inputs of a permutation π and $k \in \mathbb{N}$, is NP-complete, while Hannenhalli and Pevzner [18, 19] showed that it is fixed-parameter tractable when parameterised by k. The natural extremal problem was solved by Bafna and Pevzner [3] who proved that for a permutation π 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 γ_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 [21] 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 [20] and it remains NP-complete when the input is restricted to tournaments [1, 10]. However, Chen, Liu, Lu, O'Sullivan, and Razgon [11] showed that this problem is again fixed-parameter tractable when parameterised by k. On the extremal side, it was shown by Spencer [27, 28] that the maximum cycle edge-transversal number of an n-vertex tournament is $\frac{1}{2}\binom{n}{2} - \Theta(n^{3/2})$ 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 [13, 25]).

8.2 Open problems

We have shown (in Theorem 6) that the problem which takes as inputs a tournament T and an integer $k \in \mathbb{N}$, and asks whether $\mathrm{inv}(T) \leq k$, is fixed-parameter tractable when parameterised by k. In keeping with the pattern exhibited in the settings discussed in Section 8.1, Bang-Jensen, da Silva, and Havet [4] conjectured that the full problem is NP-complete.

Conjecture 2 ([4]). *The problem of deciding whether* $\mathrm{inv}(T) \leq k$ *for inputs of* $k \in \mathbb{N}$ *and a tournament* T *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 is not polynomial in k. In fact, as noted above, the constant arising from our algorithm is doubly exponential in k. However,

again in keeping with both settings discussed in Section 8.1 (and indeed many natural fixed-parameter tractable problems), we conjecture that this constant can be taken to be singly exponential in k, perhaps with a higher power of n.

Conjecture 3. *There exist constants $c_1, c_2 > 0$ such that, for any $k \in \mathbb{N}$, k-TOURNAMENT-INVERSION can be solved in time $O(2^{k^{c_1}}|V(T)|^{c_2})$.*

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 [16] (see [24] for an English translation) and Latka [22], 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 4 ([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 $\mathrm{inv}(n)$. The best known bounds on $\mathrm{inv}(n)$ for large n are now

$$n - \sqrt{2n \log(n)} \leq \mathrm{inv}(n) \leq n - \log(n+1),$$

and it would be interesting to tighten these further.

Question 5. *What is the asymptotic behaviour of $n - \mathrm{inv}(n)$?*

In light of our improved lower bound on $\mathrm{inv}(n)$, 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 $[\overrightarrow{C_3}]_{n/3}$) is known.

Problem 6. *Construct n-vertex tournaments with inversion number closer to $\mathrm{inv}(n)$.*

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 $\mathrm{inv}(Q_n) = \left\lfloor \frac{n-1}{2} \right\rfloor$.

Conjecture 7 ([6]). *For all $n \in \mathbb{N}$ we have $\mathrm{inv}(Q_n) = \left\lfloor \frac{n-1}{2} \right\rfloor$.*

The conjecture is known to hold for $n \leq 8$ [4, 7], and it is certainly true that $\mathrm{inv}(Q_n) \leq \left\lfloor \frac{n-1}{2} \right\rfloor$ 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*, $\mathrm{inv}(T, T')$, 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

$2^{\binom{n}{2}+n-\binom{s}{2}+s\log(n)}$ 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 $\mathcal{W}$ 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 8. *What is the mixing time of $\mathcal{W}$? 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 in the cases where $\mathrm{inv}(L), \mathrm{inv}(R) \leq 2$. 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 9. *Let $k \in \mathbb{N}$ and let $D_1, \ldots, D_k$ be oriented graphs satisfying $\mathrm{inv}(D_i) \leq 2$ for all i. Then*

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

On the other hand, Theorem 1 gives a counterexample to the dijoin conjecture where $\mathrm{inv}(L) = 1$ and $\mathrm{inv}(R) = 3$. From this, we can obtain counterexamples with $\mathrm{inv}(L) = k$ and $\mathrm{inv}(R) = 3$ for any $k \in \mathbb{N}$: let $L = [\vec{C_3}]_k$ 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 $\mathrm{inv}(L) = 3$ and $\mathrm{inv}(R) = k$. We conjecture that here 3 can be replaced with any larger integer, or in other words that the only values of $\mathrm{inv}(L)$ and $\mathrm{inv}(R)$ for which the dijoin conjecture always holds are those where $\mathrm{inv}(L), \mathrm{inv}(R) \leq 2$ or where one of $\mathrm{inv}(L)$ or $\mathrm{inv}(R)$ is 0.

Conjecture 10. *For all $\ell, r \in \mathbb{N}$ with $\ell \geq 3$ or $r \geq 3$ there exist oriented graphs L and R with $\mathrm{inv}(L) = \ell$ and $\mathrm{inv}(R) = r$, but $\mathrm{inv}(L \to R) < \ell + r$.*

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

Finally, we noted in Section 1.2 that $\mathrm{inv}(D) \leq \mathrm{inv}(D - \{v\}) + 2$ for all digraphs D and vertices $v \in V(D)$. It is certainly the case that this inequality is tight for some D and v. 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 $\mathrm{inv}(D - T) = \mathrm{inv}(D) - 2|T|$. We conjecture, however, that the inequality $\mathrm{inv}(D) \leq \mathrm{inv}(D - \{v\}) + 2$ cannot be tight for all vertices v in a given digraph D.

Conjecture 11. *Let D be a digraph with at least one vertex. Then there exists $v \in V(D)$ such that $\mathrm{inv}(D - \{v\}) \geq \mathrm{inv}(D) - 1$.*

Acknowledgments

References

[1] N. Alon. Ranking tournaments. *SIAM Journal on Discrete Mathematics*, **20**(1):137–142, 2006.

[2] G. Aubian, F. Havet, F. Hörsch, F. Klingelhoefer, N. Nisse, C. Rambaud and Q. Vermande. Problems, proofs, and disproofs on the inversion number. *arXiv:2212.09188* preprint, 2022.

[3] V. Bafna and P. A. Pevzner. Genome rearrangements and sorting by reversals. *SIAM Journal on Computing*, **25**(2):272–289, 1996.

[4] J. Bang-Jensen, J. C. F. da Silva and F. Havet. On the inversion number of oriented graphs. *Discrete Mathematics & Theoretical Computer Science*, **23**(2), 2022.

[5] H. Belkhechine. *Indécomposabilité des graphes et des tournois*. Theses, Université Claude Bernard - Lyon I ; Université de Sfax. Faculté des sciences, July 2009.

[6] H. Belkhechine, M. Bouaziz, I. Boudabbous and M. Pouzet. Inversions in tournaments. Unpublished.

[7] H. Belkhechine, M. Bouaziz, I. Boudabbous and M. Pouzet. Inversion dans les tournois. *Comptes Rendus Mathematique*, **348**(13):703–707, 2010.

[8] E. Berlekamp. On subsets with intersections of even cardinality. *Canadian Mathematical Bulletin*, **12**(4):471–474, 1969.

[9] A. Caprara. Sorting by reversals is difficult. In *Proceedings of the First Annual International Conference on Computational Molecular Biology*, RECOMB '97, page 75–83, New York, NY, USA, 1997. Association for Computing Machinery.

[10] P. Charbit, S. Thomassé and A. Yeo. The minimum feedback arc set problem is NP-hard for tournaments. *Combinatorics, Probability and Computing*, **16**(1):1–4, 2007.

[11] J. Chen, Y. Liu, S. Lu, B. O'sullivan and I. Razgon. A fixed-parameter algorithm for the directed feedback vertex set problem. In *Proceedings of the fortieth annual ACM symposium on Theory of computing*, pages 177–186, 2008.

[12] T. H. Cormen, C. E. Leiserson, R. L. Rivest and C. Stein. *Introduction to Algorithms, Third Edition*. The MIT Press, 3rd edition, 2009.

[13] W. F. de la Vega. On the maximum cardinality of a consistent set of arcs in a random tournament. *Journal of Combinatorial Theory, Series B*, **35**(3):328–332, 1983.

[14] R. G. Downey and M. R. Fellows. *Fundamentals of Parameterized Complexity*. Springer London, 2013.

[15] G. Fertin, A. Labarre, I. Rusu, E. Tannier and S. Vialette. *Combinatorics of Genome Rearrangements*. The MIT Press, 2009.

[16] T. Gallai. Transitiv orientierbare Graphen. *Acta Mathematica Academiae Scientiarum Hungarica*, **18**:25–66, 1967.

[17] J. E. Graver. Boolean designs and self-dual matroids. *Linear Algebra and its Applications*, **10**(2):111–128, 1975.

[18] S. Hannenhalli and P. Pevzner. To cut...or not to cut (applications of comparative physical maps in molecular evolution). In *Proceedings of the Seventh Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '96, page 304–313, USA, 1996. Society for Industrial and Applied Mathematics.

[19] S. Hannenhalli and P. A. Pevzner. Transforming cabbage into turnip: Polynomial algorithm for sorting signed permutations by reversals. *J. ACM*, **46**(1):1–27, jan 1999.

[20] R. M. Karp. Reducibility among combinatorial problems. In R. E. Miller, J. W. Thatcher and J. D. Bohlinger, editors, *Complexity of Computer Computations: Proceedings of a symposium on the Complexity of Computer Computations, held March 20–22, 1972, at the IBM Thomas J. Watson Research Center, Yorktown Heights, New York, and sponsored by the Office of Naval Research, Mathematics Program, IBM World Trade Corporation, and the IBM Research Mathematical Sciences Department*, pages 85–103, Boston, MA, 1972. Springer US.

[21] R. Kudelić. *Feedback Arc Set*. SpringerBriefs in Computer Science. Springer Cham, 1st edition, 2022.

[22] B. J. Latka. Structure theorem for tournaments omitting N_5. *Journal of Graph Theory*, **42**(3):165–192, 2003.

[23] J. MacWilliams. Orthogonal matrices over finite fields. *The American Mathematical Monthly*, **76**(2):152–164, 1969.

[24] F. Maffray and M. Preissmann. A translation of Gallai's paper: "Transitiv orientierbare Graphen". In J. L. Ramírez-Alfonsín and B. A. Reed, editors, *Perfect Graphs*, chapter 3, pages 25–66. Wiley, 2001.

[25] S. Poljak, V. Rödl and J. Spencer. Tournament ranking with expected profit in polynomial time. *SIAM Journal on Discrete Mathematics*, **1**(3):372–376, 1988.

[26] M. Pouzet, H. S. Kaddour and B. Thatte. On the Boolean dimension of a graph and other related parameters. *Discrete Mathematics & Theoretical Computer Science*, **23**(2), 2022.

[27] J. Spencer. Optimal ranking of tournaments. *Networks*, **1**(2):135–138, 1971.

[28] J. Spencer. Optimally ranking unrankable tournaments. *Periodica Mathematica Hungarica*, **11**(2):131–144, 1980.