Polynomial Kernel for Immersion Hitting in Tournaments

Łukasz Bożyk ✉ ⓘ
University of Warsaw, Poland

Michał Pilipczuk ✉ ⓘ
University of Warsaw, Poland

───── **Abstract** ─────

For a fixed simple digraph H without isolated vertices, we consider the problem of deleting arcs from a given tournament to get a digraph which does not contain H as an immersion. We prove that for every H, this problem admits a polynomial kernel when parameterized by the number of deleted arcs. The degree of the bound on the kernel size depends on H.

1 Introduction

Kernelization is an algorithmic framework for describing preprocessing procedures that given an instance of a hard problem, identify and reduce easily resolvable parts. The usual formalization of the concept is based on the paradigm of parameterized complexity. A *kernelization procedure*, or a *kernel* for short, for a parameterized decision problem L is a polynomial-time algorithm that given an instance (I, k) of L, where k is the parameter, outputs an equivalent instance (I', k') such that both $|I'|$ and k' are bounded by a computable function of k. If this function is a polynomial in k, we say that the kernel is *polynomial*. The search for (polynomial) kernels is an established research area within the field of parameterized algorithms. We refer to the textbook of Fomin et al. [13] for a broader discussion of classic results and techniques.

A particularly fruitful line of research within kernelization concerns the methodology of *protrusions* and *protrusion replacement*. The idea is to find a large *protrusion*: a piece of graph that is simple – for instance, has bounded treewidth – and communicates with the rest of the graph only through a small interface. If found, a protrusion can be fully understood – for instance, using dynamic programming on its tree decomposition – and replaced with a smaller one with the same functionality. So if one proves that, provided the given instance is large, a large protrusion can be efficiently found and replaced with a strictly smaller one, then applying this strategy exhaustively eventually arrives at a kernel. Protrusion-based techniques were pioneered by Bodlaender et al. [5], but by now have become a part of the standard toolbox of kernelization. We refer the interested reader to [13, Part 2] for more information.

A particularly important achievement in the development of protrusion-based kernelization procedures is the result of Fomin et al. [12], who gave a polynomial kernel for the PLANAR $\mathcal{F}$-DELETION problem, defined as follows. Let $\mathcal{F}$ be a fixed family of graphs containing

at least one planar graph. Then in the problem we are given a graph G and an integer k (considered to be the parameter), and the question is whether one can hit all minor models of graphs from $\mathcal{F}$ in G using a hitting set consisting of at most k vertices. Fomin et al. gave a polynomial kernel for this problem for every fixed family $\mathcal{F}$ as above. The degree of the polynomial bound on the kernel size depends on $\mathcal{F}$ and it is known that under certain complexity-theoretical assumptions, this is unavoidable [15]. The assumption that $\mathcal{F}$ contains a planar graph is crucial in the approach: under this assumption, graphs not containing any graph from $\mathcal{F}$ as a minor have treewidth bounded by a constant, which unlocks a multitude of tools related to tree decomposition using which one can understand the structure of the instance.

The concept of a protrusion, as described above, is quite capacious and can be applied in different settings as well. For instance, Giannopolou et al. [18] considered the $\mathcal{F}$-Immersion Deletion problem, where for a given graph G and parameter k, one wishes to hit all immersion models of graphs from $\mathcal{F}$ using a hitting set of edges of size at most k (Recall that an *immersion model* of a graph H in a graph G consists of mapping the vertices of H to distinct vertices of G and edges of H to pairwise edge-disjoint paths in G so that the image of an edge uv connects the image of u with the image of v.). By loosely following the approach of [12], Giannopolou et al. [18] gave a *linear* kernel for $\mathcal{F}$-Immersion Deletion for every family $\mathcal{F}$ that contains a subcubic planar graph. Here, the main idea was to adjust the notions of protrusions to the graph parameter *tree-cutwidth* and corresponding *tree-cut decompositions*, which play the same role for immersions as treewidth and tree decompositions play for minors.

Motivated by this, it is interesting to consider other settings where the protrusion methodology could be applied. A particularly tempting area is that of *directed graphs*, where natural analogues of problems considered in undirected graphs can be easily stated. Unfortunately, the structural theory of directed graphs is much less understood than that of undirected graphs, and many problems become inherently harder; see e.g. [16, 17, 20]. In particular, there is even a scarcity of fixed-parameter tractability results, not to mention kernelization results.

However, there is a particular class of directed graphs where a sound structural theory has been developed: *tournaments*. (Here, recall that a *tournament* is a directed graph where every pair of vertices is connected by exactly one arc.) This theory[1] was pioneered by Chudnovsky, Ovetsky Fradkin, Kim, and Seymour [9, 10, 26, 27, 21], while structural and algorithmic aspects connected to parameterized complexity were investigated by Fomin and Pilipczuk [14]. See the introductory section of [14] for an overview.

In particular, as proved in the aforementioned works, there are two main width notions for tournaments: *cutwidth* and *pathwidth*. The first one is tightly connected to (directed) immersions as follows: if a tournament T excludes a digraph H as an immersion, then the cutwidth of T is bounded by a constant depending only on H. Pathwidth is connected to *topological minors* and *strong minors* in the same sense. These structural results were used for the design of parameterized algorithms for containment problems in tournaments in [9, 26, 14]. Later, they were used by Raymond [28] and by Bożyk and Pilipczuk [8] to establish Erdős-Pósa properties for immersions and topological minors in tournaments.

The goal of this work is to explore the applicability of the structural theory of tournaments for kernelization, with a particular focus of developing a sound protrusion-based methodology.

---

[1] In this line of work, most results concern the class of *semi-complete digraphs*, which differ from tournaments by allowing that a pair of vertices can be also connected by two oppositely-oriented arcs. In this article we focus on the setting of tournaments for simplicity.

Our contribution.   For a simple directed graph H without isolated vertices we define the following parameterized problem:

H-HITTING IMMERSIONS IN TOURNAMENTS
Input: A tournament T and a positive integer k.   **Parameter:** k **Output:** Is there a set $F \subseteq A(T)$, such that $|F| \leqslant k$ and $T - F$ is H-immersion-free?

That this problem is fixed-parameter tractable is proved in [14]. Our main result states that for every fixed H, H-HITTING IMMERSIONS IN TOURNAMENTS admits a polynomial kernel, of degree dependent on H. Formally, we prove the following theorem.

▶ **Theorem 1.** *For every simple digraph H without isolated vertices there exists a constant c and an algorithm that given an instance (T, k) of H-HITTING IMMERSIONS IN TOURNAMENTS, runs in polynomial time and returns an equivalent instance (T', k) with $|T'| \leqslant c \cdot k^c$.*

We remark that when H is the directed triangle, H-HITTING IMMERSIONS IN TOURNAMENTS is equivalent to the FEEDBACK ARC SET IN TOURNAMENTS (FAST) problem. There is a sizeable literature on the parameterized complexity of FAST, see e.g. [1, 11, 14, 19], mostly due to the fact that it admits a subexponential parameterized algorithm with running time $2^{\mathcal{O}(\sqrt{k})} \cdot n^{\mathcal{O}(1)}$. Kernelization procedures for FAST were investigated by Bessy et al. in [4], while kernelization of the dual problems of packing arc-disjoint triangles and packing arc-disjoint cycles in tournaments were recently studied by Bessy et al. in [3].

On a very high conceptual level, the proof of Theorem 1 follows the classic blueprint of protrusion-based kernelization, like in e.g. [12, 18]. That is, if (T, k) is a given instance of H-HITTING IMMERSIONS IN TOURNAMENTS, we perform the following steps.

- We may assume that the cutwidth of T is bounded polynomially in k, for otherwise in T one can find $k + 1$ arc-disjoint immersion models of H; these witness a negative answer to the instance.
- Assuming that T is large – of size superpolynomial in k – but has cutwidth bounded polynomially in k, we may find in T a large *protrusion*. Here, a protrusion is an interval I in the vertex ordering σ witnessing small cutwidth such that σ restricted to I witnesses that $T[I]$ has constant cutwidth, and there is only a constant number of σ-backward arcs with one endpoint in I and the other outside of I. These are instantiations of the two desired properties of a protrusion: it has to have bounded width and communicate with the rest of the graph through a boundary of bounded size.
- We can replace the obtained protrusion with a strictly smaller one of the same "type", thus obtaining a strictly smaller equivalent instance. Applying this strategy exhaustively eventually yields a kernel of polynomial size.

Compared to the previous works, the main difficulty is to tame the interaction between a protrusion and the remainder of the instance. Namely, this interaction is not restricted to a set of vertices or arcs of constant size: as we work with tournaments, every vertex of a protrusion will necessarily have an arc connecting it to every vertex outside of the protrusion. The idea is that all but a constant number of those arcs will be forward arcs in the fixed vertex ordering σ with bounded cutwidth. We call those well-behaved forward arcs *generic*, while the remaining constantly many backward arcs are *singular*. Understanding the interaction between a protrusion and the rest of the tournament as being governed by few singular arcs and a large number of well-behaved generic arcs is the crux of our approach.

In particular, while looking for a large replaceable protrusion, we have to be extremely careful when arguing about how such a protrusion may interact with optimum solutions. Here, a key step is to find several protrusions that appear consecutively in σ (recall that our protrusions are intervals in σ), have the same *type* (in the sense of admitting partial immersions of H), and such that their union is a protrusion of again the same type. This step is done using Simon Factorization, a tool commonly used in the theory of automata and formal languages. Simon Factorization was recently used a few times in structural graph theory [7, 23, 24], but we are not aware of any previous application in the context of kernelization.

The application of Simon Factorization is also the only step in the reasoning that causes the degree of the polynomial bounding the size of our kernel to depend on H. It is an interesting open question whether this can be improved, or in other words, whether there is a kernel of size at most $c \cdot k^d$, where c may depend on H but d does not. Judging by the results on hitting immersions in undirected graphs [18], we expect that this might be the case.

We remark that the sign $\not\prec$ denotes statements whose proofs are deferred to the full version of the paper.

2 Preliminaries

We use the standard terminology and notation for describing immersions in tournaments and for cutwidth of digraphs and of tournaments. This terminology and notation is borrowed mostly, and often in a verbatim form, from the work of Bożyk and Pilipczuk [8].

For a positive integer n, we denote $[n] := \{1, \ldots, n\}$ and $[-n] = \{-1, \ldots, -n\}$.

We use standard graph terminology and notation. All graphs considered in this paper are finite, simple (i.e. without self-loops or multiple arcs with same head and tail), and directed (i.e. are *digraphs*). For a digraph D, by $V(D)$ and $A(D)$ we denote the vertex set and the arc set of D, respectively. We denote

$$|D| := |V(D)| \qquad \text{and} \qquad \|D\| := |A(D)|.$$

For $X \subseteq V(D)$, the subgraph *induced* by X, denoted $D[X]$, comprises of the vertices of X and all the arcs of D with both endpoints in X. By $D - X$ we denote the digraph $D[V(D) \setminus X]$. Further, if F is a subset of arcs of D, then by $D - F$ we denote the digraph obtained from D by removing all the arcs of F. For $X, Y \subseteq V(D)$ we denote by $\overrightarrow{A}(X, Y)$ the set of all arcs $(u, v) \in A(D)$ such that $u \in X$ and $v \in Y$ and moreover $A(X, Y) := \overrightarrow{A}(X, Y) \cup \overrightarrow{A}(Y, X)$. For an arc $a = (u, v) \in A(D)$ we define $\mathsf{tail}(a) := u$ and $\mathsf{head}(a) := v$. For a directed (not necessarily simple) path P we denote by $\mathsf{first}(P)$ and $\mathsf{last}(P)$ the first and the last arcs on path P, respectively.

Tournaments. A simple digraph $T = (V, A)$ is called a *tournament* if for every pair of distinct vertices $u, v \in V$, either $(u, v) \in A$, or $(v, u) \in A$ (but not both). Alternatively, one can represent the tournament T by providing a pair $(\sigma, \overleftarrow{A}_\sigma(T))$, where $\sigma \colon V \to [|V|]$ is an *ordering* of the set V and $\overleftarrow{A}_\sigma(T)$ is the set of σ-*backward arcs*, that is,

$$\overleftarrow{A}_\sigma(T) := \{ (u, v) \in A \mid \sigma(u) > \sigma(v) \}.$$

All the remaining arcs are called σ-*forward*. If the choice of ordering σ is clear from the context, we will call the arcs simply *backward* or *forward*.

Cutwidth. Let $T = (V, A)$ be a tournament and σ be an ordering of V. For $\alpha, \beta \in \{0, 1, \ldots, |V|\}$, $\alpha \leqslant \beta$, we define

$$\sigma(\alpha, \beta] := \{v \in V \mid \alpha < \sigma(v) \leqslant \beta\}.$$

Sets $\sigma(\alpha, \beta]$ as defined above will be called σ-*intervals*. The set

$$\mathsf{cut}_\sigma[\alpha] = \{(u, v) \in A \mid \sigma(u) > \alpha \geqslant \sigma(v)\} \subseteq \overleftarrow{A}_\sigma(T)$$

is called the α-*cut* of σ. The *width* of the ordering σ is equal to $\max_{0 \leqslant \alpha \leqslant |V|} |\mathsf{cut}[\alpha]|$, and the *cutwidth* of T, denoted $\mathbf{ctw}(T)$, is the minimum width among all orderings of V.

It is perhaps a bit surprising that in tournaments, there is a very simple algorithm to compute an ordering of optimum width: just sort the vertices by outdegrees.

▶ **Lemma 2** (see [2, 25]). *Let T be a tournament and σ be an ordering of T satisfying the following for every pair of vertices u and v: if u appears before v in σ, then the outdegre of u is not smaller than the outdegree of v. Then the width of σ is equal to $\mathbf{ctw}(T)$.*

If $I = \sigma(\alpha, \beta]$, then we denote

$$\partial^+(I) := \overrightarrow{A}(I, \sigma(0, \alpha]) \quad \text{and} \quad \partial^-(I) := \overrightarrow{A}(\sigma(\beta, |V|], I).$$

Note that $\partial^+(I) \subseteq \mathsf{cut}_\sigma[\alpha]$ and $\partial^-(I) \subseteq \mathsf{cut}_\sigma[\beta]$ and therefore $|\partial^+(I)| \leqslant c$ and $|\partial^-(I)| \leqslant c$, where c is the width of σ. These inclusions may be proper, as the arcs from the set $\widehat{\partial}(I) := \overrightarrow{A}(\sigma(\beta, |V|], \sigma(0, \alpha])$ contribute to the cuts but are not incident with I. We define $\partial(I) := \partial^+(I) \cup \partial^-(I)$ and call the elements of $\partial(I)$ I-*singular* (or simply *singular*) arcs. Moreover, we define

$$\Gamma^+(I) := \overrightarrow{A}(I, \sigma(\beta, |V|]), \qquad \Gamma^-(I) := \overrightarrow{A}(\sigma(0, \alpha], I), \quad \text{and} \quad \Gamma(I) = \Gamma^+(I) \cup \Gamma^-(I),$$

and call the elements of $\Gamma(I)$ I-*generic* (or simply *generic*) arcs.

If $I' = V - I$ where $I = \sigma(\alpha, \beta]$, then we call the set I' a *co-interval* and define I'-*singular* and I'-*generic* arcs as follows

$$\partial^-(I') := \partial^+(I), \quad \partial^+(I') := \partial^-(I), \quad \partial(I') := \partial^+(I') \cup \partial^-(I') = \partial(I),$$

$$\Gamma^-(I') := \Gamma^+(I), \quad \Gamma^+(I') := \Gamma^-(I), \quad \Gamma(I') := \Gamma^-(I') \cup \Gamma^+(I') = \Gamma(I).$$

Immersions. A digraph $\widehat{H}$ is an *immersion model* (or briefly a *copy*) of a digraph H if there exists a mapping ϕ, called an *immersion embedding*, such that:

- vertices of H are mapped to pairwise different vertices of $\widehat{H}$;
- each arc $(u, v) \in A(H)$ is mapped to a directed path in $\widehat{H}$ starting at $\phi(u)$ and ending at $\phi(v)$; and
- each arc of $\widehat{H}$ belongs to exactly one of the paths $\{\phi(a) \colon a \in A(H)\}$.

If it does not lead to misunderstanding, we will sometimes slightly abuse the above notation by identifying ϕ and $\widehat{H}$ and calling ϕ the immersion model of H.

Let H be a digraph. We say that a digraph G *contains H as an immersion* (or H can be *immersed* in G) if G has a subgraph that is an immersion model of H. Digraph G is called H-*immersion-free* (or H-*free* for brevity) if it does not contain H as an immersion.

We will use the following result of Fomin and Pilipczuk [14].

▶ **Theorem 3** (Theorem 7.3 of [14]). *Let T be a tournament which does not contain a digraph H as an immersion. Then $\mathbf{ctw}(T) \in \mathcal{O}((|H| + \|H\|)^2)$.*

▶ **Corollary 4.** *For every digraph H there exists a constant c_H such that for every H-free tournament T, we have $\mathbf{ctw}(T) \leqslant c_H$.*

Throughout this paper we fix a simple digraph H without isolated vertices and an integer $k \in \mathbb{N}$. For a tournament T, a set $F \subseteq A(T)$ is called a *solution* if $T - F$ is H-free. Moreover, F is an *optimal solution* if it is a solution of the smallest possible size. So (T, k) is a YES-instance of H-HITTING IMMERSIONS IN TOURNAMENTS if and only if there exists an optimal solution in T of size at most k.

Monoids and Simon factorization. Simon factorization was originally developed by Simon in [29] and the currently best bounds are due to Kufleitner [22]. See also the work of Bojańczyk [6] for a nice exposition; we mostly follow the notation from that source.

Let S be a finite monoid (i.e., a finite set equipped with an associative binary operation $\cdot$ and a neutral element 1). An element $e \in S$ is called *idempotent* if $e \cdot e = e$. For a finite alphabet A, by $A^\star$ we denote the set of all finite words over A, and a *morphism* $\alpha \colon A^\star \to S$ is a function satisfying $\alpha(\varepsilon) = 1$ (ε being the empty word) and $\alpha(w_1 w_2) = \alpha(w_1) \cdot \alpha(w_2)$ for every $w_1, w_2 \in A^\star$. Note that a morphism is uniquely defined by the images of single symbols from A. The following lemma is a direct consequence of Simon Factorization.

▶ **Lemma 5.** *Let S be a finite monoid, A be a finite alphabet, and $\alpha \colon A^\star \to S$ be a morphism. Suppose $w \in A^\star$ is a word of length at least $\ell^{3|S|}$. Then there exists a subword w' of w and an idempotent $e \in S$ such that $w' = w_1 w_2 \ldots w_\ell$, where $w_i \in A^\star$ are nonempty subwords of w and*

$$\alpha(w_1) = \alpha(w_2) = \ldots = \alpha(w_\ell) = e.$$

Note that in the setting of Lemma 5, given a word $w \in A^\star$ of length $n \geqslant \ell^{3|S|}$, one can easily find w' and a suitable decomposition $w' = w_1 w_2 \ldots w_\ell$ in time $\mathcal{O}(|S| \cdot n^3)$ assuming unit cost of operations in S. Indeed, one can guess e (by trying at most $|S|$ possibilities) and the first position of w' within w (by trying n possibilities), and then for every subword w'' starting at this position compute the longest possible decomposition of the form $w'' = w_1 w_2 \ldots w_{\ell'}$ such that $\alpha(w_1) = \ldots = \alpha(w_{\ell'}) = e$, if existent. The latter can be done by a standard left-to-right dynamic programming in time $\mathcal{O}(n^2)$.

3 Partial immersions

Our goal in this section is to extend the notion of an immersion to *partial immersions*. These will be used to understand possible behaviors of immersion models in a tournament T with respect to different intervals in an ordering of the vertex set of T. Let then $T = (V, A)$ be a tournament and let σ be an ordering of V. For now, fix a σ-interval $I := \sigma(\alpha, \beta]$.

Partial immersions.

▶ **Definition 6.** *A* scattered path *in I of size $q \geqslant 0$ is a sequence $\widetilde{P} = (P_i)_{i=1}^q$ satisfying the following properties:*
- *for each $i \in [q]$, P_i is a directed (simple) path of length at least 1 consisting of arcs that belong to $A(T[I]) \cup \partial(I) \cup \Gamma(I)$;*
- *paths P_i for $i \in [q]$ are pairwise arc-disjoint;*
- *for every $i \in [q]$, $i \neq 1$, we have $\mathsf{first}(P_i) \in \Gamma^-(I) \cup \partial^-(I)$;*
- *for every $i \in [q]$, $i \neq q$, we have $\mathsf{last}(P_i) \in \Gamma^+(I) \cup \partial^+(I)$.*

Each term in the sequence $(P_i)_{i=1}^q$ will be called a piece *of $\widetilde{P}$. The set of arcs of all pieces of $\widetilde{P}$ is denoted $A(\widetilde{P})$. If $\mathsf{first}(P_i) \in \Gamma^-(I)$ and $\mathsf{last}(P_i) \in \Gamma^+$, then the piece P_i is called* generic.

Note that $\mathsf{first}(P_1)$ is allowed to be an arc in the set $A(T[I]) \cup \Gamma^+(I) \cup \partial^+(I)$. If it is such, we call the vertex $\mathsf{tail}(\mathsf{first}(P_1)) \in I$ the *beginning* of $\widetilde{P}$ and denote it by $\mathsf{start}(\widetilde{P})$. Similarly, if $\mathsf{last}(P_q) \in A(T[I]) \cup \Gamma^-(I) \cup \partial^-(I)$, then we call the vertex $\mathsf{head}(\mathsf{last}(P_q))$ the *end* of $\widetilde{P}$ and denote it by $\mathsf{end}(\widetilde{P})$. Note that the empty sequence is a scattered path of size 0. Also, a scattered path with only one piece, whose both beginning and end exist, is just a path in $T[I]$. By $\mathcal{P}_I$ we denote the family of all scattered paths in I.

We say that a scattered path $\widetilde{P} = (P_i)_{i=1}^q$ in I can be *shortened* to a scattered path $\widetilde{P}'$ (or that $\widetilde{P}'$ is a *shortening* of $\widetilde{P}$) if:

- $\mathsf{start}(\widetilde{P}) = \mathsf{start}(\widetilde{P}')$ and $\mathsf{end}(\widetilde{P}) = \mathsf{end}(\widetilde{P}')$ (meaning either equal or simultaneously undefined);
- for each piece P_s of $\widetilde{P}'$ there exist indices $i_s^- \leqslant i_s^+$ such that $\mathsf{tail}(\mathsf{first}(P_s)) = \mathsf{tail}(\mathsf{first}(P_{i_s^-}))$ and $\mathsf{head}(\mathsf{last}(P_s)) = \mathsf{head}(\mathsf{last}(P_{i_s^+}))$; and
- whenever $s < s'$, we have $i_s^+ < i_{s'}^-$ and P_s appears before $P_{s'}$ in $\widetilde{P}'$.

Intuitively, shortening of the path means removing some pieces and replacing several contiguous subsequences of the pieces with single pieces, keeping the tail of the beginning and the head of the end of the replaced subsequence. Note that in particular, some pieces of $\widetilde{P}$ can be simply omitted in $\widetilde{P}'$ (other than the initial and the terminal one).

▶ **Definition 7.** *A partial immersion embedding of H in I (or shortly, a partial immersion in I) is a mapping $\phi \colon A(H) \to \mathcal{P}_I$ such that*

- *all scattered paths $\phi(a)$ for $a \in A(H)$ are pairwise arc-disjoint;*
- *if $\mathsf{tail}(a) = \mathsf{tail}(a')$ then $\mathsf{start}(\phi(a))$ and $\mathsf{start}(\phi(a'))$ are either equal, or simultaneously undefined;*
- *if $\mathsf{start}(\phi(a))$ and $\mathsf{start}(\phi(a'))$ are defined and equal, then $\mathsf{tail}(a) = \mathsf{tail}(a')$;*
- *if $\mathsf{head}(a) = \mathsf{head}(a')$ then $\mathsf{end}(\phi(a))$ and $\mathsf{end}(\phi(a'))$ are either equal, or simultaneously undefined;*
- *if $\mathsf{end}(\phi(a))$ and $\mathsf{end}(\phi(a'))$ are defined and equal, then $\mathsf{head}(a) = \mathsf{head}(a')$.*

Intuitively, we can think of a partial immersion as of a "trace" which some immersion model $\widehat{H}$ of H in T leaves on the interval I. Some edges of H have images being paths in $\widehat{H}$ non-incident with I (these correspond to empty scattered paths in the partial immersion embedding). Some images of arcs of H come back and forth to I, intersecting with I along a non-empty scattered path (the ordering of paths on a single scattered path corresponds to the order of their appearance along the image of the respective arc of H). Finally, some arc images begin or end within I, which corresponds to the case when the beginning or the end of a scattered path is defined and is a vertex of I.

We call a partial immersion ϕ' in I a *shortening* of ϕ in I if for every $a \in A(H)$, the scattered path $\phi'(a)$ is a shortening of $\phi(a)$. We call ϕ *minimal* if there is no shortening of ϕ with at least one scattered path of strictly smaller size. Note that ϕ may be minimal even if some piece of some $\phi(a)$ can be replaced by a different single piece with equal first and last vertices. Shortening which does not decrease the size of any scattered path will be called *trivial*.

Note that each immersion model ϕ of H in T is a partial immersion in $V(T)$, in which all scattered paths $\phi(a)$, $a \in A(H)$, are paths in T beginning and ending at $\phi(\mathsf{tail}(a))$ and $\phi(\mathsf{head}(a))$, respectively. Moreover, each partial immersion ϕ in I gives rise to a natural partial immersion of H in $J \subseteq I$ in which all paths $\phi(a)$ where $a \in A(H)$ are "trimmed" to scattered paths consisting of precisely those arcs which are incident with J.

Formally, let I and J be σ-intervals such that $J \subseteq I$. If P is a path in I, then define the *trace* $P|_J$ of P on J to be the scattered path consisting of all arcs of P incident with J, arranged in the order of appearance along P. If $\widetilde{P} = (P_i)_{i=1}^q$ is a scattered path in I, then

define the *trace* $\widetilde{P}|_J$ of $\widetilde{P}$ on J to be the concatenation of scattered paths $P_i|_J$. The *trace* $\phi|_J$ of a partial immersion ϕ of H in I on J is defined by setting $\phi|_J(a) = (\phi(a))|_J$ for every $a \in A(H)$.

Consider any σ-intervals I_1, I_2 with the property that there exist $\alpha, \beta, \gamma \in \{0, 1, \ldots, |V|\}$ such that $I_1 = \sigma(\alpha, \gamma]$ and $I_2 = \sigma(\gamma, \beta]$. We will call such a pair of intervals *consecutive*. Equivalently, two intervals are consecutive if their union $I = I_1 \cup I_2$ is a σ-interval as well. Let $\widetilde{P}_1$ be a scattered path in I_1 and $\widetilde{P}_2$ be a scattered path in I_2. We say that $\widetilde{P}_1$ and $\widetilde{P}_2$ are *compatible* if

- $A(\widetilde{P}_1) \cap A(I_1, I_2) = A(\widetilde{P}_2) \cap A(I_1, I_2)$;
- the set of pieces whose arc set is $A(\widetilde{P}_1) \cup A(\widetilde{P}_2)$ can be ordered to form a scattered path $\widetilde{P}$ in I with the property that all pieces of $\widetilde{P}_i$ appear in $\widetilde{P}$ in the same order as they do in $\widetilde{P}_i$, for $i = 1, 2$.

Every $\widetilde{P}$ described above will be called a *gluing* of $\widetilde{P}_1$ and $\widetilde{P}_2$. Note that a gluing is not necessarily uniquely defined, which can be seen particularly well when $A(\widetilde{P}_1) \cap A(I_1, I_2) = A(\widetilde{P}_2) \cap A(I_1, I_2) = \emptyset$ – in this case the pieces of both paths can be "shuffled" in any way only keeping the order of pieces originating from the same path.

▶ **Observation 8.** *If $\widetilde{P}_1$ is compatible with $\widetilde{P}_2$ and $\widetilde{P}_2'$ is a shortening of $\widetilde{P}_2$, then there exists a shortening $\widetilde{P}_1'$ of $\widetilde{P}_1$ such that $\widetilde{P}_1'$ and $\widetilde{P}_2'$ are compatible and every gluing of them is a shortening of some gluing of $\widetilde{P}_1$ and $\widetilde{P}_2$.*

Proof. To construct $\widetilde{P}_1'$ from $\widetilde{P}_1$ it is enough to omit the pieces which do not share arcs with $\widetilde{P}_2'$. ◀

We will say that two partial immersions ϕ_1 in I_1 and ϕ_2 in I_2 are *compatible* if there exists a partial immersion ϕ in I such that $\phi_1 = \phi|_{I_1}$ and $\phi_2 = \phi|_{I_2}$, or – in other words – that for every $a \in A(H)$ the scattered path $\phi(a)$ is a gluing of $\phi_1(a)$ and $\phi_2(a)$. We will call every such ϕ a *gluing* of ϕ_1 and ϕ_2. Note that gluing is not necessarily uniquely defined. Denote the set of all gluings of ϕ_1 and ϕ_2 by $\phi_1 \oplus \phi_2$.

▶ **Observation 9** (✂). *If $\phi \in \phi_1 \oplus \phi_2$ and ϕ is minimal, then ϕ_1 and ϕ_2 are minimal.*

The notions of a scattered path, partial immersion and trace can be naturally extended to co-intervals, by applying all definitions verbatim. If I is a σ-interval and $I' = V - I$ is the corresponding co-interval, then partial immersions ϕ_1 in I and ϕ_2 in I' are *compatible* if there exists an immersion ϕ in T such that $\phi_1 = \phi|_I$ and $\phi_2 = \phi|_{I'}$. Again, every such ϕ is called a *gluing* of ϕ_1 and ϕ_2.

Types of intervals. The key ingredient of our analysis is a constant-size encoding of the set of possible "behaviors" of partial immersions in intervals.

A σ-interval I shall be called ℓ-*long* if $|I| \geqslant \ell$. Further, we shall call I c-*flat* if $|\partial^+(I)| \leqslant c$, $|\partial^-(I)| \leqslant c$, and σ restricted to $T[I]$ has width at most c.

Note that if $I = \sigma(\alpha, \beta]$ is $2r$-long, then the intervals $I_r^- := \sigma(\alpha, \alpha+r]$ and $I_r^+ := \sigma(\beta-r, \beta]$ are disjoint. On the other hand, if I is c-flat, then we can color all backward arcs incident with I with at most $3c$ colors in such a way that each γ-cut of σ restricted to those arcs contains arcs of mutually different colors. This can be achieved e.g. by greedy coloring the γ-cuts for consecutive $\gamma = \alpha, \ldots, \beta + 1$. Formally, there exists a function $\xi \colon \overleftarrow{A}_\sigma(T) \cap A(I, V) \to [3c]$ such that for every $\gamma \in [|V| - 1]$ and every two distinct arcs $a_1, a_2 \in \mathsf{cut}_\sigma[\gamma] \cap A(I, V)$ we have $\xi(a_1) \neq \xi(a_2)$. In the following fix such a function.

▶ **Definition 10.** *Let ϕ be a partial immersion in an interval I that is $2r$-long and c-flat. For each $a \in A(H)$ we define the (r,c)-type $\tau^{(r,c)}(\phi(a))$ of the scattered path $\phi(a) = (P_i)_{i=1}^{q}$ as the following sequence of length $2q$:*

$$(f_-(\mathsf{first}(P_1)), f_+(\mathsf{last}(P_1)), f_-(\mathsf{first}(P_2)), f_+(\mathsf{last}(P_2)), \dots, f_-(\mathsf{first}(P_q)), f_+(\mathsf{last}(P_q))),$$

where the functions $f_\pm \colon A(T[I]) \cup \partial(I) \cup \Gamma(I) \to [-3c] \cup [r] \cup \{\mathsf{X}, \mathsf{H}\}$ are defined as follows

$$f_-(a) = \begin{cases} -\xi(a) & \text{if } a \in \partial^-(I), \\ \sigma(\mathsf{head}(a)) - \alpha & \text{if } a \in \Gamma^-(I) \text{ and } \mathsf{head}(a) \in I_r^-, \\ \mathsf{X} & \text{if } a \in \Gamma^-(I) \text{ and } \mathsf{head}(a) \notin I_r^-, \\ \mathsf{H} & \text{otherwise;} \end{cases}$$

$$f_+(a) = \begin{cases} -\xi(a) & \text{if } a \in \partial^+(I), \\ r + \sigma(\mathsf{tail}(a)) - \beta & \text{if } a \in \Gamma^+(I) \text{ and } \mathsf{tail}(a) \in I_r^+, \\ \mathsf{X} & \text{if } a \in \Gamma^+(I) \text{ and } \mathsf{tail}(a) \notin I_r^+, \\ \mathsf{H} & \text{otherwise.} \end{cases}$$

Let S be the set of all terms of the sequences $(\tau^{(r,c)}(\phi(a)))_{a \in A(H)}$, where by term we mean a value with an assigned position (i.e. equal values in different sequences are considered different terms). The (r,c)-type of ϕ is the collection of types $\tau^{(r,c)}(\phi) = (\tau^{(r,c)}(\phi(a)))_{a \in A(H)}$ equipped with a pair of equivalence relations (R_-, R_+) on the set $S \cup [c]$ defined as follows.

- *If a piece P_1 of $\phi(a_1)$ and a piece P_2 of $\phi(a_2)$ satisfy $\mathsf{head}(\mathsf{first}(P_1)) = \mathsf{head}(\mathsf{first}(P_2))$, then the corresponding terms in $\tau^{(r,c)}(\phi)$ (elements of the set $\{\mathsf{X}\} \cup [r]$ with assigned positions in respective sequences) are in relation R_-.*
- *If a piece P of $\phi(a)$ is such that $\mathsf{head}(\mathsf{first}(P))$ is the tail of a singular arc of color $x \in [3c]$, then the corresponding term in $\tau^{(r,c)}(\phi)$ is in relation R_- with x.*
- *Analogously, if $\mathsf{tail}(\mathsf{last}(P_1)) = \mathsf{tail}(\mathsf{last}(P_2))$, then the terms $f_+(\mathsf{last}(P_1))$ and $f_+(\mathsf{last}(P_2))$ are in relation R_+. And $x \in [3c]$ is in relation with all terms corresponding to tails of ends of paths which are simultaneously the head of the singular arc of color x.*

Let us provide some intuition on what kind of information is stored in the type of ϕ defined above. First of all, the entire "singular interface" of this partial immersion is kept, i.e. in the type we remember precisely the singular arcs used to enter or exit I when traversing along each scattered path $\phi(a)$. Moreover, if we enter or exit I with a generic arc, we remember the precise vertex of entry/exit inside I, but only if it is close enough to the "border" of I (i.e. within the first or last r vertices in σ). Otherwise we remember the respective entry/exit as "generic", which is marked by the marker X. Moreover, we keep the information about whether the scattered path begins or ends inside I – the marker H represents that a vertex of H is mapped under the partial immersion embedding to a vertex within I. Finally if the extreme (first or last) arcs of some pieces are generic and have the same first/last vertex in I, we remember this fact in the equivalence relations $R_\pm$. We shall need this information to be able to "glue" two partial immersions without using the same generic gluing arc for different pieces. The incidence with singular arcs is also stored to avoid a situation when one attempts a generic gluing along a singular arc. The reason for colors being stored as negative integers is purely technical – it ensures that $[r] \cap [-3c] = \emptyset$.

An (r,c)-*type* is any collection of sequences of even lengths over the set $[-3c] \cup [r] \cup \{\mathsf{X}, \mathsf{H}\}$ indexed by $A(H)$ and equipped with a pair of equivalence relations (R_-, R_+) on the union of the set of all terms of these sequences and $[3c]$. An I-*admissible (r,c)-type* is every type τ such that there exists a partial immersion ϕ in I such that $\tau = \tau^{(r,c)}(\phi)$. The *size* of a type is the sum of lengths of its sequences (i.e. it does not depend on the equivalence relations).

Let γ_r be a function mapping each element from the set $[r]$ onto the single marker X, and identity otherwise. Intuitively, the function γ_r keeps only the information about generic nature of the end of a piece, and "forgets" about the closeness of this vertex to the boundary. We say that an (r, c)-type τ' is a *shortening* of an (r, c)-type τ if for every $a \in A(H)$ the sequence $\gamma(\tau'_a)$ is a subsequence of $\gamma(\tau_a)$ and the two sequences have the same first and last terms.

The following observation is a direct consequence of the definition of shortening of a partial immersion.

▶ **Observation 11.** *If a partial immersion ϕ' of (r, c)-type τ' is a shortening of a partial immersion ϕ of (r, c)-type τ, then τ' is a shortening of τ.*

▶ **Definition 12.** *An I-admissible (r, c)-type τ is called* minimal *(in I) if there is no I-admissible type τ' of strictly smaller size that would be a shortening of τ.*

▶ **Observation 13.** *Let ϕ be a partial immersion in I of (r, c)-type τ. Then ϕ is minimal if and only if τ is minimal in I.*

Proof. If ϕ is not minimal, then for any nontrivial shortening ϕ' of ϕ, $\tau^{(r,c)}(\phi')$ is a shortening of τ of strictly smaller size. Conversely, if τ is not minimal, then there exists an I-admissible type τ' of strictly smaller size. Every partial immersion ϕ' of type τ' is a nontrivial shortening of ϕ. ◀

We now observe that in a minimal partial immersion on an interval I, all scattered paths will be relatively small, that is, will visit I only a bounded number of times.

▶ **Lemma 14** ($\asymp$). *Suppose that $F \subseteq A(T)$ has the property that $|F| \leqslant f$, I is $4\|H\|(c + f + 1)$-long and c-flat, and a partial immersion ϕ in I is minimal and disjoint with F. Then for every $a \in A(H)$ the scattered path $\phi(a)$ has size at most $2c + 3$.*

We call an (r, c)-type τ *short* if for every $a \in A(H)$ the length of $\tau(\phi(a))$ is at most $4c + 6$.

▶ **Corollary 15** ($\asymp$). *For $r \geqslant 1$ we have the following:*
1. *If an interval I is $4\|H\|(r + c)$-long and c-flat, and a partial immersion ϕ in I is minimal, then $\tau^{(r,c)}(\phi)$ is short.*
2. *For every pair of integers $r, c \in \mathbb{N}$ there is a constant $t(r, c)$ such that there are exactly $t(r, c)$ short (r, c)-types. In particular, for any interval I as above, there are at most $t(r, c)$ different I-admissible minimal (r, c)-types.*

Let I_1, I_2 be two consecutive c-flat, $2(r + c)$-long σ-intervals and $I = I_1 \cup I_2$. We say that two (r, c)-types τ_1 of I_1 and τ_2 of I_2 are *compatible* if there exists a partial immersion ϕ in I such that $\phi|_{I_1}$ and $\phi|_{I_2}$ are compatible, $\phi|_{I_1}$ has type τ_1, and $\phi|_{I_2}$ has type τ_2.

The *gluing* of the types $\tau_1 \oplus \tau_2$ is defined as the set of all (r, c)-types of all such ϕ. The following lemma proves that this definition is correct, i.e. that types of two immersions store enough information to ensure the possibility of gluing them.

▶ **Lemma 16** ($\asymp$). *Let I_1, I_2 be two consecutive c-flat, $4\|H\|(r + c)$-long σ-intervals and $I = I_1 \cup I_2$. If two (r, c)-types τ_1 of I_1 and τ_2 of I_2 are* compatible*, then for every partial immersion ϕ_1 in I_1 of type τ_1 and for every partial immersion ϕ_2 in I_2 of type τ_2, the immersions ϕ_1 and ϕ_2 are compatible.*

Boundaried intervals and signatures.   In the process of replacing protrusions we will need to consider intervals not as subsets of an ordered vertex set of a larger tournament, but as standalone structures which can be used to replace one another. We introduce the notion of an (r, c)-boundaried interval to enable such considerations.

▶ **Definition 17.** *An (r, c)-boundaried interval is a digraph D on vertex set $V(D) = S^+ \cup I \cup S^-$ equipped with an ordering σ_I of I. Furthermore, we require the following:*

- *$|I| \geqslant 4\|H\|(r + c)$;*
- *$D[I]$ is a tournament;*
- *σ_I has width at most c;*
- *$S^- \subseteq [3c] \times \{-\}$ and $S^+ \subseteq [3c] \times \{+\}$;*
- *each vertex of S^+ has only one incident arc and this arc belongs to the set $\overrightarrow{A}(I, S^+)$;*
- *each vertex of S^- has only one incident arc and this arc belongs to the set $\overrightarrow{A}(S^-, I)$.*

The r-boundary of D is the pair of sets I^- and I^+ consisting of the first and the last r vertices of I in σ_I, respectively. For D defined above, we will shortly write $D = I \cup S^\pm$.

Note that this notion emulates a $4\|H\|(r + c)$-long c-flat interval I in the following sense. Arcs whose heads are contained in S^+ correspond to $\partial^+(I)$, and arcs whose tails are contained in the set S^- correspond to $\partial^-(I)$. The names of the auxiliary vertices in $S^\pm$ correspond to the ξ-colors of the respective backward arcs. The r-boundary consists of precisely those vertices whose generic entry or exit is remembered in the (r, c)-type of a partial immersion.

Formally, every $4\|H\|(r + c)$-long c-flat σ-interval I in T can be uniquely encoded with an (r, c)-boundaried interval $D^{(r,c)}(I)$, whose structure resembles the structure of $T[I]$ and $\partial(I)$ as follows:

- the vertices and arcs of I are kept along with their ordering in T;
- every singular arc $a \in \partial^\pm(I)$ is mapped to an arc joining $(\xi(a), \pm) \in S^\pm$ with the endpoint of a contained in I;
- the projection of $S^\pm$ onto the first coordinate is precisely $\{\xi(a) \mid a \in \partial^\pm(I)\}$.

The notion of a partial immersion can be naturally adjusted to the setting of boundaried intervals. The only difference is the lack of the "generic interface" i.e. there are no auxiliary edges in boundaries intervals used to emulate $\Gamma(I)$. These can be, however, emulated by storing the information from the type (marker X or number in $[r]$ if the generic arc is incident with the r-boundary, and the equivalence relations $R_\pm$) instead of the identity of particular generic arcs. Formally, a piece of a scattered path in (r, c)-boundaried interval can begin or end with an element in $\{\mathsf{X}\} \cup [r]$ instead of a generic arc. In particular, this slightly modified variant of partial immersions can be equipped with precisely the same definition of admissible (r, c)-type as in the former case.

▶ **Definition 18.** *An (r, c)-signature is a subset of the set of all short (r, c)-types. The (r, c)-signature $\Sigma^{(r,c)}(I)$ of a $4\|H\|(r + c)$-long c-flat σ-interval I is the set of all I-admissible minimal (r, c)-types. The (r, c)-signature $\Sigma^{(r,c)}(D)$ of an (r, c)-boundaried interval $D = I \cup S^\pm$ is the set of all I-admissible minimal (r, c)-types.*

The intuition behind this definition is that if I is appropriately long and flat, then the signature of I stores the information about all possible interactions of I with minimal partial immersions. Note that $\Sigma^{(r,c)}(D^{(r,c)}(I)) = \Sigma^{(r,c)}(I)$.

We say that two (r, c)-boundaried intervals $I \cup S^\pm$ and $I' \cup S'^\pm$ are *exchangeable* if they have equal (r, c)-signatures, $S^\pm = S'^\pm$ (both intervals use precisely the same colors on the boundary), and the incidence structure of r-boundaries of those intervals with backward arcs

is the same, i.e. for every $i \in [r]$ the set of colors of singular arcs incident with both $S^{\mp}$ and the i-th vertex of $I^{\pm}$ is the same as analogously defined set of colors for i-th vertex of $I'^{\pm}$. Intuitively this means that in T we may replace the interval I with I'.

▶ **Corollary 19.** *For every pair of integers $r, c \in \mathbb{N}$ there is a constant $s(r, c)$ such that there are exactly $s(r, c)$ different (r, c)-signatures.*

Proof. We may set $s(r, c) = 2^{t(r,c)}$, where $t(r, c)$ is the constant provided by Corollary 15. ◀

For future discussion of algorithmic aspects, we will need the following observation.

▶ **Lemma 20** (⋉). *Consider r and c fixed and let T, σ, and I be as in Definition 18. Then given T, σ, and I, the signature $\Sigma^{(r,c)}(I)$ can be computed in polynomial time.*

Let $\mathcal{S}^{(r,c)}$ be the set of all (r, c)-signatures; we have $|\mathcal{S}^{(r,c)}| = s(r, c)$, where $s(r, c)$ is the constant given by Corollary 19. Let $\mathcal{S}_{\sigma}^{(r,c)}$ be the set of those (r, c)-signatures which are equal to $\Sigma^{(r,c)}(I)$ for some σ-interval I, i.e. contain only I-admissible minimal (r, c)-types. Then $\mathcal{S}_{\sigma}^{(r,c)} \subseteq S^{(r,c)}$, so $|\mathcal{S}_{\sigma}^{(r,c)}| \leqslant s(r, c)$. We argue that $\mathcal{S}_{\sigma}^{r,c}$ has a structure of a semigroup in the following sense.

▶ **Lemma 21** (⋉). *Let I_1, I_2 be two $4\|H\|(r + c)$-long c-flat σ-intervals such that $I = I_1 \cup I_2$ is a c-flat σ-interval. Then $\Sigma^{(r,c)}(I)$ is uniquely determined by $\Sigma^{(r,c)}(I_1)$ and $\Sigma^{(r,c)}(I_2)$.*

Lemma 21 implies that the set $\mathcal{S}_{\sigma}^{(r,c)} \cup \{0\}$ can be endowed with an associative binary product operation such that for every two consecutive intervals I_1, I_2, the product of their signatures is the signature of their union $I_1 \cup I_2$. Formally, we set the product for all pairs of consecutive intervals as above; Lemma 21 enures that this is well-defined. Next, for all pairs of elements $\tau_1, \tau_2 \in \mathcal{S}_{\sigma}^{r,c}$ for which their product is not yet defined, we set $\tau_1 \cdot \tau_2 = 0$. Also, we set $0 = 0 \cdot 0 = 0 \cdot \tau = \tau \cdot 0$ for all $\tau \in \mathcal{S}_{\sigma}^{(r,c)}$. In this way, $\mathcal{S}_{\sigma}^{(r,c)} \cup \{0\}$ becomes a monoid; the empty signature is the neutral element of multiplication.

By Lemma 5 we obtain the following.

▶ **Corollary 22** (⋉). *Suppose I is a c-flat $4\|H\|(r+c)\ell^{3s(r,c)}$-long σ-interval. Then there exists a sequence of consecutive $4\|H\|(r + c)$-long c-flat σ-intervals $(I_i)_{i=1}^{\ell}$ whose (r, c)-signatures are equal and equal to the signature of their union. Moreover, given r, c, T, σ, and I, such a sequence can be found in polynomial time.*

4   Finding protrusions

In order to find an appropriately large subgraph of T which does not "affect" the behavior of T with respect to H-HITTING IMMERSIONS IN TOURNAMENTS, we roughly proceed as follows. First, we find a suitable ordering σ of $V(T)$ and an appropriately long interval X in σ such that X has a constant-size singular interface towards the remainder of T. Then, inside X, we find (again, an appropriately long) subinterval I of a very specific structure: I can be divided into $2k + 3$ subintervals with the same signatures as itself. This is where we use Simon Factorization through Corollary 22. In the next part we use this extra structure to prove that one of these subintervals can be replaced with a strictly smaller replacement in such a way that after the substitution, we obtain an equivalent instance of the problem.

We proceed to a formal implementation of this plan. The first lemma gives the ordering σ and the interval X.

▶ **Lemma 23** (✂). *Let T be a tournament with $\mathbf{ctw}(T) \leqslant c$ and $|T| \geqslant (2c+1)(x+1)(k+1)$.*
If T contains at most k arc-disjoint immersion copies of H, then there exists an ordering σ
of $V(T)$ and an H-free σ-interval X such that $|X| \geqslant x$ and X is c_H-flat with respect to σ.

Moreover, given T, k, c, x as above, one can in polynomial time either conclude that T
contains more than k arc-disjoint immersion copies of H, or find an ordering σ and an
interval X satisfying the above properties.

In the remainder of this section let $x \geqslant 4\|H\|$, r, c be fixed positive integers. Moreover,
let T be a tournament for which there exists an optimal solution of size at most k and let
σ be an ordering of T. Finally let X be an H-free c-flat $x(3c + k + 1)(2k + 3)^{3s(r,c)}$-long
σ-interval. We assume that there is a coloring ξ mapping all σ-backward arcs incident to X
to colors in $[3c]$ such that not two such arcs of the same color participate in the same σ-cut.

We may apply Corollary 22 to X to find a collection of $2k + 3$ consecutive σ-intervals I_i
with $|I_i| = x(3c + k + 1)$ for all $i \in [2k + 3]$ such that all I_i have equal (r, c)-signatures, and
this common signature, call it Σ, is equal to the (r, c)-signature of their union I. Then I is an
H-free c-flat $x(3c + k + 1)(2k + 3)$-long σ-interval. That such I can be found in polynomial
time (given r, c, x, T, X, σ, and ξ) follows from Corollary 22.

Now comes a key step in the proof: we argue that from the equality of types of
$I, I_1, \ldots, I_{2k+3}$ it follows that every optimum solution will contain a bounded number of arcs
incident with I.

▶ **Lemma 24** (✂). *For every optimal solution $F \subseteq A(T)$, we have $|F \cap A(I, V(T))| \leqslant 2c$.*

We define a digraph T° based on T and $I = I_1 \cup \ldots \cup I_{2k+3}$ as follows: start with T, and
- remove all vertices of I_2;
- for every arc $a \in \partial^+(I_2)$, replace a with an arc with the same head as a and tail in a
  fresh vertex $s^+_{\xi(a)}$;
- for every arc $a \in \partial^-(I_2)$, replace a with an arc with the same tail as a and head in a
  fresh vertex $s^-_{\xi(a)}$.

We call the constructed graph a *c-boundaried co-interval*. The intuition of this construction
is as follows. We pinch off one of the $2k + 3$ intervals and keep the singular arc interface in a
fashion similar as in (r, c)-boundaried intervals. The only difference is that we do not keep
track of the r-boundary vertices.

Now we can define the *gluing* of T° with an (r, c)-boundaried interval $B = I \cup S^\pm$ with
signature Σ, simply by identifying the singular arcs of the same color (note that different
vertices from $S^\pm$ can be therefore mapped to the same vertex) and completing the obtained
structure to a tournament by making all missing arcs generic. Note that in order for this to
be well-defined, we need to require that the sets of colors of the singular arcs in T° and in
B are identical – if it is so, we will say that T° and B are *compatible*. Also, note that we
require that the signature of B is Σ: that is, the possible types of partial immersions present
in B are exactly the same as in the substituted interval I_2.

Denote by $T^\circ \oplus B$ the tournament obtained from gluing T° and B. Note that in this
tournament we have naturally defined ordering of vertices: in T° it is inherited from the
ordering σ of T and within B it is inherited from the ordering σ_I of the boundaried interval.
Finally all the vertices of the substituted interval appear in the ordering between the two
interval parts (prefix and suffix) of the co-interval. The following observation is obvious.

▶ **Observation 25.** *If two exchangable (r, c)-boundaried intervals $B = I \cup S^\pm$ and $B' = I' \cup S^\pm$*
are compatible with T°, then the (r, c)-signatures of I in $T^\circ \oplus B$ and I' in $T^\circ \oplus B'$ are equal.

We now observe, by inspecting the proof of Lemma 24, that if in I we replace one of its subintervals with an exchangeable interval, then the conclusion of Lemma 24 – that the modified I' will still have constant incidence with every optimal solution in T – will still hold. Let us summarize this section with putting together these observations and recalling all needed assumptions.

▶ **Corollary 26** (⋎). *Suppose that $x \geqslant 4\|H\|$, r, c are fixed positive integers, T is a tournament of for which there exists an optimal solution of size at most k and σ is an ordering of T. Suppose further that X is an H-free c-flat $x(3c + k + 1)(2k + 3)^{3s(r,c)}$-long σ-interval with all incident σ-backward arcs colored according to ξ with colors $[3c]$ so that no two arcs of the same color participate in the same σ-cut.*

Then there exists an (r, c)-signature Σ, a c-boundaried co-interval T° of this signature, and an (r, c)-boundaried interval $B = I \cup S^\pm$ such that $T = T^\circ \oplus B$ and moreover for every $B' = I' \cup S'^\pm$ exchangeable with B and for every optimal solution $F \subseteq A(T')$ where $T' = T^\circ \oplus B'$ we have: $|F \cap A(I', T')| \leqslant 2c$. Moreover, given x, r, c, T, σ, and X, such Σ, T°, and B can be computed in polynomial time.

5 Replacing protrusions

In the entire section we fix $c := c_H$, where c_H is the constant from Corollary 4, and $r := 6\|H\|c$. Moreover we assume that $x \geqslant 4\|H\|$; the precise value of x will be determined later.

▶ **Definition 27.** *Protrusion is any (r, c)-boundaried interval $X = I \cup S^\pm$ which is H-free and $x(3c + k + 1)$-long. For brevity we will refer to X as to I. The set $\Sigma^{(r,c)}(X)$ is the signature of the protrusion.*

Let T be a tournament equipped with a vertex ordering σ and let I be an H-free interval and such that $T = T^\circ \oplus X$, where X is a protrusion of signature Σ. Let $I \subseteq V(T)$ be the interval defined by the protrusion. Fix the coloring ξ of σ-backward arcs incident with I.

Recall that from Corollary 26 follows that for every (r, c)-boundaried co-interval T° compatible with X if $T^\circ \oplus X$ admits an optimal solution F of size not greater than k, then there are at most $2c$ arcs in F incident with X.

For every protrusion $X = I \cup S^\pm$ define a function $f_X : 2^{\mathcal{S}^{(r,c)}} \to \{0, 1, 2, \ldots, 2c\} \cup \{\infty\}$ as follows: $f_X(S)$ is the minimum number of arcs in $A(I, I)$ needed to hit all partial immersions in I whose signatures belong to S, or ∞ if this number is greater than $2c$.

We introduce an equivalence relation $\sim$ on the set of all protrusions. Let $X = I \cup S^\pm$, $X' = I' \cup S'^\pm$. We say that $X \sim X'$ if:
- $S^\pm = S'^\pm$;
- $\Sigma^{(r,c)}(X) = \Sigma^{(r,c)}(X')$; and
- $f_X(S) = f_{X'}(S)$ for every $S \subseteq \mathcal{S}^{(r,c)}$.

Note that the number of equivalence classes of $\sim$ is finite and bounded by a constant depending only on H. This means that if in each class we pick a representative with the minimal number of vertices (call each such element a *small* protrusion), then all small protrusions will have size bounded from above uniformly by a constant s_H depending on the digraph H only.

Same arguments as in the proof of Lemma 20 give the following.

▶ **Lemma 28.** *Given protrusions X and X' it can be decided in polynomial time whether $X \sim X'$.*

We now argue that equivalent protrusions are replaceable.

▶ **Lemma 29** ($\bowtie$). *Suppose that $T = T^\circ \oplus X$ is a tournament satisfying the conclusion of Corollary 26, where X is a boundaried (r, c)-interval of signature Σ and T° is a c-boundaried co-interval which is compatible with X.*

Then for every X' such that $X \sim X'$, the optimal solution in $T^\circ \oplus X$ is of size not greater than k if and only if the optimal solution in $T^\circ \oplus X'$ is of size not greater than k.

We are now set up with all tools needed to prove our main theorem.

Proof of Theorem 1. We prove that provided $|T| > C \cdot k^C$, where the constant C will be defined later, one can compute an instance (T', k) equivalent to (T, k) and satisfying $|T'| < |T|$. The conclusion will follow from applying such reduction (at most) $|T|$ times.

First of all note that from Theorem 3 for a digraph being a disjoint union of k exemplars of H, we conclude that if T does not contain k arc-disjoint immersion copies of H, then $\mathbf{ctw}(T) \leqslant c_0 k^2$ for some constant c_0 (depending on H). In particular, it follows that if $\mathbf{ctw}(T) > k^2 c_0$ (which can be established in polynomial time using Lemma 2), then T is a NO-instance. So from now on we may assume that $\mathbf{ctw}(T) \leqslant c_0 k^2$.

Let $c = c_H$, $r = 6\|H\|c$, $x' = \max\{4\|H\|, s_H + 1\}$ and $x = x'(3c + k + 1)(2k + 3)^{3s(r,c)}$, where $s(r, c)$ is defined in Corollary 19. Let C be a constant satisfying $Ck^C \geqslant (2k^2 c_0 + 1)(x + 1)(k + 1)$, e.g. $C = \max\{3s(r, c) + 4, 5^{3s(r,c)} \cdot 3c_0 \cdot 4x' \cdot (6c + 2)\}$.

Suppose that T is a tournament satisfying $\mathbf{ctw}(T) \leqslant k^2 c_0$ and $|T| > Ck^C$. Applying Lemma 23, we either conclude that T admits more than k arc-disjoint copies of H (so (T, k) is a NO-instance), or find an ordering σ of $V(T)$ and an H-free c-flat x-long σ-interval J. Both conclusions can be effectively gained in polynomial time.

In the latter case, we may use Corollary 22 in a manner described in Section 4 to find in J an H-free c-flat $x'(3c + k + 1)(2k + 3)$-long σ-interval I of (r, c)-signature Σ, which can be decomposed to $2k + 3$ consecutive $x'(3c + k + 1)$-long σ-intervals I_i, each of (r, c)-signature Σ. Both I and Σ are found in polynomial time.

Let $T = T^\circ \oplus X$ be the decomposition where T° is a c-boundaried co-interval and $X = (I_2 \cup S^\pm, \Sigma)$ is a protrusion corresponding to I_2. Clearly, X is compatible with T°. Using Lemma 28 we may check in polynomial time all small protrusions to find one X' such that $X' \sim X$. Let us define $T' = T^\circ \oplus X'$.

By Lemma 29 we conclude that (T', k) is an instance of H-HITTING IMMERSIONS IN TOURNAMENTS equivalent to (T, k). Moreover as $|X'| \leqslant s_H < |X|$, we have that $|T'| < |T|$. ◀

───── **References** ─────

1   Noga Alon, Daniel Lokshtanov, and Saket Saurabh. Fast FAST. In *36th International Colloquium on Automata, Languages and Programming, ICALP 2009*, volume 5555 of *Lecture Notes in Computer Science*, pages 49–58. Springer, 2009. `doi:10.1007/978-3-642-02927-1_6`.

2   Florian Barbero, Christophe Paul, and Michał Pilipczuk. Exploring the complexity of layout parameters in tournaments and semicomplete digraphs. *ACM Trans. Algorithms*, 14(3):38:1–38:31, 2018. `doi:10.1145/3196276`.

3   Stéphane Bessy, Marin Bougeret, R. Krithika, Abhishek Sahu, Saket Saurabh, Jocelyn Thiebaut, and Meirav Zehavi. Packing arc-disjoint cycles in tournaments. *Algorithmica*, 83(5):1393–1420, 2021. `doi:10.1007/s00453-020-00788-2`.

4   Stéphane Bessy, Fedor V. Fomin, Serge Gaspers, Christophe Paul, Anthony Perez, Saket Saurabh, and Stéphan Thomassé. Kernels for feedback arc set in tournaments. *Journal of Computer and System Sciences*, 77(6):1071–1078, 2011. `doi:10.1016/j.jcss.2010.10.001`.

5   Hans L. Bodlaender, Fedor V. Fomin, Daniel Lokshtanov, Eelko Penninkx, Saket Saurabh, and Dimitrios M. Thilikos. (Meta) Kernelization. *J. ACM*, 63(5):44:1–44:69, 2016. `doi:10.1145/2973749`.

6   Mikolaj Bojanczyk. Factorization forests. In *13th International Conference on Developments in Language Theory, DLT 2009*, volume 5583 of *Lecture Notes in Computer Science*, pages 1–17. Springer, 2009. `doi:10.1007/978-3-642-02737-6_1`.

7   Marthe Bonamy and Michał Pilipczuk. Graphs of bounded cliquewidth are polynomially χ-bounded. *Advances in Combinatorics*, 2020:8, 2020.

8   Łukasz Bożyk and Michał Pilipczuk. On the Erdős-Pósa property for immersions and topological minors in tournaments. *Discrete Mathematics & Theoretical Computer Science*, vol. 24, no. 1, April 2022. `doi:10.46298/dmtcs.7099`.

9   Maria Chudnovsky, Alexandra Ovetsky Fradkin, and Paul D. Seymour. Tournament immersion and cutwidth. *Journal of Combinatorial Theory, Series B*, 102(1):93–101, 2012. `doi:10.1016/j.jctb.2011.05.001`.

10   Maria Chudnovsky and Paul D. Seymour. A well-quasi-order for tournaments. *Journal of Combinatorial Theory, Series B*, 101(1):47–53, 2011. `doi:10.1016/j.jctb.2010.10.003`.

11   Uriel Feige. Faster FAST (Feedback Arc Set in Tournaments). *CoRR*, abs/0911.5094, 2009. `arXiv:0911.5094`.

12   Fedor V. Fomin, Daniel Lokshtanov, Neeldhara Misra, and Saket Saurabh. Planar $\mathcal{F}$-Deletion: Approximation, kernelization and optimal FPT algorithms. In *53rd Annual IEEE Symposium on Foundations of Computer Science, FOCS 2012*, pages 470–479. IEEE Computer Society, 2012. `doi:10.1109/FOCS.2012.62`.

13   Fedor V. Fomin, Daniel Lokshtanov, Saket Saurabh, and Meirav Zehavi. *Kernelization: Theory of Parameterized Preprocessing*. Cambridge University Press, 2019.

14   Fedor V. Fomin and Michał Pilipczuk. On width measures and topological problems on semi-complete digraphs. *Journal of Combinatorial Theory, Series B*, 138:78–165, 2019. `doi:10.1016/j.jctb.2019.01.006`.

15   Archontia C. Giannopoulou, Bart M. P. Jansen, Daniel Lokshtanov, and Saket Saurabh. Uniform kernelization complexity of hitting forbidden minors. *ACM Transactions on Algorithms*, 13(3):35:1–35:35, 2017. `doi:10.1145/3029051`.

16   Archontia C. Giannopoulou, Ken-ichi Kawarabayashi, Stephan Kreutzer, and O-joung Kwon. The directed flat wall theorem. In *2020 ACM-SIAM Symposium on Discrete Algorithms, SODA 2020*, pages 239–258. SIAM, 2020. `doi:10.1137/1.9781611975994.15`.

17   Archontia C. Giannopoulou, Ken-ichi Kawarabayashi, Stephan Kreutzer, and O-joung Kwon. Directed tangle tree-decompositions and applications. In *2022 ACM-SIAM Symposium on Discrete Algorithms, SODA 2022*, pages 377–405. SIAM, 2022. `doi:10.1137/1.9781611977073.19`.

18   Archontia C. Giannopoulou, Michał Pilipczuk, Jean-Florent Raymond, Dimitrios M. Thilikos, and Marcin Wrochna. Linear kernels for edge deletion problems to immersion-closed graph classes. *SIAM Journal on Discrete Mathematics*, 35(1):105–151, 2021. `doi:10.1137/18M1228839`.

19   Marek Karpinski and Warren Schudy. Faster algorithms for Feedback Arc Set Tournament, Kemeny Rank Aggregation and Betweenness Tournament. In *21st International Symposium on Algorithms and Computation, ISAAC 2010*, volume 6506 of *Lecture Notes in Computer Science*, pages 3–14. Springer, 2010. `doi:10.1007/978-3-642-17517-6_3`.

20   Ken-ichi Kawarabayashi and Stephan Kreutzer. The directed grid theorem. In *47th Annual ACM on Symposium on Theory of Computing, STOC 2015*, pages 655–664. ACM, 2015. `doi:10.1145/2746539.2746586`.

21   Ilhee Kim and Paul D. Seymour. Tournament minors. *Journal of Combinatorial Theory, Series B*, 112:138–153, 2015. `doi:10.1016/j.jctb.2014.12.005`.

22   Manfred Kufleitner. The height of factorization forests. In *33rd International Symposium Mathematical Foundations of Computer Science 2008, MFCS 2008*, volume 5162 of *Lecture Notes in Computer Science*, pages 443–454. Springer, 2008. `doi:10.1007/978-3-540-85238-4_36`.

23 Jaroslav Nešetřil, Patrice Ossona de Mendez, Michał Pilipczuk, Roman Rabinovich, and Sebastian Siebertz. Rankwidth meets stability. In *2021 ACM-SIAM Symposium on Discrete Algorithms, SODA 2021*, pages 2014–2033. SIAM, 2021. `doi:10.1137/1.9781611976465.120`.

24 Jaroslav Nešetřil, Roman Rabinovich, Patrice Ossona de Mendez, and Sebastian Siebertz. Linear rankwidth meets stability. In *2020 ACM-SIAM Symposium on Discrete Algorithms, SODA 2020*, pages 1180–1199. SIAM, 2020. `doi:10.1137/1.9781611975994.72`.

25 Alexandra Ovetsky Fradkin. *Forbidden structures and algorithms in graphs and digraphs*. PhD thesis, Princeton University, 2011.

26 Alexandra Ovetsky Fradkin and Paul D. Seymour. Tournament pathwidth and topological containment. *Journal of Combinatorial Theory, Series B*, 103(3):374–384, 2013. `doi:10.1016/j.jctb.2013.03.001`.

27 Alexandra Ovetsky Fradkin and Paul D. Seymour. Edge-disjoint paths in digraphs with bounded independence number. *Journal of Combinatorial Theory, Series B*, 110:19–46, 2015. `doi:10.1016/j.jctb.2014.07.002`.

28 Jean-Florent Raymond. Hitting minors, subdivisions, and immersions in tournaments. *Discrete Mathematics and Theoretical Computer Science*, 20(1), 2018. URL: `http://dmtcs.episciences.org/4212`.

29 Imre Simon. Factorization forests of finite height. *Theoretical Computer Science*, 72(1):65–94, 1990. `doi:10.1016/0304-3975(90)90047-L`.