Trees with many leaves in tournaments

Alistair Benford* and Richard Montgomery†

Abstract

Sumner's universal tournament conjecture states that every $(2n-2)$-vertex tournament should contain a copy of every n-vertex oriented tree. If we know the number of leaves of an oriented tree, or its maximum degree, can we guarantee a copy of the tree with fewer vertices in the tournament? Due to work initiated by Häggkvist and Thomason (for number of leaves) and Kühn, Mycroft and Osthus (for maximum degree), it is known that improvements can be made over Sumner's conjecture in some cases, and indeed sometimes an $(n+o(n))$-vertex tournament may be sufficient.

In this paper, we give new results on these problems. Specifically, we show

i) for every $\alpha > 0$, there exists $n_0 \in \mathbb{N}$ such that, whenever $n \geqslant n_0$, every $((1+\alpha)n+k)$-vertex tournament contains a copy of every n-vertex oriented tree with k leaves, and

ii) for every $\alpha > 0$, there exists $c > 0$ and $n_0 \in \mathbb{N}$ such that, whenever $n \geqslant n_0$, every $(1+\alpha)n$-vertex tournament contains a copy of every n-vertex oriented tree with maximum degree $\Delta(T) \leqslant cn$.

Our first result gives an asymptotic form of a conjecture by Havet and Thomassé, while the second improves a result of Mycroft and Naia which applies to trees with polylogarithmic maximum degree.

1 Introduction

When the edges of a complete graph are oriented in any manner, giving a tournament, which oriented trees must appear within its edges? The study of this question has been motivated by Sumner's universal tournament conjecture from 1971, which states that every $(2n-2)$-vertex tournament should contain a copy of every n-vertex oriented tree (see, e.g., [15]). The extremal examples showing this conjecture would be tight, the n-vertex stars whose root vertex has in- or out-degree 0, also maximise the number of leaves and the maximum degree among the n-vertex oriented trees. In this paper, we consider whether fewer vertices are required in the tournament for trees with fewer leaves or a lower maximum degree.

The first major step towards Sumner's conjecture was taken by Häggkvist and Thomason [6] in 1991, who showed that $O(n)$ vertices in a tournament are sufficient to find a copy of any n-vertex oriented tree. The constant implicit in this result has been improved in the intervening years (see [9, 7, 4]), with the best current bound applicable for all n by Dross and Havet [3], who showed that any $\lceil \frac{21}{8}n - \frac{47}{16} \rceil$-vertex tournament contains a copy of any n-vertex oriented tree. Significantly, however, Sumner's conjecture has been proved exactly for all sufficiently large n, by Kühn, Mycroft and Osthus [13], so that the conjecture remains open for only finitely many oriented trees.

That trees with fewer leaves generally require fewer vertices in the tournament was also first demonstrated by Häggkvist and Thomason [6] in 1991. That is, they showed that there is some smallest $g(k)$ such that every $(n+g(k))$-vertex tournament contains a copy of every n-vertex oriented tree with k leaves. Due to Thomason [16], it is known that every $(n+1)$-vertex tournament contains a copy of every n-vertex oriented path, and so $g(2) = 1$ (noting that we must have $g(k) \geqslant k-1$ as a consequence of the previous example of an n-vertex star). Motivated in part by this result, Havet and Thomassé [8] generalised Sumner's conjecture, suggesting that $g(k) = k-1$, that is, that every $(n+k-1)$-vertex tournament should contain a copy of every n-vertex oriented tree with k leaves.

Improving the initial bound given by Häggkvist and Thomason [6], which was exponential in k^3, Dross and Havet [3] showed that $g(k) = O(k^2)$, before the current authors showed that $g(k) = O(k)$ [2]. That is, it

*School of Mathematics, University of Birmingham, Birmingham, B15 2TT, UK. AXB1433@bham.ac.uk

†Mathematics Institute, University of Warwick, Coventry, CV4 7AL, UK. richard.montgomery@warwick.ac.uk. Supported by the European Research Council (ERC) under the European Union Horizon 2020 research and innovation programme (grant agreement No. 947978) and the Leverhulme trust.

is now known that every $(n + O(k))$-vertex tournament contains a copy of every n-vertex oriented tree with k leaves. Further evidence towards the conjecture of Havet and Thomassé is that $(n + k - 1)$-vertices in the tournament are known to be sufficient if n is much larger than k [2] or if the tree is an arborescence [3] (that is, if the tree either has all paths branching outwards, or all paths branching inwards, from some designated root vertex).

The largest gap between these results and the conjecture of Havet and Thomassé occurs whenever $k = \Omega(n)$. In this paper, we reduce the required number of vertices in the tournament to $n + k + o(n)$, giving an asymptotic form of the Havet-Thomassé conjecture, as follows.

Theorem 1.1. *Let $\alpha > 0$. There exists $n_0 \in \mathbb{N}$ such that for any $n \geqslant n_0$, if G is a $((1 + \alpha)n + k)$-vertex tournament and T is an n-vertex oriented tree with k leaves, then G contains a copy of T.*

We turn now to consider whether oriented trees with low maximum degree are guaranteed to appear in tournaments with fewer vertices than the extremal cases for Sumner's conjecture. This appears more difficult than the study of oriented trees based on the number of leaves. Indeed, it is not known whether there is a function $h(\Delta)$ such that any $(n + h(\Delta))$-vertex tournament contains a copy of every oriented tree with maximum degree at most Δ, despite this question being raised by Kühn, Mycroft and Osthus [12]. Mycroft and Naia [14] asked whether $h(\Delta) = 2\Delta - 4$ is sufficient as long as n is much larger than Δ, while recalling extremal examples due to Allen and Cooley that demonstrate this would be tight for each Δ (see also [12]). When the maximum degree of the tree is sufficiently tightly bounded, it is known that few additional vertices are required in the tournament. Specifically, Kühn, Mycroft and Osthus [12] proved that, if Δ is a fixed constant, then every $(1 + o(1))n$-vertex tournament contains a copy of every n-vertex oriented tree with maximum degree $\Delta(T) \leqslant \Delta$ (however many leaves it has). Mycroft and Naia [14] later showed that the same conclusion holds even if the bound on $\Delta(T)$ is relaxed to one polylogarithmic in n. Here, we will relax the bound on $\Delta(T)$ much further still, showing that a degree bound linear in n is sufficient, as follows.

Theorem 1.2. *Let $\alpha > 0$. There exists $c > 0$ and $n_0 \in \mathbb{N}$ such that for any $n \geqslant n_0$, if G is a $(1+\alpha)n$-vertex tournament and T is an n-vertex oriented tree with $\Delta(T) \leqslant cn$, then G contains a copy of T.*

We prove Theorem 1.1 and Theorem 1.2 under a common framework, using regularity methods and random homomorphisms to reduce these theorems to critical cases amenable to a more direct study. Though our methods apply to all the trees covered by these theorems, the results discussed above imply that the critical case for consideration are those trees with $\Omega(n)$ leaves. The most difficult cases arise when these leaves are all close to each other within the tree, connected via some smaller core tree (see Figure 1). The challenge is then to be able to distribute these leaves around the tournament despite their location being quite tightly restricted by the location of the vertices in the core tree. For Theorem 1.2, the maximum degree condition will imply the core tree cannot be too small, and we exploit this to distribute the vertices of the core tree around the tournament. Here, the key novelty in our methods is the identification of the small core tree in the most challenging cases, and its embedding around the tournament.

For Theorem 1.1, we will be able to contract this small core tree in the most challenging cases to a single vertex without increasing the number of leaves. As the core tree is small, if we can find a copy of this contracted tree then we will be able to recover the original tree using suitable regularity techniques. The critical case will then be trees which have one very high degree vertex, whose removal results in components of at most constant size. Further simplification will allow us to assume that this high degree vertex has either in-degree or out-degree 0. This simplification focuses in on the hardest cases in our proof. To embed a tree T with one high out-degree vertex x with in-degree 0 into a tournament G, a natural approach is to place x at the vertex of G with highest out-degree, maximising the attachment possibilities for the components of $T - \{x\}$. Often, this is a good strategy (indeed, this approach will always succeed for the first tree depicted in Figure 1), but when many vertices of T are reached from x by travelling along a path beginning with an forwards edge followed by a backwards edge (such as for the second tree depicted in Figure 1), this may fail. Key to our proof is to use the failure in these cases to identify structural properties of the tournament, and thus a better location for the high out-degree vertex. This is the most significant novelty in our proof of Theorem 1.1, and enables the most difficult trees to be found in tournaments.

In Section 2, we will state our notation before giving a more detailed outline of the proof, and of this paper.

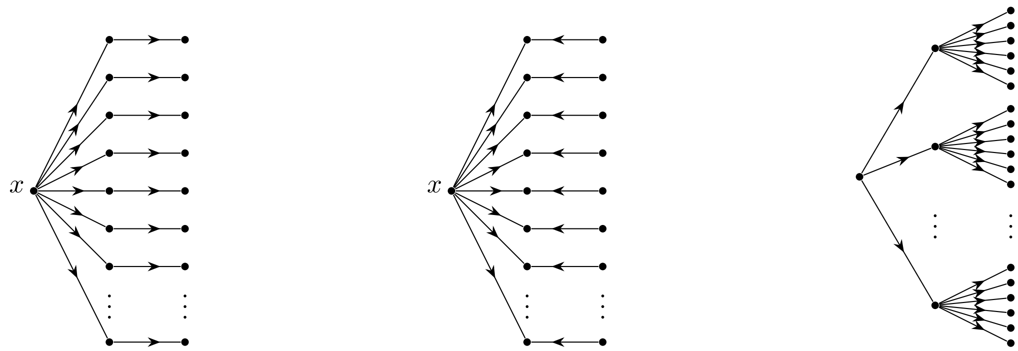

Figure 1: Examples of oriented trees with many leaves close to one another. While the first two trees have $\Delta(T) \approx n/2$, the third tree may be realised with $\Delta(T) \leqslant cn$ for a small constant c, making it a case of particular interest for Theorem 1.2.

2 Preliminaries

2.1 Notation

For a directed graph (digraph) G, we use $V(G)$ to denote the vertex set of G and $E(G)$ to denote the edge set of G. We write $|G| = |V(G)|$ for the order of G. Each element of $E(G)$ is an ordered pair (u, v) (which we write as uv, or $u \to_G v$), where $u, v \in V(G)$. If $uv \in E(G)$, then we say that v is an *out-neighbour* of u, and that u is an *in-neighbour* of v. Given $v \in V(G)$, the *out-neighbourhood* of v, written $N^+_G(v)$, is the set of out-neighbours of v in $V(G)$, and the *in-neighbourhood* of v, written $N^-_G(v)$ is the set of in-neighbours of v in $V(G)$. Throughout, we use $+$ and $-$ interchangeably with 'out' and 'in' respectively. For $X, Y \subseteq V(G)$ and $\diamond \in \{+, -\}$, we write $N^\diamond_G(X) = (\cup_{v \in X} N^\diamond_G(v)) \setminus X$ and $N^\diamond_G(X, Y) = N^\diamond_G(X) \cap Y$. For each $\diamond \in \{+, -\}$, the $\diamond$-*degree* of v in G is $d^\diamond_G(v) = |N^\diamond_G(v)|$, and for $X, Y \subseteq V(G)$ we also write $d^\diamond_G(X, Y) = |N^\diamond_G(X, Y)|$. For a vertex v, we also define its neighbourhood to be $N_G(v) = N^+_G(v) \cup N^-_G(v)$ and its degree to be $d_G(v) = |N_G(v)|$, and similarly define $N_G(X) = N^+_G(X) \cup N^-_G(X)$ for a set $X \subseteq V(G)$. We denote by $G[X]$ the induced sub-digraph of G with vertex set X and let $G - X = G[V(G) \setminus X]$. Subscripts are omitted wherever they are clear from context, as are rounding signs wherever they are not crucial.

An *oriented graph* is a digraph with at most one edge between any pair of vertices. A *tournament* G is an oriented graph whose underlying graph is a complete graph, i.e., for each $u, v \in V(G)$ with $u \neq v$, exactly one of uv or vu is in $E(G)$. An *oriented tree* (respectively, *oriented path*) is an oriented graph whose underlying graph is a tree (respectively, path). The *maximum degree* of an oriented tree T is the maximum degree of its underlying tree, and denoted $\Delta(T)$. A *directed path* from v_0 to v_ℓ is an oriented path of the form $v_0 \to v_1 \to \ldots \to v_\ell$. The *length* of a path P is $|P| - 1$. If G, H are digraphs, a *homomorphism* ϕ from G to H is a function $\phi : V(G) \to V(H)$ such that $\phi(u)\phi(v) \in E(H)$ whenever $uv \in E(G)$. We sometimes write $\phi : G \to H$ to denote a homomorphism from G to H, and refer to ϕ as an *embedding* of G into H.

Having proved, for example, a result holds for $\diamond = +$, we will occasionally deduce the same result for $\diamond = -$ by *directional duality*. That is, reversing all the relevant orientations and applying the result with $\diamond = +$ implies, after reversing the edges again, the result with $\diamond = -$. Where the symbol $\pm$ appears in a formula, we mean the formula holds for both $+$ and $-$ in place of $\pm$. For a set X and a function $f : X \to \mathbb{R}$, if $A \subseteq X$ we will often write $f(A)$ to mean $\sum_{x \in A} f(x)$ and $f(x_1, \ldots, x_k)$ to mean $f(\{x_1, \ldots, x_k\})$. For an event E depending on the parameter n, we will say that E holds *with high probability* if $\mathbb{P}(E) \to 1$ as $n \to \infty$. We also use standard hierarchy notation. That is, for $a, b \in (0, 1]$, we write $a \ll b$ to mean that there is a non-decreasing function $f : (0, 1] \to (0, 1]$ such that the subsequent statement holds whenever $a \leqslant f(b)$.

2.2 Proof outline

We will now sketch the proofs for both Theorem 1.1 and Theorem 1.2 together. In the introduction, we discussed how, for both results, we need to take particular care with trees which contain some small core subtree that restricts the distribution of the other vertices in the tree around the tournament. Therefore, we will identify a small core in any tree T, from which T can be recovered by first appending a collection of constant-sized trees, then connecting components by constant-length paths, and then iteratively attaching a small number of additional leaves. This decomposition is independent of the directions of the edges of T, and so we state it for non-oriented trees. More precisely, given any tree T, we find $T_0 \subseteq T_1 \subseteq T_2 \subseteq T_3 \subseteq T_4 = T$ (shown in Figure 2), such that

i) T_0 is small,
ii) T_1 is formed by adding constant-sized trees, each attached with an edge to some vertex of T_0,
iii) T_2 is formed by adding (unattached) constant-sized trees to T_1,
iv) T_3 is formed by adding long but constant-sized paths connecting the components of T_2, and
v) T_4 is formed by attaching constant-sized trees to T_3, so that few vertices are added in total.

Having found such a decomposition, we need a strategy for embedding these pieces. We note first that the vertices added in iii) and v) above pose little trouble given the spare vertices in our tournament. Indeed, within, say, any $\alpha n/2$ vertices in a tournament, any oriented tree with up to $\alpha n/6$ vertices can be found using known results (see Theorem 2.2). This allows the new constant-sized trees in iii) to be found greedily. For v), we note that setting aside a small random subset of the spare vertices preserves some in- and out-neighbours for almost all the remaining vertices in the tournament (see Proposition 2.14). Carrying out the embedding for i)–iv) within the tournament induced on these *good* remaining vertices, will then allow us to extend the embedding greedily to cover the final vertices in v) (see Corollary 2.3).

Thus, our focus is on how to embed the vertices in T_0 so that this can be extended to an embedding of T_1, and how to embed the paths at iv). In certain tournaments the paths at iv) can also be embedded straightforwardly by reserving a random set of vertices for this purpose. Where this is not possible, by removing a small set of vertices from the tournament, we will be able to partition the vertices into a sequence of linearly-sized sets, with all edges between the sets directed forwards along the sequence. This partition allows us to divide the tree naturally into pieces, which can then be found separately along the sequence of sets. We note that this is a streamlined version of a decomposition due to Kühn, Mycroft and Osthus [12, 13].

Let us assume then that the tournament is sufficiently well connected that the paths at iv) can be embedded within a reserved random set of vertices. We need then to embed T_0 so that the vertices of $T_1 - V(T_0)$ can be distributed throughout the tournament. To do this we will use the regularity lemma for digraphs, so that we may assign vertices to clusters before embedding them. The challenge is to identify some good clusters for T_0, for which we can assign the vertices in $V(T_1) \setminus V(T_0)$ across the other regularity clusters. The whole of T_1 can then be embedded using relatively standard regularity techniques, in combination with the result that any oriented tree with ℓ vertices can be found in tournaments with only $O(\ell)$ vertices.

Embedding the core T_0 of the tree and extending it to cover T_1 is the only part where the proofs of Theorem 1.1 and Theorem 1.2 that differ. For Theorem 1.1, the core tree can always be embedded within a single regularity cluster, which will allow us to reduce the problem to embedding trees where the core is a single vertex, x say (which may have very high degree), and further reduction will allow us to assume that all components of $T - \{x\}$ are attached to x by out-edges. Similar to the discussion in the introduction, here it would be natural to try embedding x to a cluster with as many out-edges as possible in a suitable 'reduced digraph' (see Section 2.5). This may fail, but we try this anyway, essentially embedding as much of T_1 as possible. If the embedding fails, it will be due to certain structural properties of the tournament which will allow us to move the embedding of x, along with some of the embedded vertices, to complete the embedding. This part of the proof, with its division into a number of detailed subcases, is the most technical aspect of our proof, but solves the key problem and allows the proof of Theorem 1.1.

Fortunately, embedding T_0 and extending this to cover T_1 is less involved for Theorem 1.2. The maximum degree condition in this case ensures that T_0 necessarily contains at least a large constant number of vertices (for example, for the third tree shown in Figure 1, we may identify T_0 with the star consisting of all non-leaf vertices). Thus, it is possible to distribute the vertices of T_0 across several regularity clusters if required for the even distribution of $V(T_1) \setminus V(T_0)$ throughout the tournament. For this, we identify a particular caterpillar-like structure which spans most of the clusters in the regularity graph (see Section 4.1).

Each aspect of the proof is discussed in more detail before it is carried out. In Section 2.3, we define

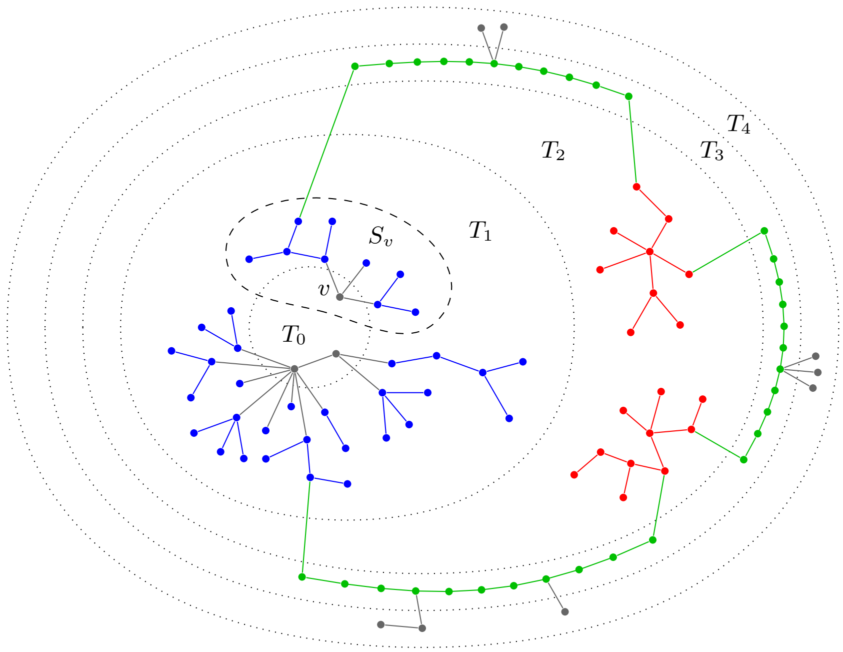

Figure 2: A simplified example of the tree decomposition $T_0 \subseteq T_1 \subseteq T_2 \subseteq T_3 \subseteq T_4 = T$ described by Lemma 2.1. In this illustration, the forest T_0 consists of a single edge together with an isolated vertex, and there are three paths, each with length $q = 12$, connecting components of T_2 to form the tree T_3.

precisely our tree decomposition and show that such a decomposition can always be found. In Section 2.4, we state some established tree embedding results. We then recall and discuss the regularity lemma is Section 2.5, before covering some probabilistic results in Section 2.6. In Section 3, we embed the core tree T_0 and extend the embedding to cover T_1 for Theorem 1.1 (see Theorem 3.1), deferring the most technical parts to Section 6 where we prove a key intermediate result, Theorem 3.4. In Section 4, we embed the core tree T_0 and extend the embedding to cover T_1 for Theorem 1.2 (see Theorem 4.1). These embeddings of T_0 extended to T_1 allow us then to prove both Theorem 1.1 and 1.2 in Section 5. We then finish with the deferred proof of Theorem 3.4 in Section 6.

2.3 Tree decomposition

We now give the tree decomposition discussed in the proof outline (see Figure 2). This is a modified version of a result of Kathapurkar and the second author (see [11]); though its proof is very similar, we include it for completion.

Lemma 2.1. *Let $1/n \ll 1/m \ll \eta, 1/q$ with $q \geqslant 2$. Then, any n-vertex tree T contains forests $T_0 \subseteq T_1 \subseteq T_2 \subseteq T_3 \subseteq T_4 = T$, such that T_3 is a tree, and the following properties hold.*

- **A1** $|T_0| \leqslant \eta n$.
- **A2** *T_1 is formed from T_0 by the vertex-disjoint addition of trees S_v, $v \in V(T_0)$, so that, for each $v \in V(T_0)$, $S_v - v$ is a forest with each component tree having size at most m.*
- **A3** *T_2 is the disjoint union of T_1 and a forest with each component tree having size at most m.*
- **A4** *T_3 is formed from T_2 by connecting components by paths of length q.*
- **A5** $|V(T_4) \setminus V(T_2)| \leqslant \eta n$.

Proof. Choose $\varepsilon > 0$ and $k \in \mathbb{N}$ such that $1/m \ll \varepsilon \ll 1/k \ll \eta, 1/q$. Fix an arbitrary vertex $t \in V(T)$. We start by finding a subtree T' of T which includes t and has few leaves, and is such that $T - V(T')$ is a forest of components each having size at most m. We do this by including in T' every vertex which appears on the path in T from t to many other vertices. That is, for each $v \in V(T)$, let $w(v)$ be the number of vertices $u \in V(T)$ whose path from t to u includes v (in particular, v is such a vertex). Let T' be the subgraph of T induced on all the vertices $v \in V(T)$ with $w(v) \geqslant m + 1$.

For each $v \in V(T')$, let S_v be the tree containing v in $T - (V(T') \setminus \{v\})$. Note that $S_v - v$ is a forest with each component tree having size at most m. Indeed, suppose T'' is a component of $S_v - v$, and let v' be the neighbour of v in T''. Since every path from a vertex $u \in V(T'')$ to t in T goes through v' (and then v), we have that $m \geqslant w(v') \geqslant |T''|$ (and, in fact, the final inequality is an equality). Observe further that, for any leaf v of T', $|S_v - v| = w(v) - 1 \geqslant m$, and, therefore, T' can have at most $n/m \leqslant \varepsilon n$ leaves.

We say a subpath P of T' is a bare path if all of the internal vertices v of P have $d_{T'}(v) = 2$, and we denote by $T' - P$ the graph formed from T' by removing all the edges and internal vertices of P. Using [11, Lemma 2.8], find in T' vertex disjoint bare paths $P_1, \ldots, P_r$ with length k such that

$$|T' - P_1 - \ldots - P_r| \leqslant 6k \cdot \varepsilon n + 2n/(k+1) \leqslant \eta n/4 \tag{1}$$

Note that $r \leqslant n/k$. For each path P_i, if possible, find within P_i a path P_i' with length at least $k - 2\eta^3 k$, such that, letting X_i, Y_i be the subpaths of P_i' induced by the first and last $q-1$ vertices of P_i', the following hold.

(i) $\sum_{v \in V(X_i)} |S_v|, \sum_{v \in V(Y_i)} |S_v| \leqslant \eta k/4$.

(ii) Letting Q_i be the component of $T - X_i - Y_i$ containing $P_i' - X_i - Y_i$, we have $|Q_i| \leqslant m$.

Say, with relabelling, these paths are $P_1', \ldots, P_{r'}'$. Let $T_0 = T' - P_1' - \ldots - P_{r'}'$. We will show that $|T_0| \leqslant \eta n$. Consider first the number of paths P_i which do not have length $q-2$ subpaths X_i, Y_i, each contained within $\eta^3 k$ of each end of P_i, and for which $\sum_{v \in V(X_i)} |S_v|, \sum_{v \in V(Y_i)} |S_v| \leqslant \eta k/4$. There are at most $n/(\lfloor \eta^3 k/(q-1) \rfloor (\eta k/4)) \leqslant \eta n/4k$ such paths. Of the remaining paths, at most n/m may fail to produce a P_i' due to having $|Q_i| > m$. Thus, we have $r' \geqslant r - \eta n/4k - n/m \geqslant r - \eta n/2k$.

Note that, for each $i \in [r']$, $|V(P_i) \setminus V(P_i')| \leqslant 2\eta^3 k$. Therefore

$$|T_0| \leqslant |T' - P_1 - \ldots - P_r| + k(r - r') + r'(2\eta^3 k) \leqslant \eta n/4 + k(\eta n/2k) + r(2\eta^3 k) \leqslant \eta n,$$

and hence **A1** holds. Let $T_1 = T[\cup_{v \in V(T_0)} V(S_v)]$. Recall that for each $v \in V(T')$, $S_v - v$ is a forest with each component tree having size at most m. Therefore, **A2** holds. Let $T_2 = T_1 \cup (\cup_{i \in [r']} Q_i)$, and note that **A3** holds. Note that

$$|V(T) \setminus V(T_2)| = \sum_{i \in [r']} \sum_{v \in V(X_i) \cup V(Y_i)} |S_v| \leqslant 2r(\eta k/4) \leqslant \eta n,$$

and hence **A5** holds. Let $T_3 = T[V(T_2) \cup (\cup_{i \in [r']} (V(X_i) \cup V(Y_i)))]$ and note that **A4** holds. Finally, the only vertices missing from T_3 are those in $S_v - v$ for each $v \in \cup_{i \in [r']} (V(X_i) \cup V(Y_i))$, and hence T_3 is a tree. □

2.4 Tree embedding results

We will often embed small parts of a tree into a subset of a tournament with many spare vertices. To do this we could use any result embedding an n-vertex tree into a tournament with $O(n)$ vertices, but for convenience we will use the following result of El Sahili [4].

Theorem 2.2 ([4, Corollary 2])**.** *For each $n \geqslant 2$, every $(3n-3)$-vertex tournament contains a copy of every n-vertex oriented tree.*

The following corollary shows how Theorem 2.2 can be used to extend a partial copy of a tree to a full copy, provided each vertex in the partial copy has sufficient out- and in-degree to the remaining vertices in the tournament.

Corollary 2.3. *Let G be a tournament with disjoint subsets $U, V \subseteq V(G)$. Let T be a tree, and suppose $T' \subseteq T$ is a subtree such that there is a copy S' of T' in $G[V]$. If $d^{\pm}_G(v, U) \geqslant 3|V(T) \setminus V(T')|$ for every $v \in V$, then S' can be extended to a copy of T in G, with $T - V(T')$ copied to U.*

Proof. Label the components of $T - V(T')$ as $T_1, \dots, T_r$, and take the largest $s \leqslant r$ such S' can be extended to a copy S of $T[V(T') \cup (\cup_{i\in[s]} V(T_i))]$. Suppose that $s < r$. Then, if $\diamond \in \{+, -\}$ is such that T_{s+1} is attached to T' by a $\diamond$-neighbour, and $v \in V(S')$ is the copy of the attachment point, then

$$d^{\diamond}_G(v, U \setminus V(S)) \geqslant 3|V(T) \setminus V(T')| - |T_1| - \ldots - |T_s| \geqslant 3|T_{s+1}|,$$

and so, by Theorem 2.2, $N^{\diamond}_G(v, U \setminus V(S))$ contains a copy of T_{s+1}, contradicting the maximality of s. Thus, S is a copy of T in G. □

We will also need to embed a tree into a subset of a tournament with a number of spare vertices depending on the number of leaves of the tree. Any such bound would suffice, but we will use the following result.

Theorem 2.4 ([2, Theorem 1.1])**.** *There is some $C > 0$ such that every $(n+Ck)$-vertex tournament contains a copy of every n-vertex oriented tree with k leaves.*

2.5 Regularity

Our embeddings use the regularity lemma for digraphs, by now a well-established tool in the study of tournaments (see, for example, [12, 13, 14]). As with the regularity lemma for graphs, this partitions most of the vertices of a tournament into clusters so that edges behave pseudorandomly between most pairs of clusters. We will now recall the notation needed to state the regularity lemma for digraphs.

Let G be a digraph. For disjoint subsets $A, B \subseteq V(G)$, define the *directed density* from A to B to be

$$d(A, B) = \frac{|E(A, B)|}{|A||B|},$$

where $E(A, B)$ denotes the set of edges of G directed from A towards B. Note that, if G is tournament, then $d(B, A) = 1 - d(A, B)$. We say that (A, B) forms an *ε-regular pair* if, for every $X \subseteq A$ such that $|X| \geqslant \varepsilon|A|$ and every $Y \subseteq B$ such that $|Y| \geqslant \varepsilon|B|$, we have $|d(X, Y) - d(A, B)| \leqslant \varepsilon$. Note that, for tournaments, $|d(X, Y) - d(A, B)| \leqslant \varepsilon$ if and only if $|d(Y, X) - d(B, A)| \leqslant \varepsilon$. We say that (A, B) forms an ε-regular pair of density at least μ if, in addition to forming an ε-regular pair, we also have $d(A, B) \geqslant \mu$.

We will use the following directed version of Szemerédi's regularity lemma proved by Alon and Shapira [1].

Theorem 2.5 (Regularity lemma for digraphs)**.** *Let $1/r_2 \ll 1/r_1 \ll \varepsilon$. Every digraph on a vertex set V of order at least r_1 partitions as $V = V_0 \cup V_1 \cup \ldots \cup V_r$, with $r_1 \leqslant r \leqslant r_2$, satisfying the following.*

B1 $|V_0| \leqslant \varepsilon|V|$.

B2 $|V_1| = \ldots = |V_r|$.

B3 *All but at most εr^2 pairs (V_i, V_j) with $1 \leqslant i < j \leqslant r$ are ε-regular.*

We now state the definition of an *ε-regular partition*. For convenience, we use a slightly different definition of an ε-regular partition of a tournament than is directly produced by Theorem 2.5, but which is gained through the removal of few clusters (see Corollary 2.7).

Definition 2.6. *An* ε-regular partition *of a tournament G is a partition $V(G) = V_1 \cup \ldots \cup V_r$ with $|V_1| = \ldots = |V_r|$ such that, for each fixed $i \in [r]$, (V_i, V_j) forms an ε-regular pair for all but at most $\sqrt{\varepsilon}r$ many $j \in [r]$.*

Corollary 2.7. *Let $\alpha > \beta > 0$ and $1/n \ll 1/r_2 \ll 1/r_1 \ll \varepsilon \ll \beta$. Let G be a $(1 + \alpha)n$-vertex tournament. Then, there is a subtournament $G' \subseteq G$ with $|G'| \geqslant (1+\alpha-\beta)n$, and an ε-regular partition $V(G') = V_1 \cup \ldots \cup V_r$ with $r_1 \leqslant r \leqslant r_2$.*

Proof. Given a tournament G, using Theorem 2.5, take a partition $V(G) = V_0 \cup V_1 \cup \ldots \cup V_{\bar{r}}$, with $2r_1 \leqslant \bar{r} \leqslant r_2$, satisfying **B1-B3**. By reordering, we may suppose there is some r with $0 \leqslant r \leqslant \bar{r}$ such that, for each fixed $i \in [\bar{r}]$, (V_i, V_j) forms an ε-regular pair with at least $(1 - \sqrt{\varepsilon}/2)\bar{r}$ many $j \in [\bar{r}]$ if and only if $i \in [r]$. By **B3**, we find $(\bar{r} - r)\sqrt{\varepsilon}\bar{r}/2 \leqslant \varepsilon\bar{r}^2$, and hence $r \geqslant (1 - 2\sqrt{\varepsilon})\bar{r} \geqslant r_1$. Let $G' = G[V_1 \cup \ldots \cup V_r]$. The desired properties for G' then follow by noting that $|V(G) \setminus V(G')| \leqslant |V_0| + \frac{\bar{r}-r}{\bar{r}}|G| \leqslant (\varepsilon + 2\sqrt{\varepsilon})|G| \leqslant \beta n$, and that $\sqrt{\varepsilon}\bar{r}/2 \leqslant \sqrt{\varepsilon}r$. □

We will use the following simple proposition on vertex degrees in ε-regular partitions.

Proposition 2.8. *Let $\varepsilon, \mu > 0$ and $r, m \in \mathbb{N}$. Suppose G is a tournament with disjoint subsets $V_0, V_1, \dots, V_r \subseteq V(G)$ of size $|V_0| = |V_1| = \dots = |V_r| = m$, such that (V_0, V_i) is an ε-regular pair of density at least μ for each $1 \leqslant i \leqslant r$. Fix a subset $U \subseteq \cup_{i\in[r]} V_i$. Then, all but at most εm vertices of V_0 have at least $(\mu - \varepsilon)(|U| - \varepsilon rm)$ out-neighbours in U.*

Proof. Let W be the set of vertices of V_0 which have fewer than $(\mu - \varepsilon)(|U| - \varepsilon rm)$ out-neighbours in U, and suppose that $|W| \geqslant \varepsilon m$. Then, for each $i \in [r]$, because (V_0, V_i) is an ε-regular pair of density at least μ, there are at least $|W|(\mu - \varepsilon)(|U \cap V_i| - \varepsilon m)$ edges directed from W to $U \cap V_i$, noting that this is trivial if $|U \cap V_i| \leqslant \varepsilon m$. Therefore, there are at least $|W|(\mu - \varepsilon)(|U| - \varepsilon rm)$ edges directed from W to U. However, from the definition of W, the number of edges from W to U is less than $|W|(\mu - \varepsilon)(|U| - \varepsilon rm)$, a contradiction. □

Our proofs will often allocate the vertices of a tree to the clusters of a regularity partition, before applying variations of standard regularity methods to embed these vertices so that they are (mostly) embedded to their assigned cluster. For this we will use, in part, the following simple proposition, which embeds a tree from an assignment in this way, provided that the tree is small and also that the vertices of the tree are not assigned to too many different clusters.

Proposition 2.9. *Let $1/m \ll \varepsilon \ll \beta, \mu, 1/\ell$. Suppose G is a tournament with subsets $V_1, \dots, V_\ell \subseteq V(G)$ of size $|V_1| = \dots = |V_\ell| = m$, and, for $j \in [\ell]$, let $U_j \subseteq V_j$ have size $|U_j| \geqslant \beta m$. Let T be an oriented tree with $|T| \leqslant \varepsilon m$, and suppose $\varphi : V(T) \to [\ell]$ is such that if $uv \in E(T)$ and $\varphi(u) \neq \varphi(v)$, then $(V_{\varphi(u)}, V_{\varphi(v)})$ is an ε-regular pair of density at least μ. Then, there is an embedding $\psi : T \to G$ with $\psi(v) \in U_{\varphi(v)}$ for each $v \in V(T)$.*

Proof. Let $V(T) = Y_1 \cup \dots \cup Y_r$ be a partition such that

- **C1** For each $i \in [r]$, $T[Y_i]$ is a connected component of $T[\phi^{-1}(j)]$ for some $j \in [\ell]$.
- **C2** For each $i \in [r]$, $T[Y_1 \cup \dots \cup Y_i]$ is a tree.

Let $s \in \{0\} \cup [r]$ be maximal such that, if $T_s = T[Y_1 \cup \dots \cup Y_s]$, then there is an embedding $\psi : T_s \to G$ with $\psi(v) \in V_{\varphi(v)}$ for every $v \in V(T_s)$, and, for every $v \in V(T_s)$ and $j \in [\ell]$ for which $(V_{\varphi(v)}, V_j)$ is an ε-regular pair, we have $d^+_G(\psi(v), U_j) \geqslant (d(V_{\varphi(v)}, V_j) - \varepsilon)\beta m$ and $d^-_G(\psi(v), U_j) \geqslant (d(V_j, V_{\varphi(v)}) - \varepsilon)\beta m$. Suppose, for contradiction, that $s < r$. If $s = 0$, then let $y \in Y_1$ be arbitrary and set $Z_1 = U_{\varphi(y)}$. If instead we have $s > 0$, then let $x \in V(T_s)$, $y \in Y_{s+1}$ and $\diamond \in \{+, -\}$ be such that $y \in N^\diamond_T(x)$, and set $Z_{s+1} = N^\diamond_G(\psi(x), U_{\varphi(y)})$. In either case, we find that $|Z_{s+1}| \geqslant \beta\mu m/2$ and $Z_{s+1} \subseteq U_{\varphi(y)}$. For each $j \in [\ell]$ such that $(V_{\varphi(y)}, V_j)$ is an ε-regular pair, all but at most εm vertices z of Z_{s+1} satisfy $d^+_G(z, U_j) \geqslant (d(V_{\varphi(y)}, V_j) - \varepsilon)\beta m$ and all but at most εm vertices z of Z_{s+1} satisfy $d^-_G(z, U_j) \geqslant (d(V_j, V_{\varphi(y)}) - \varepsilon)\beta m$. Therefore, as $\varepsilon \ll \beta, \mu, 1/\ell$, there is a subset $Z'_{s+1} \subseteq Z_{s+1} \setminus \psi(V(T_s))$ with $|Z'_{s+1}| \geqslant \beta\mu m/4$, such that, for every $z \in Z'_{s+1}$ and $j \in [\ell]$ for which $(V_{\varphi(y)}, V_j)$ is an ε-regular pair, we have $d^+_G(z, U_j) \geqslant (d(V_{\varphi(y)}, V_j) - \varepsilon)\beta m$ and $d^-_G(z, U_j) \geqslant (d(V_j, V_{\varphi(y)}) - \varepsilon)\beta m$. But then, by Theorem 2.2, there is a copy of $T[Y_{s+1}]$ in $G[Z'_{s+1}]$, and so we can extend ψ to cover Y_{s+1}, a contradiction to the maximality of s. □

As is common, given an *ε-regular partition* $V_1 \cup \dots \cup V_r$ of a tournament G, we will consider the reduced digraph R for the partition which has $V(R) = [r]$, and $ij \in E(R)$ exactly when (V_i, V_j) is an ε-regular pair with density comfortably larger than ε. We will sometimes delete edges arbitrarily from R so that there is at most 1 edge between any pair of vertices. As a small proportion of pairs of clusters in an ε-regular partition may not form an ε-regular pair, this will not necessarily result in a tournament. For this, we define an *ε-almost tournament*, as follows.

Definition 2.10. *An* ε-almost tournament R *is a digraph with at most one edge between any pair of vertices, and in which, for each $v \in V(R)$, there are at most $\varepsilon|R|$ vertices $u \in V(R)$ with $vu \notin E(R)$ and $uv \notin E(R)$.*

We will use the following simple property of ε-almost tournaments, which shows they each have some vertex with a good number of both in- and out-neighbours.

Proposition 2.11. *Let R be an ε-almost tournament on r vertices. Then, there exists a vertex $v \in V(R)$ such that $d^+_R(v), d^-_R(v) \geqslant \frac{r-1}{4} - \varepsilon r$.*

Proof. Let H be any tournament with $V(H) = V(R)$ such that $R \subseteq H$, and let $m = \frac{r-1}{4}$. Any set of $2m+1$ vertices in H contains a vertex with out-degree at least m and a vertex with in-degree at least m. Therefore, all but at most $2m$ vertices of H have out-degree at least m, and all but at most $2m$ vertices of H have in-degree at least m. Therefore, as $n > 4m$, there is some $v \in V(H)$ with $d^+_H(v), d^-_H(v) \geqslant \frac{r-1}{4}$. Then, $v \in V(R)$ satisfies $d^+_R(v), d^-_R(v) \geqslant \frac{r-1}{4} - \varepsilon r$. □

2.6 Probabilistic results

Parts of our embeddings will be random, or use some reserved random set. To analyse these parts, we will use the following probabilistic bounds (see, for example, [5]). The first is a Chernoff bound, and the second is Hoeffding's inequality.

Lemma 2.12. *If X is a binomial variable with standard parameters n and p, denoted $X = \mathrm{Bin}(n,p)$, and ε satisfies $0 < \varepsilon \leqslant 3/2$, then*

$$\mathbb{P}(|X - \mathbb{E}X| \geqslant \varepsilon\mathbb{E}X) \leqslant 2\exp\left(-\varepsilon^2\mathbb{E}X/3\right).$$

Theorem 2.13. *Let $X_1, \ldots, X_n$ be independent random variables with X_i bounded by the interval $[a_i, b_i]$ for $i \in [n]$. Let $X = \sum_{i\in[n]} X_i$. Then, for any $t > 0$, we have*

$$\mathbb{P}(|X - \mathbb{E}X| \geqslant t) \leqslant 2\exp\left(-\frac{2t^2}{\sum_{i\in[n]}(b_i - a_i)^2}\right).$$

It will often be convenient for most of the vertices to have large in- and out-degree into a reserved random set, for which we use the following result.

Proposition 2.14. *Fix $p > 0$. Let G be a tournament with $n \leqslant |G| \leqslant 3n$. Let $U \subseteq V(G)$ be a random subset, with elements from $V(G)$ chosen independently at random with probability p. Let V' be the set of vertices $v \in V(G) \setminus U$ for which $d^\pm(v, U) \geqslant p^2 n$. Then, with high probability, $pn/2 \leqslant |U| \leqslant 4pn$, and $|V(G) \setminus V'| \leqslant 12pn$.*

Proof. By Lemma 2.12 and the fact that $pn \leqslant \mathbb{E}|U| \leqslant 3pn$, we have $pn/2 \leqslant |U| \leqslant 4pn$ with high probability. If $v \in V(G)$ is such that $d^\pm_G(v) \geqslant 2pn$, then, by setting $\varepsilon = 1/2$ in Lemma 2.12, the probability that $d^\pm(v, U) \geqslant p^2 n$ fails for v is at most $4\exp(-p^2 n/6)$. Any set of $4pn + 1$ vertices in G contains a vertex with out-degree at least $2pn$ and a vertex with in-degree at least $2pn$. So at most $4pn$ vertices v of G have $d^+_G(v) < 2pn$ and at most $4pn$ vertices of G have $d^-_G(v) < 2pn$. Therefore, the probability that $|V(G) \setminus V'| \leqslant |U| + 8pn$ fails is at most $12n\exp(-p^2 n/6)$. So U satisfies both $pn/2 \leqslant |U| \leqslant 4pn$ and $|V(G) \setminus V'| \leqslant 12pn$ with high probability. □

3 Theorem 1.1: embedding the core and attached small trees

In this section, following the proof outline in Section 2.2, we embed T_0 and T_1 for Theorem 1.1. In the embeddings we may assume that T_0 is connected (i.e., that it is a tree not just a forest), and so our embedding of T_0 and T_1 will follow from the following theorem applied with $T = T_1$ and $T_0 = T_0$.

Theorem 3.1. *Let $1/n \ll \eta \ll \bar{\alpha} < 1$. Suppose T is an n-vertex k-leaf oriented tree with a subtree $T_0 \subseteq T$, such that $|T_0| \leqslant \eta n$ and every component of $T - V(T_0)$ has size at most ηn. Then, any $((1+\bar{\alpha})n + k)$-vertex tournament contains a copy of T.*

We require, for Theorem 3.1, the components of $T - V(T_0)$ to be bounded above by ηn, for some appropriately small η. This is not required for its application, where these components will have constant size (see **A2**), but this small linear bound follows at no additional cost. As discussed in Section 2, to embed the core T_0 and extend this embedding to T_1, we will first allocate the vertices of the tree to regularity clusters. This allocation requires care beyond that in previous embeddings of trees in tournaments (see [12, 13, 14]) as the large degree of some vertices in the tree require edges not just to have sufficient density for regularity embedding techniques to be effective (i.e., $\varepsilon \ll \mu$ in Proposition 2.9), but sufficient density for potentially linearly many neighbours of a vertex to be embedded within the same regularity cluster. For this, we find it convenient to consider an ε-regular partition of clusters $V_1 \cup \ldots \cup V_r$ as a weighted complete looped digraph D with vertex set $[r]$ and edge weights $d(e) \in [0,1]$, $e \in E(D)$, indicating the edge density between ε-regular pairs of clusters. We call the sets of edge weights we typically encounter *ε-complete*, as follows.

Definition 3.2. *Given a complete looped digraph D on vertex set $[r]$, we say edge weights $d(e) \in [0,1]$, $e \in E(D)$, are* ε-complete *if the following holds.*

D *For each $j \in [r]$, $d(j,j) = 1$ and, for all but at most εr values of $i \in [r] \setminus \{j\}$, $d(i,j) + d(j,i) = 1$.*

From this perspective, an allocation of the vertices of an oriented tree T to the regularity clusters is a homomorphism from T to a complete looped digraph with an appropriate set of edge weights, as follows.

Definition 3.3. *Given an oriented tree T and a complete looped digraph D with associated edge-weights $d(e) \in [0,1]$, $e \in E(D)$, we say that a function $\phi : V(T) \to V(D)$ is a* homomorphism from T to D *if $d(\phi(v), \phi(w)) > 0$ whenever $vw \in E(T)$.*

We need to find such a homomorphism satisfying additional properties, such as a limit to how many vertices are assigned to each cluster. The allocation we find for Theorem 3.1 will always assign the vertices of T_0 to a single cluster, whose index we call j_t, and then distribute the vertices of the components of $T - V(T_0)$ across the other regularity clusters. One complication is the diversity of oriented trees which may appear as components of $T - V(T_0)$. Having chosen j_t, the restriction to where a vertex v in $T - V(T_0)$ can be embedded depends on the path from T_0 to v in T, and its edge directions. However, it will turn out that we need only be concerned with up to the first three edges from T_0 on such a path, and this allows us to categorise the components of $T - V(T_0)$.

Very roughly, all we need to consider is what proportion of the vertices in $T - V(T_0)$ have paths from $T - V(T_0)$ beginning with any given oriented path of length at most 3. To record this we introduce H (defined precisely below, and shown in Figure 3), a small oriented forest whose vertices roughly correspond to the groups of vertices we need to consider, with vertex weights on them to represent the proportion of vertices in this group (the vertices in H marked with a bar are used so that we can assume y^+ has more weight than x^+ and y^- has more weight than x^-). In this way, H represents an average of the components of $T - V(T_0)$. We cannot represent the embedding of H with only one homomorphism, as vertices of the same group must be spread over different regularity clusters. Instead, we find a random homomorphism ϕ of H along with a fixed index j_t for the vertices of T_0. As the location of T_0 is always into the cluster indexed by j_t, we can assign the vertices of each component of $T - V(T_0)$ according to an independent selection of ϕ. From the properties of ϕ this will (with high probability) spread the vertices of $T - V(T_0)$ out among the regularity clusters appropriately. Later, in Section 6, in some parts of the proof we will use a list of homomorphisms to allocate room for different components before selecting from these homomorphisms randomly, but away from this the random homomorphism allows us to concisely state and prove lemmas representing various cases for our main technical theorem.

The proof of the existence of the random homomorphism ϕ (Theorem 3.4 below) is the most involved part of this paper, and as noted before we defer this to Section 6. Before Theorem 3.4, we first define the graph H used to represent the components of $T - V(T_0)$ (see also Figure 3). Let H be the oriented forest with vertex and edge sets given by

$$V(H) = \{x^+, y^+, z^+, u^+, w^+, \bar{x}^+, \bar{z}^+, \bar{u}^+, \bar{w}^+, x^-, y^-, z^-, u^-, w^-, \bar{x}^-, \bar{z}^-, \bar{u}^-, \bar{w}^-\},$$
$$E(H) = \{x^+y^+, z^+x^+, z^+u^+, w^+z^+, \bar{z}^+\bar{x}^+, \bar{z}^+\bar{u}^+, \bar{w}^+\bar{z}^+, y^-x^-, x^-z^-, u^-z^-, z^-w^-, \bar{x}^-\bar{z}^-, \bar{u}^-\bar{z}^-, \bar{z}^-\bar{w}^-\}.$$

For each $\diamond \in \{+, -\}$, let $X^\diamond = \{x^\diamond, \bar{x}^\diamond\}$. Let $X = X^+ \cup X^-$. As described above, we use H with appropriate weights to represent the average component of $T - V(T_0)$. Vertices in X represent the vertices connected to T_0 in T, which may therefore be neighbours of any very high degree vertices in T_0, and so we need to pay particular attention to how often they are embedded to each regularity cluster.

Given appropriate weights for the vertices of H and a complete looped digraph D with edge weights, using Theorem 3.4 we find a $j_t \in V(D)$ and a random $\phi : H \to D$ with properties **E1**-**E4** which ensure that ϕ can be used to randomly assign the vertices of components across a regularity partition whose reduced digraph corresponds to D. By **E1**, we will have with high probability that ϕ is a homomorphism, so that the regularity properties can be used to embed any edges assigned between two regularity clusters, and, for

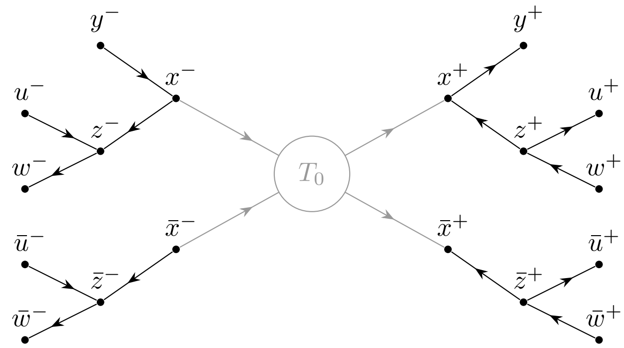

Figure 3: The oriented forest H.

convenience, no vertex in X is assigned to j_t. **E2** ensures that (on average) not too much weight is allocated to a single cluster. **E3** ensures that (on average) the weight of vertices in X^+ (i.e., those which need to be attached by an out-edge to T_0, which is allocated to j_t), or X^-, allocated to each cluster is not too much, where the limit is dictated by the appropriate density from that cluster to j_t, or from j_t to that cluster. Finally, **E4** is present to later ensure that certain vertices of $N_T(V(T_0))$ are allocated to a different cluster to their neighbours, which will assist with the embedding process.

We use the function γ (see (2) in Theorem 3.4) to control the total size of components we can embed using ϕ relative to the size of the tournament from which D is derived. As we use this for Theorem 1.1 it should be related to the number of leaves. In the application of Theorem 3.4, the weight on the vertices with base label 'x' or 'z' is distributed so that it can be bounded based on the number of leaves of the original tree (and in certain cases uses a lower bound than that required for Theorem 1.1).

Theorem 3.4. *Let $1/r \ll \varepsilon \ll \mu \ll \alpha < 1$. Let $\beta : V(H) \to [0,1]$ be a function satisfying $\sum_{v\in V(H)} \beta(v) = 1$ with $\beta(y^+) \geqslant \beta(x^+)$ and $\beta(y^-) \geqslant \beta(x^-)$, and, for every $v \in V(H)$, $\beta(v) \geqslant \mu$. Let D be a complete looped digraph on vertex set $[r]$ with ε-complete edge weights $d(e)$, $e \in E(D)$. Let*

$$\gamma = \max\{\beta(x^+, \bar{x}^+), \beta(z^+, \bar{z}^+)\} + \max\{\beta(x^-, \bar{x}^-), \beta(z^-, \bar{z}^-)\}. \tag{2}$$

Then, there is some $j_t \in [r]$ and a random $\phi : H \to D$ such that the following hold.

E1 *With probability 1, ϕ is a homomorphism from H to D, and $j_t \notin \phi(\{x^+, \bar{x}^+, x^-, \bar{x}^-\})$.*

E2 *For each $j \in [r]$, $\mathbb{E}(\beta(\phi^{-1}(j))) \leqslant \frac{1+\gamma+\alpha}{r}$.*

E3 *For each $j \in [r]$, either*

- **E3.1** *$\mathbb{E}(\beta(\phi^{-1}(j) \cap X^+)) \leqslant d(j_t, j) \cdot \frac{1+\gamma+\alpha}{r}$ and $\mathbb{E}(\beta(\phi^{-1}(j) \cap X)) \leqslant d(j, j_t) \cdot \frac{1+\gamma+\alpha}{r}$, or*
- **E3.2** *$\mathbb{E}(\beta(\phi^{-1}(j) \cap X^-)) \leqslant d(j, j_t) \cdot \frac{1+\gamma+\alpha}{r}$ and $\mathbb{E}(\beta(\phi^{-1}(j) \cap X)) \leqslant d(j_t, j) \cdot \frac{1+\gamma+\alpha}{r}$.*

E4 *With probability 1, we have $|\phi(e)| = 2$ for every $e \in E(H)$.*

As the proof for Theorem 3.4 is deferred to Section 6, here we take Theorem 3.4 as our starting point and prove Theorem 3.1 from it. In this proof, we first define a homomorphism f from $T - V(T_0)$ to H, and then use f to define an appropriate weight function $\beta : V(H) \to [0,1]$. Then, taking an ε-regular partition of a tournament G with vertex classes $V_1, \dots, V_r$, we choose the appropriate edge-weighted complete looped digraph D. Applying Theorem 3.4, we obtain $j_t \in [r]$ and a random homomorphism $\phi : H \to D$. We then embed T_0 into a subset of good vertices inside V_{j_t}. By sampling ϕ for each component of $T - V(T_0)$, we then get a homomorphism $\hat{\varphi}$ from $T - V(T_0)$ to D, which acts as the allocation of the remaining vertices to clusters of the tournament. Because T_0 may have vertices with high degree, we need to allocate room in the clusters for vertices in $N_T(V(T_0))$. This restricts their embedding, possibly preventing us from being able to follow the true random allocation. However, by switching the allocations given to identical pairs of components of $T - V(T_0)$, we can still embed most components of $T - V(T_0)$. Finally, any remaining components of $T - V(T_0)$ that are not embedded can then be handled by greedily embedding them to a random subset $U \subseteq V(G)$ reserved at the beginning of the proof.

Proof of Theorem 3.1. Let $\alpha = \bar{\alpha}/35$ and introduce constants $\mu, \varepsilon, r_1, r_2$ such that $\eta \ll 1/r_2 \ll 1/r_1 \ll \varepsilon \ll \mu \ll \alpha$. For each $x \in N_T(V(T_0))$, let S_x be the component of $T - V(T_0)$ containing x.

Let G be a $((1+35\alpha)n+k)$-vertex tournament, and note that $n \leqslant |G| \leqslant 3n$. Let $U \subseteq V(G)$ be a random subset, with elements from $V(G)$ chosen independently at random with probability 2α, and let V' be the set of vertices $v \in V(G) \setminus U$ with $d^{\pm}_G(v, U) \geqslant 4\alpha^2 n$. By Proposition 2.14, we may proceed assuming that $|U| \geqslant \alpha n$ and $|V'| \geqslant ((1+11\alpha)n + k)$.

Define $f_0 : N_T(V(T_0)) \to \{x^+, \bar{x}^+, x^-, \bar{x}^-\}$ as follows. For each $\diamond \in \{+,-\}$ and $v \in N^{\diamond}_T(V(T_0))$, set $f_0(v) = x^{\diamond}$ if $N^{\diamond}_T(v) \setminus V(T_0) \neq \emptyset$, and set $f_0(v) = \bar{x}^{\diamond}$ if $N^{\diamond}_T(v) \setminus V(T_0) = \emptyset$. Then, let $f : V(T) \setminus V(T_0) \to V(H)$ be the unique homomorphism from $T - V(T_0)$ to H extending f_0 such that $f^{-1}(\{x^+, \bar{x}^+, x^-, \bar{x}^-\}) = N_T(V(T_0))$, $f^{-1}(\{z^+, \bar{z}^+\}) = N^-_T(N^+_T(V(T_0))) \setminus V(T_0)$, and $f^{-1}(\{z^-, \bar{z}^-\}) = N^+_T(N^-_T(V(T_0))) \setminus V(T_0)$.

Let $\beta : V(H) \to [0,1]$ be given by setting, for each $v \in V(H)$,

$$\beta(v) = \frac{|f^{-1}(v)| + 2\mu n}{|V(T) \setminus V(T_0)| + 36\mu n}, \tag{3}$$

and note that β is a function satisfying $\sum_{v \in V(H)} \beta(v) = 1$, such that $\beta(v) \geqslant \mu$ for every $v \in V(H)$. Set

$$\gamma = \max\{\beta(x^+, \bar{x}^+), \beta(z^+, \bar{z}^+)\} + \max\{\beta(x^-, \bar{x}^-), \beta(z^-, \bar{z}^-)\}. \tag{4}$$

We remark that, for each $\diamond \in \{+,-\}$, if $v \in f^{-1}(x^{\diamond})$, then there is some $v' \in N^{\diamond}_T(v) \setminus V(T_0)$ with $f(v') = y^{\diamond}$. Therefore, $\beta(y^+) \geqslant \beta(x^+)$ and $\beta(y^-) \geqslant \beta(x^-)$. We also remark that, for each $x \in N_T(V(T_0))$, the number of leaves of T appearing in S_x is at least $\max\{1, |f^{-1}(\{z^+, \bar{z}^+, z^-, \bar{z}^-\}) \cap V(S_x)|\}$, and hence

$$\begin{aligned}
k &\geqslant \sum_{x \in N_T(V(T_0))} \max\{1, |f^{-1}(\{z^+, \bar{z}^+, z^-, \bar{z}^-\}) \cap V(S_x)|\} \\
&\geqslant \max\{|N^+_T(V(T_0))|, |f^{-1}(\{z^+, \bar{z}^+\})|\} + \max\{|N^-_T(V(T_0))|, |f^{-1}(\{z^-, \bar{z}^-\})|\} \\
&= \max\{|f^{-1}(\{x^+, \bar{x}^+\})|, |f^{-1}(\{z^+, \bar{z}^+\})|\} + \max\{|f^{-1}(\{x^-, \bar{x}^-\})|, |f^{-1}(\{z^-, \bar{z}^-\})|\}.
\end{aligned}$$

Therefore, by (3) and (4), $\gamma \leqslant \alpha + k/n$ and hence $|V'| \geqslant (1 + \gamma + 10\alpha) \cdot n$.

By Corollary 2.7, there is some r with $r_1 \leqslant r \leqslant r_2$ and disjoint subsets $V_1, \ldots, V_r \subseteq V'$, with $|V_1| = \ldots = |V_r| \geqslant (1 + \gamma + 9\alpha) \cdot n/r$, such that $V_1 \cup \ldots \cup V_r$ is an ε-regular partition of $G[V_1 \cup \ldots \cup V_r]$. Let D be a complete looped digraph on vertex set $[r]$ with edge weights $d(e)$, $e \in E(D)$ given by setting $d(j,j) = 1$ for every $j \in [r]$, $d(j,j') = 0$ for every $jj' \in E(D)$ for which $j \neq j'$ and $(V_j, V_{j'})$ is not an ε-regular pair, and, for every $jj' \in E(D)$ for which $(V_j, V_{j'})$ forms an ε-regular pair, setting

$$d(j,j') = \begin{cases} 1 & \text{if } d(V_j, V_{j'}) > 1 - \mu, \\ d(V_j, V_{j'}) & \text{if } \mu \leqslant d(V_j, V_{j'}) \leqslant 1 - \mu, \\ 0 & \text{if } d(V_j, V_{j'}) < \mu. \end{cases}$$

We remark that the edge weights $d(e)$, $e \in E(D)$ are $\sqrt{\varepsilon}$-complete, and, for $j \neq j'$, if $d(j,j') > 0$, then $(V_j, V_{j'})$ is an ε-regular pair with density satisfying $d(V_j, V_{j'}) \geqslant \mu$ and $d(V_j, V_{j'}) \geqslant (1-\mu) \cdot d(j,j')$.

By Theorem 3.4 applied to β and D, there is then some $j_t \in [r]$ and a random $\phi : H \to D$ such that **E1-E4** hold. Let J_1 be the set of $j \in [r]$ for which we have **E3.1**, and let $J_2 = [r] \setminus J_1$, so that **E3.2** holds for every $j \in J_2$. For $j \in [r]$, let $U_j, W_j \subseteq V_j$ be disjoint subsets with $|U_j| = (1+\gamma+4\alpha) \cdot n/r$ and $|W_j| = 3\alpha \cdot n/r$. Let Z be the set of $z \in V_{j_t} \setminus (U_{j_t} \cup W_{j_t})$ such that the following holds for ϕ with probability at least $1 - \sqrt{\varepsilon}$.

$$\begin{aligned}
d^+_G(z, U_{\phi(x^+)}) &\geqslant d(j_t, \phi(x^+)) \cdot (1+\gamma+3\alpha) \cdot n/r, & d^+_G(z, W_{\phi(x^+)}) &\geqslant 2\alpha\mu \cdot n/r, \\
d^+_G(z, U_{\phi(\bar{x}^+)}) &\geqslant d(j_t, \phi(\bar{x}^+)) \cdot (1+\gamma+3\alpha) \cdot n/r, & d^+_G(z, W_{\phi(\bar{x}^+)}) &\geqslant 2\alpha\mu \cdot n/r, \\
d^-_G(z, U_{\phi(x^-)}) &\geqslant d(\phi(x^-), j_t) \cdot (1+\gamma+3\alpha) \cdot n/r, & d^-_G(z, W_{\phi(x^-)}) &\geqslant 2\alpha\mu \cdot n/r, \\
d^-_G(z, U_{\phi(\bar{x}^-)}) &\geqslant d(\phi(\bar{x}^-), j_t) \cdot (1+\gamma+3\alpha) \cdot n/r, & d^-_G(z, W_{\phi(\bar{x}^-)}) &\geqslant 2\alpha\mu \cdot n/r.
\end{aligned} \tag{5}$$

Claim 3.5. $|Z| \geqslant \alpha \cdot n/r$.

Proof of Claim 3.5. Let $\bar{Z}$ be the set of $z \in V_{j_t}$ such that (5) fails with probability at least $\sqrt{\varepsilon}$. If $|\bar{Z}| < \alpha \cdot n/r$, then, as $|V_{j_t} \setminus (U_{j_t} \cup W_{j_t})| \geqslant 2\alpha \cdot n/r$, the claim follows. So assume for contradiction that $|\bar{Z}| \geqslant \alpha \cdot n/r$.

Let Ω be the set of homomorphisms $\bar{\phi} : H \to D$ such that $j_t \notin \bar{\phi}(\{x^+, \bar{x}^+, x^-, \bar{x}^-\})$ and $d(j_t, \bar{\phi}(x^+))$, $d(j_t, \bar{\phi}(\bar{x}^+))$, $d(\bar{\phi}(x^-), j_t))$, and $d(\bar{\phi}(\bar{x}^-), j_t)$ are all positive, and hence, by our choice of $d(e)$, $e \in E(D)$, are all at least μ. Note that, by **E1**, $\mathbb{P}(\phi \in \Omega) = 1$. Given $\bar{\phi} \in \Omega$, let $B_{\bar{\phi}}$ be the set of $z \in V_{j_t} \setminus (U_{j_t} \cup W_{j_t})$ such that (5) fails for $\phi = \bar{\phi}$. We claim that $|B_{\bar{\phi}}| \leqslant 24\varepsilon \cdot n/r$ for every $\bar{\phi} \in \Omega$. Indeed, if $\bar{\phi} \in \Omega$, then $(V_{j_t}, V_{\bar{\phi}(x^+)})$ is an ε-regular pair of density $d(V_{j_t}, V_{\bar{\phi}(x^+)}) \geqslant \min\{d(j_t, \bar{\phi}(x^+)), 1-\mu\} \geqslant \mu$, and so the number of $z \in V_{j_t}$ for which we do not have $d^+_G(z, U_{\bar{\phi}(x^+)}) \geqslant d(j_t, \bar{\phi}(x^+)) \cdot (1+\gamma+3\alpha) \cdot n/r$ is at most $\varepsilon|V_{j_t}| \leqslant 3\varepsilon \cdot n/r$ (using, for example, Proposition 2.8, with $r' = 1$ and $\mu' = d(V_{j_t}, V_{\bar{\phi}(x^+)})$). More generally, if $\bar{\phi} \in \Omega$, then $(V_{j_t}, V_{\bar{\phi}(x^+)})$, $(V_{j_t}, V_{\bar{\phi}(\bar{x}^+)})$, $(V_{\bar{\phi}(x^-)}, V_{j_t})$, and $(V_{\bar{\phi}(\bar{x}^-)}, V_{j_t})$ all form ε-regular pairs (of density at least μ), and thus each one of the inequalities of (5) fails for at most $3\varepsilon \cdot n/r$ many $z \in V_{j_t}$. Hence we have $|B_{\bar{\phi}}| \leqslant 8 \cdot 3\varepsilon \cdot n/r = 24\varepsilon \cdot n/r$ for every $\bar{\phi} \in \Omega$, as claimed. But then

$$\alpha\sqrt{\varepsilon} \cdot n/r \leqslant |\bar{Z}| \cdot \sqrt{\varepsilon} \leqslant \sum_{\bar{\phi} \in \Omega} \mathbb{P}(\phi = \bar{\phi}) \cdot |B_{\bar{\phi}}| \leqslant 24\varepsilon \cdot n/r,$$

a contradiction as $\varepsilon \ll \alpha$. ⊡

Note that, by Claim 3.5 and as $\eta \ll \alpha, 1/r_2$, $|Z| \geqslant 3\eta n \geqslant 3|T_0|$. Therefore, using Theorem 2.2, let $\psi : T_0 \to G$ be an embedding so that $\psi(V(T_0)) \subseteq Z$. For each $x \in N_T(V(T_0))$, let $z_x \in Z$ be the image under ψ of the unique neighbour of x in $V(T_0)$. Our aim now is to extend ψ to cover the components S_x, $x \in N_T(V(T_0))$, with each $\psi(x)$ in the appropriate in- or out-neighbourhood of z_x.

Given $v \in V(T) \setminus V(T_0)$, let $x(v) \in N_T(V(T_0))$ be the unique vertex such that $v \in V(S_{x(v)})$. For each $x \in N_T(V(T_0))$, choose a homomorphism $\phi_x : H \to D$ by sampling ϕ, conditioned on the event that (5) holds for $z = z_x$. Define a function $\hat{\varphi} : V(T) \setminus V(T_0) \to [r]$ by setting $\hat{\varphi}(v) = \phi_{x(v)}(f(v))$. We remark that $\hat{\varphi}$ is a homomorphism from $T - V(T_0)$ to D, with $|\hat{\varphi}(V(S_x))| \leqslant |H|$ for every $x \in N_T(V(T_0))$.

We will later see that a switching argument can be used to extend our embedding to cover all but at most $\mu^4 n$ components of $T - V(T_0)$. We should therefore take care to ensure that the components that are missed each have size at most $1/\mu^3$, so that only at most μn vertices remain uncovered by the embedding (which can be embedded greedily afterwards). Therefore, we now partition the components of $T - V(T_0)$ according to their size. Let X_0 be the set of $x \in N_T(V(T_0))$ with $|S_x| \leqslant 1/\mu^3$, and let $Y_0 = N_T(V(T_0)) \setminus X_0$, so that $|S_x| > 1/\mu^3$ whenever $x \in Y_0$. Note that $|Y_0| \leqslant \mu^3 n$. Roughly speaking, we will try to embed each $v \in V(T) \setminus V(T_0)$ into $V_{\hat{\varphi}(v)}$, with each $v \in Y_0$ embedded into $W_{\hat{\varphi}(v)}$ and each $v \in V(T) \setminus (V(T_0) \cup Y_0)$ embedded into $U_{\hat{\varphi}(v)}$. This motivates the following claim, which we will prove later. For this, for each $j \in [r]$ and $\diamond \in \{+,-\}$, let $X_j^\diamond$ (respectively, $Y_j^\diamond$) be the set of vertices in X_0 (respectively, Y_0) which are $\diamond$-neighbours of $V(T_0)$ and allocated to V_j by $\hat{\varphi}$. That is, for each $j \in [r]$ and $\diamond \in \{+,-\}$, let $X_j^\diamond = X_0 \cap f^{-1}(X^\diamond) \cap \hat{\varphi}^{-1}(j)$ and $Y_j^\diamond = Y_0 \cap f^{-1}(X^\diamond) \cap \hat{\varphi}^{-1}(j)$.

Claim 3.6. *With probability at least* $3/4$*, the following properties hold.*

- **F1** *For every* $j \in [r]$, $|\hat{\varphi}^{-1}(j)| \leqslant (1+\gamma+3\alpha) \cdot n/r$.
- **F2** *For every* $j \in [r]$ *and* $x \in X_j^+ \cup X_j^-$,
  - **F2.1** *if* $j \in J_1$, *then* $d^+_G(z_x, U_j) \geqslant |X_j^+|$ *if* $x \in X_j^+$ *and* $d^-_G(z_x, U_j) \geqslant |X_j^+ \cup X_j^-|$ *if* $x \in X_j^-$;
  - **F2.2** *if* $j \in J_2$, *then* $d^-_G(z_x, U_j) \geqslant |X_j^-|$ *if* $x \in X_j^-$ *and* $d^+_G(z_x, U_j) \geqslant |X_j^+ \cup X_j^-|$ *if* $x \in X_j^+$.
- **F3** $|Y_j^+ \cup Y_j^-| \leqslant \alpha\mu \cdot n/r$ *for every* $j \in [r]$.

We therefore proceed with the assumption that properties **F1-F3** hold. Under this assumption, $\hat{\varphi}$ now acts as a good guide for assigning vertices of $V(T) \setminus V(T_0)$ to the clusters $V_1, \ldots, V_r$. **F2** ensures only just enough room in each U_j for the allocated vertices from X_0 (embedding first X_j^+ and then X_j^- in the case of **F2.1**, or vice versa in the case of **F2.2**), and therefore we start by embedding X_0. We then embed as much of $\cup_{x \in X_0 \cup Y_0} V(S_x)$ as possible. Given $x \in X_0$, if it is not possible to embed S_x according to $\hat{\varphi}$ (for example, if the image of x does not have the appropriate neighbours in the sets $U_{\hat{\varphi}(v)}$, $v \in N_{S_x}(x)$), then we may still be able to find an embedding for S_x by switching the image of x with the image of some $x' \in X_j^\diamond$, and correspondingly switching the assignments associated to S_x and $S_{x'}$ before then embedding S_x. A new assignment φ, with many of the same properties as $\hat{\varphi}$, will keep track of these switches. Through this argument, we will extend ψ to cover most of $\cup_{x \in X_0} V(S_x)$ (see Claim 3.8). On the other hand, because each $Y_j^\diamond$ is small (see **F3**), we can greedily extend ψ to cover $\cup_{x \in Y_0} V(S_x)$ by embedding each $x \in Y_0$ into $W_{\hat{\varphi}(x)}$ (see Claim 3.7). Thus we will have an almost complete embedding of T, which can then be extended to a full embedding using Corollary 2.3 and the vertices in U we reserved at the start of the proof.

Extend ψ to cover X_0 as follows.

- For each $j \in J_1$, using **F2.1**, greedily extend ψ to first cover X_j^+, and then to cover X_j^-, so that $\psi(X_j^+ \cup X_j^-) \subseteq U_j$.
- For each $j \in J_2$, using **F2.2** greedily extend ψ to first cover X_j^-, and then to cover X_j^+, so that $\psi(X_j^+ \cup X_j^-) \subseteq U_j$.

Next, let $X' \subseteq X_0 \cup Y_0$ be a maximal set such that there exists a homomorphism φ from $T - V(T_0)$ to D and an extension of ψ covering $\cup_{x \in X'} V(S_x)$ such that the following properties hold.

G1 $\varphi(v) = \hat{\varphi}(v)$ for every $v \in X_0 \cup (\cup_{x \in Y_0} V(S_x))$.

G2 $|\varphi(V(S_x))| \leqslant |H|$ for every $x \in X_0 \cup Y_0$.

G3 $|\varphi^{-1}(j)| = |\hat{\varphi}^{-1}(j)|$ for every $j \in [r]$.

G4 $\psi(x) \in W_{\varphi(x)}$ whenever $x \in X' \cap Y_0$, and $\psi(v) \in U_{\varphi(v)}$ whenever $v \in (\cup_{x \in X'} V(S_x)) \setminus Y_0$.

We remark that this is well defined, as we may take $X' = \emptyset$ and $\varphi = \hat{\varphi}$.

For this maximal X', take (φ, ψ) so that **G1**-**G4** hold, and let $R = \psi(V(T_0) \cup X_0 \cup (\cup_{x \in X'} V(S_x))$. Note that, by **F1**, **G3** and **G4**, we have

$$|U_j \setminus R| \geqslant \alpha \cdot n/r \tag{6}$$

for every $j \in [r]$. We now show that X' includes all of Y_0 and almost all of X_0.

Claim 3.7. $Y_0 \subseteq X'$.

Proof of Claim 3.7. For any $\diamond \in \{+, -\}$ and $x \in N_T^\diamond(V(T_0)) \cap Y_0$, we have

$$|N_G^\diamond(z_x, W_{\varphi(x)}) \setminus R| \overset{(5)}{\geqslant} 2\alpha\mu \cdot n/r - |Y_j^+ \cup Y_j^-| \overset{\mathbf{F3}}{\geqslant} \alpha\mu \cdot n/r.$$

So if $x \in Y_0 \setminus X'$, then, by Proposition 2.9 and (6), ψ can be extended to cover $V(S_x)$ with $\psi(x) \in W_{\varphi(x)} \setminus R$ and $\psi(v) \in U_{\varphi(v)} \setminus R$ whenever $v \in V(S_x) \setminus \{x\}$, contradicting the maximality of X'. So we must have $Y_0 \subseteq X'$. ⊡

Claim 3.8. $|X_0 \setminus X'| \leqslant \mu^4 n$.

Proof of Claim 3.8. For each $m \in \mathbb{N}$, let $g(m)$ denote the number of rooted oriented trees with at most m vertices.

Suppose, for contradiction, that $|X_0 \setminus X'| > \mu^4 n$. Then there is some $j \in [r]$ with $|(X_j^+ \cup X_j^-) \setminus X'| > \mu^4 \cdot n/r$. Therefore, there is some rooted oriented tree S such that, if X_j^S is the set of $x \in (X_j^+ \cup X_j^-) \setminus X'$ for which S_x is isomorphic to S, then $|X_j^S| \geqslant (\mu^4/g(\lfloor 1/\mu^3 \rfloor)) \cdot n/r$.

Choose $x_1 \in X_j^S$ arbitrarily. By Proposition 2.9 and (6), there is a copy of S_{x_1} in G, with each $v \in V(S_{x_1}) \setminus \{x_1\}$ copied to $U_{\varphi(v)} \setminus R$ and x_1 copied to $\psi(X_j^S)$, and let $x_2 \in X_j^S$ be such that $\psi(x_2)$ is the image of x_1 in this copy. Because S_{x_1} and S_{x_2} are isomorphic, we may regard this as a copy of S_{x_2}, and use this copy to extend ψ to cover $V(S_{x_2})$.

Let ρ be an automorphism of $T - V(T_0)$ with $\rho(S_{x_1}) = S_{x_2}$, $\rho(S_{x_2}) = S_{x_1}$, and $\rho(v) = v$ whenever $v \notin V(S_{x_1}) \cup V(S_{x_2})$. Note that $\psi(v) \in U_{\varphi(\rho(v))}$ whenever $v \in (\cup_{x \in X' \cup \{x_2\}} V(S_x)) \setminus Y_0$, and so $\varphi \circ \rho$ is a homomorphism from $T - V(T_0)$ to D also satisfying **G1**-**G4**. So using this extension of ψ and the homomorphism $\varphi \circ \rho$, we may add x_2 to X', a contradiction. ⊡

We now have an embedding of a subtree $T[\psi^{-1}(R)] \subseteq T$ into $G[V']$, where, using Claims 3.7 and 3.8,

$$|V(T) \setminus \psi^{-1}(R)| \leqslant \sum_{x \in (X_0 \cup Y_0) \setminus X'} |S_x| \leqslant \mu n.$$

Recall that we also have $d^{\pm}_G(v, U) \geqslant 4\alpha^2 n \geqslant 3\mu n$ for every $v \in V'$. Therefore, by Corollary 2.3, this embedding can be extended to an embedding of T into G with the vertices of $V(T) \setminus \psi^{-1}(R)$ embedded into U. All that remains now is to prove Claim 3.6.

Proof of Claim 3.6. We will prove that each of the properties **F1**-**F3** fails with probability at most 1/16, and so the claim then follows.

F1: As each ϕ_x was chosen previously by sampling ϕ conditioned on an event which holds with probability at least $(1 - \sqrt{\varepsilon})$, we have that for any $v \in V(T) \setminus V(T_0)$, $j \in [r]$,

$$\mathbb{P}(\phi_{x(v)}(f(v)) = j) \leqslant (1 - \sqrt{\varepsilon})^{-1} \mathbb{P}(\phi(f(v)) = j). \tag{7}$$

Thus, we find that, for any $w \in V(H)$, $j \in [r]$,

$$\begin{aligned}\mathbb{E}(|\hat{\varphi}^{-1}(j) \cap f^{-1}(w)|) &= \sum_{v \in f^{-1}(w)} \mathbb{P}(\phi_{x(v)}(f(v)) = j) \\ &\overset{(7)}{\leqslant} \sum_{v \in f^{-1}(w)} (1 - \sqrt{\varepsilon})^{-1} \mathbb{P}(\phi(f(v)) = j) \overset{(3)}{\leqslant} (1 + \sqrt{\mu}) \cdot \mathbb{E}(\beta(\phi^{-1}(j) \cap \{w\})) \cdot n. \end{aligned} \tag{8}$$

For any $j \in [r]$, we have

$$\mathbb{E}(|\hat{\varphi}^{-1}(j)|) = \sum_{w \in V(H)} \mathbb{E}(|\hat{\varphi}^{-1}(j) \cap f^{-1}(w)|) \overset{(8)}{\leqslant} \mathbb{E}(\beta(\phi^{-1}(j))) \cdot n + \alpha \cdot n/r \overset{\mathbf{E2}}{\leqslant} (1 + \gamma + 2\alpha) \cdot n/r. \tag{9}$$

Therefore, as

$$\sum_{x \in X_0 \cup Y_0} |S_x|^2 \leqslant \sum_{x \in X_0 \cup Y_0} |S_x| \cdot \max_{x \in X_0 \cup Y_0} |S_x| \leqslant \eta n^2, \tag{10}$$

we find that

$$\begin{aligned}\mathbb{P}(|\hat{\varphi}^{-1}(j)| > (1 + \gamma + 3\alpha) \cdot n/r) &\overset{(9)}{\leqslant} \mathbb{P}(||\hat{\varphi}^{-1}(j)| - \mathbb{E}(|\hat{\varphi}^{-1}(j)|)| \geqslant \alpha \cdot n/r) \\ &\overset{\text{Theorem 2.13}}{\leqslant} 2\exp\left(-\frac{2\alpha^2 \cdot n^2/r^2}{\sum_{x \in X_0 \cup Y_0} |S_x|^2}\right) \overset{(10)}{\leqslant} 2\exp\left(-\frac{2\alpha^2}{\eta r^2}\right),\end{aligned}$$

and so the probability that **F1** fails is at most $2r \cdot \exp(-2\alpha^2/\eta r^2) < 1/16$.

F2: We first note that, for any $j \in [r]$ and $\diamond \in \{+, -\}$,

$$\mathbb{E}(|X_j^\diamond|) = \sum_{w \in X^\diamond} \mathbb{E}(|\hat{\varphi}^{-1}(j) \cap f^{-1}(w)|) \overset{(8)}{\leqslant} (1 + \sqrt{\mu}) \cdot \mathbb{E}(\beta(\phi^{-1}(j) \cap X^\diamond)) \cdot n. \tag{11}$$

Also,

$$\mathbb{P}(||X_j^\diamond| - \mathbb{E}(|X_j^\diamond|)| \geqslant \mu^2 \cdot n/r) \overset{\text{Theorem 2.13}}{\leqslant} 2\exp\left(-\frac{2\mu^4 \cdot n^2/r^2}{|X_0 \cap f^{-1}(X^\diamond)|}\right) \leqslant 2\exp\left(-\frac{2\mu^4}{r^2} \cdot n\right).$$

Therefore, with probability at least $1 - 4r \cdot \exp(-(2\mu^4/r^2) \cdot n) > 15/16$, we have

$$||X_j^\diamond| - \mathbb{E}(|X_j^\diamond|)| \leqslant \mu^2 \cdot n/r \qquad \text{for every } j \in [r], \diamond \in \{+, -\}. \tag{12}$$

Thus, it is enough to show that **F2** follows from (12). Indeed, for $j \in J_1$, if $\mathbb{P}(|X_j^+| > 0), \mathbb{P}(|X_j^-| > 0) > 0$, then $d(j_t, j), d(j, j_t) > \mu$, and so for any $x \in X_j^+$ we have

$$|X_j^+| \overset{(12)}{\leqslant} \mathbb{E}(|X_j^+|) + \mu^2 \cdot n/r \overset{(11),\mathbf{E3.1}}{\leqslant} (1 + \sqrt{\mu}) d(j_t, j)(1 + \gamma + \alpha) \cdot n/r + \mu^2 \cdot n/r \overset{(5)}{\leqslant} d^+_G(z_x, U_j),$$

and for any $x \in X_j^-$ we have

$$|X_j^+ \cup X_j^-| \overset{(12)}{\leqslant} \mathbb{E}(|X_j^+ \cup X_j^-|) + 2\mu^2 \cdot n/r \overset{(11),\textbf{E3.1}}{\leqslant} (1+\sqrt{\mu})d(j,j_t)(1+\gamma+\alpha)\cdot n/r + 2\mu^2 \cdot n/r \overset{(5)}{\leqslant} d_G^-(z_x, U_j),$$

and so **F2.1** holds. If instead $|X_j^+| = 0$ with probability 1 or $|X_j^-| = 0$ with probability 1, then the same conclusion holds. Similarly, if $j \in J_2$ then (12) implies **F2.2**.

F3: Note that if $x \in N_T(V(T_0))$ and $j \in [r]$, then, because $\beta(f(x)) \geqslant \mu$,

$$\begin{aligned}\mathbb{P}(\hat{\varphi}(x) = j) = \mathbb{P}(\phi_x(f(x)) = j) &\overset{(7)}{\leqslant} (1-\sqrt{\varepsilon})^{-1}\mathbb{P}(\phi(f(x)) = j) \\ &= (1-\sqrt{\varepsilon})^{-1}\frac{1}{\beta(f(x))} \cdot \mathbb{E}(\beta(\phi^{-1}(j) \cap \{f(x)\})) \overset{\textbf{E2}}{\leqslant} 6/\mu r,\end{aligned}$$

and so, for any $j \in [r]$, we have $\mathbb{E}(|Y_j^+ \cup Y_j^-|) \leqslant 6\mu^2 \cdot n/r$. Therefore, for any $j \in [r]$,

$$\begin{aligned}\mathbb{P}(|Y_j^+ \cup Y_j^-| > \alpha\mu \cdot n/r) &\leqslant \mathbb{P}(||Y_j^+ \cup Y_j^-| - \mathbb{E}(|Y_j^+ \cup Y_j^-|)| > \mu^2 \cdot n/r) \\ &\overset{\text{Theorem 2.13}}{\leqslant} 2\exp\left(-\frac{2\mu^4 \cdot n^2/r^2}{|Y_0|}\right) \leqslant 2\exp\left(-\frac{2\mu^4}{r^2} \cdot n\right),\end{aligned}$$

and so, the probability that **F3** fails is at most $r \cdot \exp(-(2\mu^4/r^2) \cdot n) < 1/16$. ⊡ □

4 Theorem 1.2: embedding the core and attached small trees

In this section, following the proof outline in Section 2.2, we embed T_0 and T_1 for Theorem 1.2, doing so in the form of the following result, Theorem 4.1. (This compares to our work in Section 3 for Theorem 1.1, proving Theorem 3.1.)

Theorem 4.1. *Let $1/n \ll \eta \ll \alpha$. Suppose T is an n-vertex oriented tree with a subtree $T_0 \subseteq T$, such that $|T_0| \leqslant \eta n$ and T is formed from T_0 by attaching to each vertex v of T_0 a tree S_v with $|S_v| \leqslant \eta n$. Then, any $(1+\alpha)n$-vertex tournament contains a copy of T.*

Note that there is no direct maximum degree imposed on T in Theorem 4.1, but as (exactly) one tree is attached to each vertex in T_0 to get T, it follows that $\Delta(T) \leqslant 2\eta n$. As with Theorem 3.1, the proof of Theorem 4.1 is broken into two main parts – in Section 4.1 we allocate the vertices of T to regularity clusters, before embedding the vertices according to this allocation in Section 4.2.

4.1 Allocating vertices for Theorem 4.1

To allocate the vertices of an oriented tree T from Theorem 4.1 to regularity clusters in some ε-regular partition $V_1 \cup \ldots \cup V_r$, we first find a homomorphism from T to a simpler 'caterpillar-like' digraph (see Figure 4). This maps the vertices of the core $T_0 \subset T$ into a small transitive tournament, with the components of $T - V(T_0)$ assigned to an in- or out-leaf from this transitive tournament according to the direction of the edge from T_0 to the component. The number of in- and out-leaves from each transitive tournament vertex is chosen so that the number of vertices of $V(T) \setminus V(T_0)$ mapped onto each one is approximately even.

We ultimately find the 'caterpillar-like' digraph within the reduced digraph R for an ε-regular partition $V_1 \cup \ldots \cup V_r$ (see Section 2.5), and therefore we wish to find the 'caterpillar-like' digraph in any ε-almost tournament R. The method for finding such a 'caterpillar-like' digraph is presented in Lemma 4.2, which is then applied, to a weight function naturally arising from the simplification of T discussed above, to produce a full description of the 'caterpillar-like' digraph in Corollary 4.3. The transitive tournament of the 'caterpillar-like' digraph is found with vertex set $\{j_1, j_2, \ldots, j_s\}$ (where **I2** guarantees it is a transitive tournament), with sets of out-leaves I_i^+ and in-leaves I_i^- of j_i, for each $i \in [s]$. The condition **I3** ensures there are enough in- and out-leaves to allow the approximately even distribution of $V(T) \setminus V(T_0)$ in the simplification of T.

Lemma 4.2. *Let $\varepsilon > 0$ and $\bar{s}, m \in \mathbb{N}$. Let $n_i^+, n_i^- \in \mathbb{N}$, $i \in [\bar{s}]$, be such that $m \leqslant n_i^+ + n_i^- \leqslant 4m$ for each $i \in [\bar{s}]$. Suppose that R is an oriented graph in which, for each $v \in V(R)$,*

$$d_R^+(v) + d_R^-(v) \geqslant |R| - m \geqslant (25 + 1000 \log \bar{s})m + \sum_{i \in [\bar{s}]} (n_i^+ + n_i^-). \tag{13}$$

Then, there is some $s \in [\bar{s}]$ for which there exists $0 = i_0 < i_1 < \ldots < i_{s-1} < i_s = \bar{s}$, and subsets $\{j_\ell\}, I_\ell^+, I_\ell^- \subseteq [r]$ for $\ell \in [s]$, all disjoint, with the following properties.

H1 *$j_{\ell_1} \to_R j_{\ell_2}$ whenever $\ell_1 < \ell_2$.*

H2 *For each $\ell \in [s]$ and $\diamond \in \{+,-\}$, we have $I_\ell^\diamond \subseteq N_R^\diamond(j_\ell)$, and*

$$|I_\ell^\diamond| = \sum_{i=i_{\ell-1}+1}^{i_\ell} n_i^\diamond.$$

Proof. Fix $m \in \mathbb{N}$ and $\varepsilon > 0$. We will show, by strong induction on $\bar{s}$, that the lemma holds for each $\bar{s} \geqslant 1$.

First, suppose $\bar{s} = 1$. It follows from (13) that R is a $(1/25)$-almost tournament with $|R| \geqslant 25m$, and therefore, by Proposition 2.11, there is some $j_1 \in [r]$ such that $d_R^+(j_1), d_R^-(j_1) \geqslant |R|/5 \geqslant 4m$. If we set $I_1^\diamond \subseteq N_R^\diamond(j_1)$ with $|I_1^\diamond| = n_1^\diamond$ for $\diamond \in \{+,-\}$, then all the required properties are satisfied.

Suppose then that $\bar{s} > 1$. It follows from (13) that R is an $(1/25)$-almost tournament with $|R| \geqslant 25m$, and therefore, by Proposition 2.11, there is some $j \in [r]$ such that $d_R^+(j), d_R^-(j) \geqslant |R|/5$. Now, by (13), we have $d_R^+(j) + d_R^-(j) \geqslant \sum_{i \in [\bar{s}]}(n^+(i) + n^-(i))$. Therefore, at least one of $\sum_{i \in [\bar{s}]} n^+(i) \leqslant d_R^+(j)$ or $\sum_{i \in [\bar{s}]} n^-(i) \leqslant d_R^-(j)$ holds. If both inequalities hold, then the desired result follows by taking $s = 1$, $j_1 = j$, and $I_1^\diamond \subseteq N_R^\diamond(j)$ with $|I_1^\diamond| = \sum_{i \in [\bar{s}]} n_i^\diamond$ for each $\diamond \in \{+,-\}$. Otherwise, by directional duality, we may assume that $\sum_{i \in [\bar{s}]} n^+(i) \leqslant d_R^+(j)$ and $\sum_{i \in [\bar{s}]} n^-(i) > d_R^-(j)$.

Then, let $s' \in [\bar{s} - 1]$ be maximal such that

$$\sum_{i \in [s']} n^-(i) \leqslant d_R^-(j).$$

Note that, as

$$d_R^-(j) \geqslant \frac{|R|}{5} \overset{(13)}{\geqslant} \frac{1}{5} \sum_{i \in [\bar{s}]} (n_i^+ + n_i^-) \geqslant \frac{\bar{s}m}{5}$$

and $n_i^- \leqslant (n_i^+ + n_i^-) \leqslant 4m$ for each $i \in [\bar{s}]$, we have $s' \geqslant \bar{s}/20$. Furthermore, by the maximality of s', we have

$$\sum_{i \in [s']} n^-(i) \geqslant d_R^-(j) - 4m. \tag{14}$$

Let $j_0 = 0$ and $j_1 = s'$. Let $I_1^- \subseteq d_R^-(j)$ have size $\sum_{i \in [s']} n^-(i)$. Using that $d_R^+(j) \geqslant \sum_{i \in [\bar{s}]} n^+(i)$, let $I_1^+ \subseteq d_R^+(j)$ have size $\sum_{i \in [s']} n^+(i)$ and let $I = N_R^+(j) \setminus I_1^+$.

Now, $\bar{s} - s' \leqslant \bar{s} - \bar{s}/20 = 19\bar{s}/20$ so that $1000 \log(\bar{s} - s') \leqslant -5 + 1000 \log \bar{s}$, and hence

$$\begin{aligned}
|I| = d_R^+(j) - |I_I^+| &\overset{(14)}{\geqslant} d_R^+(j) - |I_1^+| + d_R^-(j) - |I_1^-| - 4m \\
&\overset{(13)}{\geqslant} (25 + 1000 \log \bar{s})m + \sum_{i \in [\bar{s}]} (n_i^+ + n_i^-) - |I_1^+| - |I_1^-| - 4m \\
&= (21 + 1000 \log \bar{s})m + \sum_{i \in [\bar{s}] \setminus [s']} (n_i^+ + n_i^-) \\
&\geqslant (26 + 1000 \log(\bar{s} - s'))m + \sum_{i \in [\bar{s}] \setminus [s']} (n_i^+ + n_i^-).
\end{aligned}$$

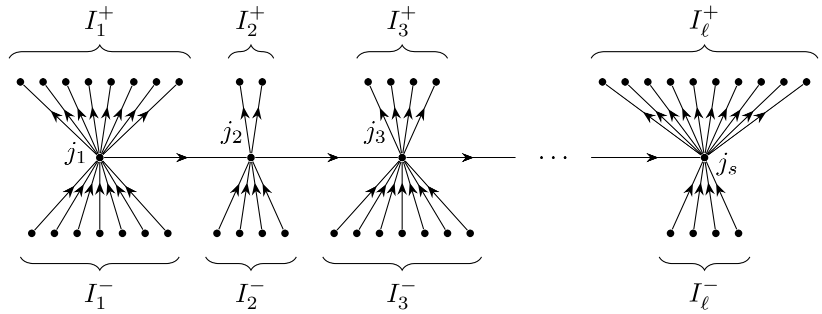

Figure 4: A 'caterpillar-like' digraph, as appearing in Lemma 4.2 and Corollary 4.3. While the other edges are omitted for legibility, $j_{\ell_1} \to j_{\ell_2} \to \ldots \to j_s$ is the underlying directed path of a transitive tournament.

Let $R' = R[I]$, and note that, for each $j \in V(R')$, by (13) we have $d^+_{R'}(j) + d^-_{R'}(j) \geqslant |R'| - m = |I| - m$, so that, in combination with the above calculation,

$$d^+_{R'}(v) + d^-_{R'}(v) \geqslant |R'| - m \geqslant (25 + 1000\log(\bar{s} - s'))m + \sum_{i\in[\bar{s}]\setminus[s']} (n_i^+ + n_i^-),$$

for each $v \in V(R')$. Therefore, by the inductive hypothesis for $\bar{s} - s'$, there is (with relabelling) some $s \in [\bar{s}]$ for which there exists $s' = i_1 < i_2 < \ldots < i_s = \bar{s}$ and subsets $\{j_\ell\}, I^+_\ell, I^-_\ell \subseteq V(R') = N^+_R(j) \setminus I^+_1$ for $\ell \in [s] \setminus [1]$, all disjoint, such that $j_{\ell_1} \to_R j_{\ell_2}$ whenever $\ell_1 < \ell_2$, and, for each $\ell \in [s]$ and $\diamond \in \{+, -\}$, we have $|I^\diamond_\ell| = \sum_{i=i_{\ell-1}+1}^{i_\ell} n_i^\diamond$. Thus, the required properties are satisfied, completing the proof. □

Corollary 4.3. *Let $1/n \ll \varepsilon, \eta, 1/r \ll \alpha \leqslant 1$. Suppose T is an n-vertex oriented tree with a subtree $T_0 \subseteq T$, such that $|T_0| \leqslant \eta n$ and T is formed from T_0 by attaching to each vertex v of T_0 a tree S^+_v in which v only has out-neighbours and a tree S^-_v in which v only has in-neighbours, so that $|S^+_v|, |S^-_v| \leqslant \eta n$. Let R be a ε-almost tournament with vertex set $[r]$.*

Then, there is some $s \leqslant \alpha/100\varepsilon$ for which there exists a partition $V(T_0) = X_1 \cup \ldots \cup X_s$ and subsets $\{j_\ell\}, I^+_\ell, I^-_\ell \subseteq [r]$ for $\ell \in [s]$, all disjoint, with the following properties.

- **I1** *There are no edges of T_0 directed from X_i to X_j with $i, j \in [s]$ and $i > j$.*
- **I2** *$j_{\ell_1} \to_R j_{\ell_2}$ whenever $\ell_1 < \ell_2$.*
- **I3** *For each $\ell \in [s]$ and $\diamond \in \{+, -\}$, we have $I^\diamond_\ell \subseteq N^\diamond_R(j_\ell)$, and*

$$|I^\diamond_\ell| \geqslant \frac{r}{(1 + \alpha/4)n} \cdot \sum_{v\in X_\ell} |S^\diamond_v|.$$

Proof. Pick $c \geqslant 2\eta$ such that $\varepsilon, 1/r \ll c \ll \alpha$. Let $\bar{m} = cn$. Let $n_0 = |T_0|$ and let $v_1, \ldots, v_{n_0}$ order $V(T_0)$ such that $i < j$ whenever $v_i \to_T v_j$. Let $\bar{s}$ be the largest integer for which there are integers $0 = k_0 < k_1 < \ldots < k_{\bar{s}} \leqslant n_0$ such that $\bar{m} \leqslant \sum_{k=k_{\ell-1}+1}^{k_\ell} (|S^+_{v_k}| + |S^-_{v_k}|) \leqslant 2\bar{m}$ for each $\ell \in [\bar{s}]$. Now, as $|S^+_{v_k}| + |S^-_{v_k}| \leqslant \bar{m}$ for each $k \in [n_0]$, we must have by this maximality that $\sum_{k=k_{\bar{s}}+1}^{n_0} |S^+_{v_k}| + |S^-_{v_k}| < \bar{m}$, and therefore, as T has n vertices, we have that $\bar{s} \geqslant n/3\bar{m} = 1/3c \geqslant 1$. Furthermore, setting $W_\ell = \{v_{k_{\ell-1}+1}, \ldots, v_{k_\ell}\}$ for each $\ell \in [\bar{s} - 1]$ and $W_{\bar{s}} = \{v_{k_{\bar{s}-1}+1}, \ldots, v_{n_0}\}$, we have, for each $\ell \in [\bar{s}]$, that

$$\bar{m} \leqslant \sum_{v\in W_\ell} (|S^+_v| + |S^-_v|) \leqslant 3\bar{m}.$$

Finally, note that

$$\frac{n}{3\bar{m}} \leqslant \bar{s} \leqslant \frac{2n}{\bar{m}}. \tag{15}$$

Now, for each $i \in [\bar{s}]$, let

$$n_i^\diamond = \left\lceil \frac{r}{n(1+\alpha/4)} \sum_{v \in W_i} |S_v^\diamond| \right\rceil.$$

Let $m = r\bar{m}/n(1+\alpha/4)$, so that $cr/2 \leqslant m \leqslant cr$ and, for each $i \in [\bar{s}]$, $m \leqslant n_i^+ + n_i^- \leqslant 4m$. From (15), we have $r/4m \leqslant \bar{s} \leqslant 2r/m$. Therefore, as $\bar{s} \geqslant 1/3c$ and $1/r \ll c \ll \alpha$, we have

$$2\bar{s} + (25 + 1000\log \bar{s})m \leqslant \frac{4r}{m} + \left(\frac{10^5 \log \bar{s}}{\bar{s}}\right) r \leqslant \frac{8}{c} + \frac{\alpha r}{16} \leqslant \frac{\alpha r}{8}. \tag{16}$$

Note that

$$\sum_{i \in [\bar{s}]} (n_i^+ + n_i^-) \leqslant 2\bar{s} + \frac{r}{(1+\alpha/4)n} \cdot \sum_{i \in [\bar{s}]} \sum_{v \in W_i} (|S_v^+| + |S_v^-|) \leqslant 2\bar{s} + \frac{r(1+\eta)}{(1+\alpha/4)} \leqslant 2\bar{s} + (1-\alpha/8)r,$$

as $\eta, 1/r \ll \alpha$, so that, by (16), we have

$$r \geqslant (25 + 1000\log \bar{s})m + \sum_{i \in [\bar{s}]} (n_i^+ + n_i^-).$$

Finally, we have $m \geqslant cr/2 \geqslant \varepsilon r$, so that, as R is an ε-almost tournament, for each $v \in V(R)$, we have $d_R^+(v) + d_R^-(v) \geqslant |R| - \varepsilon|R| \geqslant |R| - m$.

Thus, by Lemma 4.2, there is some $s \in [\bar{s}]$ for which there exists $0 = i_0 < i_1 < \ldots < i_{s-1} < i_s = \bar{s}$, and subsets $\{j_\ell\}, I_\ell^+, I_\ell^- \subseteq [r]$ for $\ell \in [s]$, all disjoint, such that **H1** and **H2** hold. Letting $X_\ell = \cup_{i=j_{\ell-1}}^{j_\ell} W_i$ for each $\ell \in [s]$ then gives the required partition. □

4.2 Embedding vertices for Theorem 4.1

We can now prove Theorem 4.1 using Corollary 4.3. We first prove Lemma 4.4 which can embed the vertices of T that have been mapped to a single vertex of the small transitive tournament and its in- and out-neighbours in the 'caterpillar-like' digraph (here corresponding to V_0, with in-neighbours corresponding to $V_1^+, \ldots, V_k^+$ and out-neighbours corresponding to $V_1^-, \ldots, V_\ell^-$), before using this repeatedly for each vertex of the small transitive tournament produced by Corollary 4.3 to prove Theorem 4.1.

Lemma 4.4. *Fix $\alpha \geqslant \beta > 0$, $\mu > 0$ and let $1/m \ll \eta \ll 1/r \ll \varepsilon \ll \gamma \ll \mu, \beta$. Let G be a tournament. Suppose, for some $k, \ell \leqslant r$, there are disjoint subsets $V_0, V_1^+, \ldots, V_k^+, V_1^-, \ldots, V_\ell^-$ of $V(G)$, all of size $(1+\alpha)m$, such that (V_0, V_i^+) is an ε-regular pair of density at least μ for $i \in [k]$, and (V_i^-, V_0) is an ε-regular pair of density at least μ for $i \in [\ell]$.*

Suppose T is an oriented tree with a subtree $T_0 \subseteq T$, such that $|T_0| \leqslant \eta m$, and T is formed from T_0 by attaching to each vertex v of T_0 trees S_v^+, S_v^- with $d^-_{S_v^+}(v) = 0$, $d^+_{S_v^-}(v) = 0$, and $|S_v^+|, |S_v^-| \leqslant \eta m$.

Let $W \subseteq V_0$ be a set with $|W| \geqslant \gamma m$, and let $U_i^+ \subseteq V_i^+$, $i \in [k]$, and $U_i^- \subseteq V_i^-$, $i \in [\ell]$ be sets such that $\sum_{i \in [k]} |U_i^+| \geqslant \sum_{v \in V(T_0)} |S_v^+| + k\beta m$ and $\sum_{i \in [\ell]} |U_i^-| \geqslant \sum_{v \in V(T_0)} |S_v^-| + \ell\beta m$.

Then, there is a copy of T in G, with T_0 copied to W and $T - V(T_0)$ copied to $U_1^+ \cup \ldots \cup U_k^+ \cup U_1^- \cup \ldots \cup U_\ell^-$.

Proof. For the smallest possible p, take a partition $V(T_0) = X_1 \cup \ldots \cup X_p$ such that, for each $j \in [p]$, $T_0[X_1 \cup \ldots \cup X_j]$ is a tree, $\sum_{v \in X_j} |S_v^+| \leqslant k\beta\mu m/4$, and $\sum_{v \in X_j} |S_v^-| \leqslant \ell\beta\mu m/4$. This is possible for $p = |T_0|$, so a smallest such p will exist. We in fact claim that $p \leqslant 32/\beta\mu$. Indeed, for this smallest possible p, take a partition that minimises $\sum_{j \in [p]} j|X_j|$. Suppose there is some $j < p$ for which both $\sum_{v \in X_j} |S_v^+| \leqslant k\beta\mu m/8$ and $\sum_{v \in X_j} |S_v^-| \leqslant \ell\beta\mu m/8$. Let $x \in X_{j'+1}$ be such that $T_0[X_1 \cup \ldots \cup X_{j'} \cup \{x\}]$ is a tree. Then moving x from $X_{j'+1}$ to $X_{j'}$ produces a partition which contradicts the minimality of $\sum_{j \in [p]} j|X_j|$. Thus we have

$$p \leqslant 1 + \frac{\sum_{v \in V(T_0)} |S_v^+|}{(k\beta\mu m/8)} + \frac{\sum_{v \in V(T_0)} |S_v^-|}{(\ell\beta\mu m/8)} \leqslant 1 + \frac{8\sum_{i \in [k]} |U_i^+|}{k\beta\mu m} + \frac{8\sum_{i \in [\ell]} |U_i^-|}{\ell\beta\mu m} \leqslant 32/\beta\mu.$$

Now find vertex sets $W_1 \subseteq W_2 \subseteq \ldots \subseteq W_p = W$ such that $|W_j| \geqslant |W_{j+1}|/8$ for each $j \in [p-1]$, and $d^\pm(w, W_{j+1}) \geqslant |W_{j+1}|/8$ for each $j \in [p-1]$, $w \in W_j$. This is possible by starting with W_p and iteratively

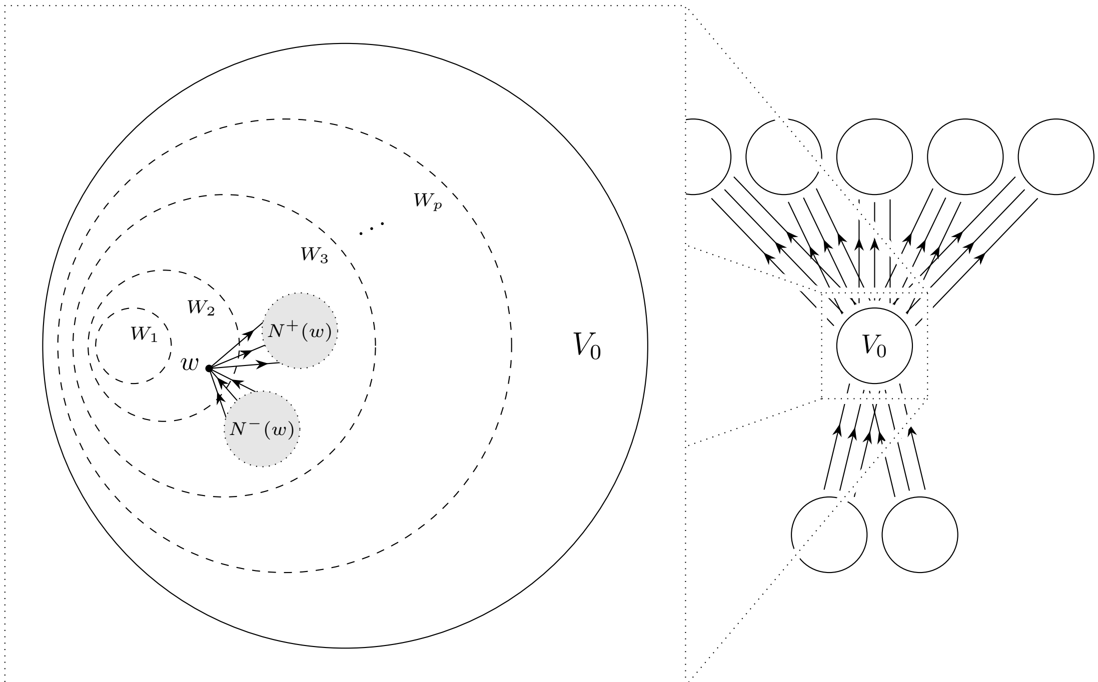

Figure 5: The sets $W_1 \subseteq \ldots \subseteq W_p$ in the proof of Lemma 4.4. The sets are chosen so that each vertex $w \in W_j$ has sufficiently many in- and out-neighbours in W_{j+1}.

using that fact that at most $|W_{j+1}|/4$ vertices w of W_{j+1} have $d^+(w, W_{j+1}) \leqslant |W_{j+1}|/8$, and at most $|W_{j+1}|/4$ vertices w of W_{j+1} have $d^-(w, W_{j+1}) \leqslant |W_{j+1}|/8$.

We will now embed T in p stages as follows. At stage j, suppose we have already embedded $T[\cup_{j'=1}^{j-1} \cup_{v \in X_{j'}} (V(S_v^+) \cup V(S_v^-))]$. For each $v \in X_1 \cup \ldots \cup X_{j-1}$ in turn, consider the forest F_v^+ consisting of trees of $T_0[X_j]$ attached to v by out-neighbours of v, and suppose v has already been copied to some $w \in W_{j-1}$ (for the case $j = 1$, regard all of $T_0[X_j]$ as components attached to a single auxiliary vertex v by out-neighbours, where v has already been copied to an auxiliary vertex w satisfying $W_1 \subseteq N_G^+(w)$). Let $Z_{j,v}^+$ be the set of unoccupied out-neighbours of w in W_j which each have at least $3k\beta\mu m/4$ unoccupied out-neighbours in $\cup_{i\in[k]} U_i^+$ as well as $3\ell\beta\mu m/4$ unoccupied in-neighbours in $\cup_{i\in[\ell]} U_i^-$. Because there are always at least $k\beta m$ unoccupied vertices in $\cup_{i\in[k]} U_i^+$ and $\ell\beta m$ unoccupied vertices in $\cup_{i\in[\ell]} U_i^-$, Proposition 2.8 implies $|Z_{j,v}^+| \geqslant |N^+(w, W_j)| - |T_0| - 2\varepsilon m \geqslant |W_1|/8 - 3\varepsilon m \geqslant |W|/8^{p+1} \geqslant 3\eta m$. Therefore, by Theorem 2.2, there is a copy of F_v^+ in $Z_{j,v}^+$. Then, for each $v' \in V(F_v)$, if v' has now been copied to w', find a copy of $S_{v'}^+ - v'$ in the unoccupied vertices of $N^+(w', \cup_{i\in[k]} U_i^+)$. Because $w' \in Z_{j,v}^+$, and only at most $\sum_{v''\in X_j, v''\neq v'} |S_{v''}^+|$ additional vertices of $N^+(w', \cup_{i\in[k]} U_i^+)$ may become occupied since choosing $Z_{j,v}^+$, at least $3k\beta\mu m/4 - \sum_{v''\in X_j, v''\neq v'} |S_{v''}^+| \geqslant 3|S_{v'}^+|$ vertices of $N^+(w', \cup_{i\in[k]} U_i^+)$ remain unoccupied, allowing the copy of $S_{v'}^+ - v'$ to be found using Theorem 2.2. Similarly, find a copy of $S_{v'}^- - v'$ in the unoccupied vertices of $N^-(w', \cup_{i\in[\ell]} U_i^-)$. We then do the same for the forest F_v^- consisting of trees of $T_0[X_j]$ attached to $X_1 \cup \ldots \cup X_{j-1}$ by in-neighbours. Performing this process for each $v \in X_1 \cup \ldots \cup X_{j-1}$ completes stage j of the embedding procedure. Upon the completion of stage p, we obtain a copy of T in G, with T_0 copied to W and $T - |T_0|$ copied to $U_1^+ \cup \ldots \cup U_k^+ \cup U_1^- \cup \ldots \cup U_\ell^-$. $\square$

We now combine Lemma 4.2 and Lemma 4.4 to prove Theorem 4.1.

Proof of Theorem 4.1. Set $\beta = \alpha/4$, $\mu = 1/2$, and introduce constants ε, r_1, r_2 such that $\eta \ll 1/r_2 \ll 1/r_1 \ll \varepsilon \ll \beta$. Let G be a $(1+\alpha)n$-vertex tournament. By Corollary 2.7, there is a subtournament $G' \subseteq G$ with $|G'| \geqslant (1+3\beta)n$, and an ε-regular partition $V(G') = V_1 \cup \ldots \cup V_r$ with $r_1 \leqslant r \leqslant r_2$. Let R be a $\sqrt{\varepsilon}$-almost tournament with vertex set $[r]$, such that (V_i, V_j) is an ε-regular pair of density at least μ whenever $i \to_R j$.

Fix disjoint subsets $U_j, W_j \subseteq V_j$ for each $j \in [r]$ with $|U_j| = (1 + 2\beta) \cdot n/r$ and $|W_j| = \beta \cdot n/r$.

For each $v \in V(T_0)$, let $S_v^+ \subseteq S_v$ be the subtree of S_v induced by the vertices whose path from v begins with an out-edge, and let $S_v^- \subseteq S_v$ be the subtree of S_v induced by the vertices whose path from v begins with an in-edge. Note that we have $|S_v^+|, |S_v^-| \leqslant \eta n$ for every $v \in V(T_0)$. By Corollary 4.3, there is some $s \leqslant \alpha/100\varepsilon$ for which there exists a partition $V(T_0) = X_1 \cup \ldots \cup X_s$ and subsets $\{j_\ell\}, I_\ell^+, I_\ell^- \subseteq [r]$ for $\ell \in [s]$, all disjoint, satisfying properties **I1-I3**. In particular, for each $\ell \in [s]$ and $\diamond \in \{+, -\}$, we have

$$\sum_{v \in X_\ell} |S_v^\diamond| + |I_\ell^\diamond|\beta \cdot n/r \overset{\textbf{I3}}{\leqslant} |I_\ell^\diamond|(1 + 2\beta) \cdot n/r = \sum_{j \in I_\ell^\diamond} |U_j^\diamond| \tag{17}$$

Set $\gamma = \beta\mu/8r$. Using **I2**, Proposition 2.8, and $s \leqslant \alpha/100\varepsilon$, for each $\ell \in [s]$ at most $s\varepsilon(1 + \alpha) \cdot n/r \leqslant 2\gamma n$ vertices w of W_{j_ℓ} have either some $\ell' > \ell$ for which $d^+(w, W_{j_{\ell'}}) \leqslant 4\gamma n$, or some $\ell' < \ell$ for which $d^-(w, W_{j_{\ell'}}) \leqslant 4\gamma n$. Therefore, we may take subsets $W'_{j_\ell} \subseteq W_{j_\ell}$ for $\ell \in [s]$ such that, $d^+(w, W'_{j_{\ell_2}}) \geqslant 2\gamma n$ whenever $\ell_2 > \ell_1$ and $w \in W'_{j_\ell}$, and $d^-(w, W'_{j_{\ell_2}}) \geqslant 2\gamma n$ whenever $\ell_2 < \ell_1$ and $w \in W'_{j_\ell}$.

Now obtain a partition $V(T_0) = Y_1 \cup \ldots \cup Y_\tau$ such that

J1 For each $t \in [\tau]$, $T[Y_t]$ is a connected component of $T_0[X_\ell]$ for some $\ell \in [s]$.

J2 For each $t \in [\tau]$, $T_0[Y_1 \cup \ldots \cup Y_t]$ is a tree.

We will now embed T into G so that X_ℓ is copied to W'_{j_ℓ} for each $\ell \in [s]$, and $\cup_{v \in X_\ell} V(S_v^\diamond)$ is copied to $\cup_{j \in I_\ell^\diamond} U_j^\diamond$ for each $\ell \in [s]$ and $\diamond \in \{+, -\}$. The embedding is given in τ stages as follows. Let R_0 be the empty graph. Suppose after stage $t-1$, we have embedded $T[\cup_{v \in Y_1 \cup \ldots \cup Y_{t-1}} V(S_v)]$ to get R_{t-1}. Let $\ell \in [s]$ be such that $Y_t \subseteq X_\ell$. If $t = 1$, set $A_t = W'_{j_\ell}$. Otherwise, if $t > 1$, let y_t be the unique vertex of $Y_1 \cup \ldots \cup Y_{t-1}$ with a neighbour in Y_t, let $\diamond \in \{+, -\}$ be such that the neighbour in Y_t is a $\diamond$-neighbour, let z_t be the image of y_t in R_{t-1}, and set $A_t = N^\diamond(z_t, W'_{j_\ell}) \setminus V(R_{t-1})$. Note that in both cases we find $|A_t| \geqslant \gamma n$. Also, we find for $\diamond \in \{+, -\}$ that

$$\sum_{j \in I_\ell^\diamond} |U_j^\diamond \setminus V(R_{t-1})| \overset{(17)}{\geqslant} \sum_{v \in Y_t} |S_v^\diamond| + |I_\ell^\diamond|\beta n/r.$$

Therefore, by **H2** and Lemma 4.4, there is a copy of $T[Y_t \cup (\cup_{v \in Y_t} S_v)]$ in G with Y_t copied to $A_t \subseteq W'_{j_\ell}$, and $(\cup_{v \in Y_t} V(S_v^\diamond)) \setminus Y_t$ copied to $(\cup_{j \in I_\ell^\diamond} U_j^\diamond) \setminus V(R_{t-1})$ for $\diamond \in \{+, -\}$. Thus we obtain a copy of T after stage τ. □

5 Proof of Theorem 1.1 and Theorem 1.2

Recall the decomposition of our tree T from Section 2.2 as $T_0 \subseteq T_1 \subseteq T_2 \subseteq T_3 \subseteq T_4 = T$. In Sections 3 and 4 respectively, we showed how to embed T_0 and extend this to T_1 for both Theorems 1.1 and 1.2. In this section, we will show how a copy of T_1 can be extended to a copy of T, completing the proof of both theorems. As noted in the proof outline, the main challenge here is to embed the vertices in $V(T_3) \setminus V(T_2)$, where these vertices form paths with constant length between vertices in T_2. Indeed, firstly, $T_2 - V(T_1)$ is a forest of constant-sized components not directly connected to T_1 (see **A3**), which can be embedded greedily using, for example, Theorem 2.2. Secondly, to reach T_4 from T_3 we add small tree components on to T_3, which is already connected. This can be done by reserving a small random subset of vertices U (using Proposition 2.14) and carrying out the rest of the embedding in the vertices with sufficient out- and in-degree to U. Such an embedding can then be completed greedily, giving an embedding of $T_4 = T$.

Thus, most of this section will be dedicated to showing how we can extend a copy of T_2 to a copy of T_3 (using a method effective for both Theorems 1.1 and 1.2). Recall that T_3 is obtained from T_2 by attaching paths of fixed length by their endpoints (see **A4**), but such that the total number of vertices contained in such paths is only a small proportion of the resulting tree (see **A5**). Thus, with a copy of T_2 already found, we will often wish to find paths of a fixed length between certain attachment points. By ensuring these attachment points have plenty of out- and in-neighbours, we need only to be able to connect linear-sized sets with paths of fixed but small length, while avoiding some small set of vertices already used in some paths. As we will see, paths with changes of direction are comparatively easy to find, so we only consider whether we can find such paths so that they are directed paths. We will call tournaments with this connection property *well-connected*, as follows.

Definition 5.1. *We say a tournament G is* (a, b, ℓ)-well-connected *if, for every $A_1, A_2 \subseteq V(G)$ with $|A_1|, |A_2| \geqslant a$ and $B \subseteq V(G)$ with $|B| \leqslant b$, there is a directed path in $V(G) \setminus B$ from A_1 to A_2 with length ℓ.*

In Lemma 5.7, we will see that both of our main theorems hold if the tournament G is well-connected. Of course, not every tournament is well-connected, but, in Lemma 5.5 we will see that any tournament that is not well-connected contains a bipartition of most of its vertices, so that all the relevant edges are directed in the same direction across the bipartition. Through the repeated application of Lemma 5.5, we can then decompose the vertices of any tournament as $V(G) = B \cup W_1 \cup \ldots \cup W_r$, so that B is small, all possible edges are directed from W_i to W_j for $1 \leqslant i < j \leqslant r$, and each $G[W_i]$ is either small or well-connected (see Lemma 5.6). We then assign the vertices of T to the sets $W_1, \ldots, W_r$, so that any edge of T assigned between some W_i and W_j with $i < j$ is to be embedded as directed from W_i into W_j. Thus, we can embed the vertices of T assigned to W_i into $G[W_i]$ independently for each $i \in [r]$, while knowing the other edges of T can then be embedded. As noted in the proof sketch, this is a streamlined version of techniques by Kühn, Mycroft and Osthus [12, 13]. In [12, 13], a notion of robust out-expansion is used, from which our well-connected property can be derived. As we do not need any other results of robust out-expansion (most notably, we do not use a Hamilton, or almost-spanning, cycle in the reduced digraph), we use the well-connected property directly. This allows the decomposition of [13, Lemma 5.2] to be simplified to find bipartitions with all the edges directed from one side to another, rather than just most of the edges.

In Section 5.1, we will prove a number of results on well-connected tournaments, including the tournament decomposition discussed above. Then, in Section 5.2, after showing our main results hold for well-connected tournaments (i.e., Lemma 5.7), we prove both Theorems 1.1 and 1.2.

5.1 Well-connected tournaments

We start by proving two simple properties of well-connected tournaments in Lemma 5.2. The first is that removing a small number of vertices from a well-connected tournament maintains some (potentially slightly weaker) connection property. The second shows that (a, b, ℓ)-well-connected tournaments robustly contain paths of length ℓ, regardless of the desired orientation of the paths' edges. While Definition 5.1 only refers to directed paths, Thomason [16] showed that a path with at least one change of direction can be found between two sufficiently large subsets of any tournament, covering all other cases.

Lemma 5.2. *Let $a, b, \ell \geqslant 0$, and suppose G is a (a, b, ℓ)-well-connected tournament.*

i) *If $C \subseteq V(G)$ has size $c \leqslant b$, then $G - C$ is $(a, b - c, \ell)$-well-connected.*

ii) *Suppose P is an oriented path of length ℓ, and $A_1, A_2, B \subseteq V(G)$ satisfy $|A_1|, |A_2| \geqslant a$, $|B| \leqslant b$. If $a \geqslant b + \ell + 3$, then there is a copy of P in $G - B$, with its first vertex in A_1 and its last vertex in A_2.*

Proof. First, fix a subset $C \subseteq V(G)$ with size c. Then, if $A_1, A_2, B \subseteq V(G)$ satisfy $|A_1|, |A_2| \geqslant a$ and $|B| \leqslant b - c$, then, because G is (a, b, ℓ)-well-connected, there is a directed path in $V(G) \setminus (B \cup C)$ from A_1 to A_2 with length ℓ. Therefore, $G - C$ is $(a, b - c, \ell)$-well-connected and **i)** holds.

Next, suppose P is an oriented path of length ℓ, and $A_1, A_2, B \subseteq V(G)$ satisfy $|A_1|, |A_2| \geqslant a$, $|B| \leqslant b$. If P is a directed path, then, because G is (a, b, ℓ)-well-connected there is a copy of P in $G - B$ with first vertex in A_1 and last vertex in A_2. On the other hand, if P has a change of direction, then by one of [16, Theorem 3], [16, Theorem 4], or [16, Theorem 5], there is a copy of P in $G[(A_1 \cup A_2) \setminus B]$, with first vertex in A_1 and last vertex in A_2. Therefore, **ii)** holds. □

We will need to set aside a random subset of vertices to use to attach paths to T_2 to obtain T_3. We need therefore to show that random subsets of well-connected tournaments can be used to find connecting paths of this sort. To do this, we use *median orders*. Given a tournament G, an ordering $v_1, \ldots, v_n$ of $V(G)$ is a median order if it maximises the number of pairs $i < j$ with $v_i v_j \in E(G)$. We use median orders only to recall the following lemma from [2] and apply it to prove Lemma 5.4. For further useful properties of median orders applicable to embedding trees in tournaments see [10, 3, 2].

Lemma 5.3 ([2, Lemma 2.9])**.** *Suppose G is a tournament with a median order $v_1, \ldots, v_n$. Then, for any $1 \leqslant i < j \leqslant n$ with $j - i \geqslant 7$, and $A \subset V(G) \setminus \{v_i, v_j\}$ with $|A| \leqslant (j - i - 7)/6$, there is a directed v_i, v_j-path in $G - A$ with length 3.*

We are now ready to state and prove our lemma, showing that, with high probability, a random subset of vertices in a well-connected tournament induces a well-connected tournament, as follows.

Lemma 5.4. *Let $1/n \ll \eta \ll 1/\ell \ll \varepsilon \ll p$. Suppose G is a $(\varepsilon n, \eta n, \ell)$-well-connected tournament with $|G| \leqslant 3n$, and that $U \subseteq V(G)$ is a random subset with vertices included uniformly at random with probability p. Then, with high probability, $G[U]$ is $(6\varepsilon n, \eta^2 n, \ell+6)$-well-connected.*

Proof. Let $v_1, \dots, v_m$ be a median order for G. Let W_1 and W_2 respectively denote the first and last εn vertices of the median order. Let V' be the middle $m - 4\varepsilon n$ vertices of the median order.

It is enough to show that, with high probability, for every $v, w \in V'$, there are at least $2\eta^2 n$ internally vertex-disjoint directed v,w-paths with length $\ell + 6$ and with all internal vertices in U. Indeed, then for any $A_1, A_2 \subseteq U$ with $|A_1|, |A_2| \geqslant 6\varepsilon n$ and $B \subseteq U$ with $|B| \leqslant \eta^2 n$, there is some $v \in (V' \cap A_1) \setminus B$ and $w \in (V' \cap A_2) \setminus (B \cup \{v\})$, and hence at least $2\eta^2 n$ internally vertex-disjoint directed paths from $A_1 \setminus B$ to $A_2 \setminus B$ in $G[U]$ with length $\ell + 6$. Of these paths, at most $\eta^2 n$ contain some internal vertex in B, and so there is some directed path in $U \setminus B$ from A_1 to A_2 of length $\ell + 6$, thus demonstrating $G[U]$ is $(6\varepsilon n, \eta^2 n, \ell + 6)$-well-connected.

Fix $v, w \in V'$. Because G is $(\varepsilon n, \eta n, \ell)$-well-connected, and $|W_1|, |W_2| \geqslant \varepsilon n$, we can greedily find at least $\eta n / 2\ell$ disjoint directed paths in $V(G) \setminus \{v, w\}$ from W_2 to W_1 with length ℓ. Using Lemma 5.3, we can greedily and disjointly connect v to the first vertex of each path by a directed path of length 3, while avoiding all other vertices used so far. Indeed, at least εn vertices in the median order lie between v and the first vertex of each path, while the total number of vertices to be avoided each time is at most $\eta n \leqslant (\varepsilon n - 7)/6$. Similarly, we can also disjointly connect the last vertex of each path to w by a directed path of length 3, also avoiding any vertex used previously. Therefore, we have at least $\eta n / 2\ell$ internally disjoint directed paths in $V(G)$ from v to w, each with length $\ell + 6$.

Let $X_{v,w}$ be the number of these directed v,w-paths which additionally have all internal vertices in U, and note that $X_{v,w}$ is a binomial variable with $\mathbb{E}X_{v,w} \geqslant p^{\ell+5} \eta n / 2\ell > 3\eta^2 n$. From Lemma 2.12, we have

$$\mathbb{P}(X_{v,w} \leqslant 2\eta^2 n) \leqslant \mathbb{P}(|X_{v,w} - \mathbb{E}X_{v,w}| \geqslant \mathbb{E}X_{v,w}/3) \leqslant 2\exp\left(-\mathbb{E}X_{v,w}/27\right) \leqslant 2\exp\left(-\eta^2 n/9\right).$$

Thus, the probability that the desired property fails is at most $18n^2 \exp\left(-\eta^2 n/9\right)$, and so, as $1/n \ll \eta$, the conclusion of the lemma holds with high probability. $\square$

Next we will show that, if a tournament is not well-connected, then, except for a small subset of vertices, we may partition the vertices in two so that all the edges between the parts are directed into the same part.

Lemma 5.5. *Let $\varepsilon > 0$, $\ell \in \mathbb{N}$, and $\eta \ll \varepsilon, 1/\ell$. Suppose G is a tournament with $|G| \leqslant 3n$ that is not $(\varepsilon n, \eta n, \ell)$-well-connected. Then, there is a partition $V(G) = W_1 \cup W_2 \cup B$ so that $|W_1|, |W_2| \geqslant \varepsilon n/2$, $|B| \leqslant 4\ell^{-1} n$, and $x \to y$ for every $x \in W_1$, $y \in W_2$.*

Proof. Using that G is not $(\varepsilon n, \eta n, \ell)$-well-connected, let $A_1, A_2, B_0 \subseteq V(G)$ be sets such that $|A_1|, |A_2| \geqslant \varepsilon n$, $|B_0| \leqslant \eta n$, and there is no directed path in $V(G) \setminus B$ from A_1 to A_2 with length ℓ. Construct a chain of subsets $U_0 \subseteq U_1 \subseteq \ldots \subseteq U_\ell$ as follows. Let $U_0 = A_1 \setminus B_0$, and, for $i \in [\ell]$, let U_i be the set of vertices $x \in V(G)$ such that there is a directed path from A_1 to x in $V(G) \setminus B_0$ with length at most i. We remark that $|(U_r \setminus U_{r-1}) \cap A_2| \leqslant \ell + 1$ for any $r \leqslant \ell$, else, by taking a directed path of length $\ell - r$ in $(U_r \setminus U_{r-1}) \cap A_2$ together with a path of length r from A_1 to that path's initial vertex, we would be able to find a directed path in $V(G) \setminus B_0$ from A_1 to A_2 of length ℓ. In particular, we have $|U_r \cap A_2| \leqslant (\ell + 1)^2$ for any $r \leqslant \ell$.

Let $r \in [\ell]$ be minimal such that $|U_r \setminus U_{r-1}| \leqslant 3\ell^{-1} n$. Set $W_2 = U_{r-1}$, $B = B_0 \cup (U_r \setminus U_{r-1})$, and $W_1 = V(G) \setminus (U_r \cup B_0)$, so that $W_1 \cup W_2 \cup B$ is a partition of $V(G)$. Because $A_1 \setminus B_0 \subseteq W_2$, we have $|W_2| \geqslant \varepsilon n/2$. Because $A_2 \setminus (U_r \cup B_0) \subseteq W_1$, we have $|W_1| \geqslant \varepsilon n - (\ell+1)^2 - \eta n \geqslant \varepsilon n/2$. From the choice of r, we have $|B| \leqslant 3\ell^{-1} n + \eta n \leqslant 4\ell^{-1} n$. Finally, the fact that $x \to y$ for every $x \in W_1$, $y \in W_2$ follows from the definition of U_r and U_{r-1}. $\square$

Using a repeated application of Lemma 5.5, we are now ready to state and prove the tournament decomposition referred to at the start of this section.

Lemma 5.6. *Suppose $\eta \ll \varepsilon$ and let $\ell = \lceil \varepsilon^{-3} \rceil$. Suppose G is a tournament with $|G| \leqslant 3n$. Then, there is a partition $V(G) = B \cup W_1 \cup \ldots \cup W_r$ so that $|B| \leqslant \varepsilon n$ and the following properties hold.*

K1 *If $1 \leqslant i < j \leqslant r$ and $x \in W_i$, $y \in W_j$, then $x \to y$.*

K2 *For $i \in [r]$, if $|W_i| \geqslant \sqrt{\varepsilon} n$, then $G[W_i]$ is $(\varepsilon n, \eta n, \ell)$-well-connected.*

Proof. Initially, set $B^{(1)} = \emptyset$ and $W_1^{(1)} = V(G)$. Then, for $r \geqslant 1$, do the following. We are given a partition $V(G) = B^{(r)} \cup W_1^{(r)} \cup \ldots \cup W_r^{(r)}$ with $|B^{(r)}| \leqslant 5r\varepsilon^3 n$, such that $|W_i^{(r)}| \geqslant \varepsilon n/2$ for each $i \in [r]$, and, if $1 \leqslant i < j \leqslant r$ and $x \in W_i^{(r)}$, $y \in W_j^{(r)}$, then $x \to y$. If we have that $G[W_i^{(r)}]$ is $(\varepsilon n, \eta n, \ell)$ well-connected whenever $|W_i^{(r)}| \geqslant \sqrt{\varepsilon} n$, then set $B = B^{(r)}$ and $W_i = W_i^{(r)}$ for $i \in [r]$. Otherwise, let $j \in [r]$ be such that $W_j^{(r)}$ is not $(\varepsilon n, \eta n, \ell)$-well-connected, with $|W_j^{(r)}|$ maximal (so $|W_j^{(r)}| \geqslant \sqrt{\varepsilon} n$). By Lemma 5.5, there is a partition $W_j^{(r)} = U_1 \cup U_2 \cup B_r$ so that $|U_1|, |U_2| \geqslant \varepsilon n/2$, $|B_r| \leqslant 4\ell^{-1} n \leqslant 5\varepsilon^3 n$, and $x \to y$ for every $x \in U_1$ and $y \in U_2$. We then set

$$B^{(r+1)} = B^{(r)} \cup B_r$$

$$W_i^{(r+1)} = \begin{cases} W_i^{(r)} & \text{if } 1 \leqslant i < j \\ U_1 & \text{if } i = j \\ U_2 & \text{if } i = j+1 \\ W_{i-1}^{(r)} & \text{if } j+1 < i \leqslant r+1 \end{cases}$$

We remark that $V(G) = B^{(r+1)} \cup W_1^{(r+1)} \cup \ldots \cup W_{r+1}^{(r+1)}$ is a partition with $|B^{(r+1)}| \leqslant 5(r+1)\varepsilon^2 n$, such that $|W_i^{(r+1)}| \geqslant \varepsilon n/2$ for each $i \in [r]$, and, if $1 \leqslant i < j \leqslant r+1$ and $x \in W_i^{(r+1)}$, $y \in W_j^{(r+1)}$, then $x \to y$, and so the procedure may continue.

On the r^{th} iteration of this procedure, the largest $|W_i^{(r)}|$ that is not $(\varepsilon n, \eta n, \ell)$-well-connected has size at most $3n - (r-1) \cdot \varepsilon n/2$, and so the procedure will terminate after at most $6\varepsilon^{-1}$ iterations, at which point we find $|B^{(r)}| \leqslant 30\varepsilon^2 n \leqslant \varepsilon n$. □

5.2 Proof of Theorem 1.1 and Theorem 1.2

We first prove that our two main theorems hold when the tournament is well-connected.

Lemma 5.7. *Suppose $1/n \ll \eta \ll \varepsilon \ll \alpha$ and let $\ell = \lceil \varepsilon^{-3} \rceil$. Suppose G is a tournament which is $(\varepsilon n, 5\eta^{1/4} n, \ell)$-well-connected, and that T is an n-vertex oriented tree.*

(1) *Suppose that $|G| = ((1+\alpha)n + k)$ where k is the number of leaves of T. Then, G contains a copy of T.*

(2) *Suppose that c is a constant such that $1/n \ll c \ll \eta$, that $|G| = (1+\alpha)n$, and that $\Delta(T) \leqslant cn$. Then, G contains a copy of T.*

Proof. The proof for each statement of this theorem is nearly identical, so here we will present a proof for (1), and explain in the footnotes any places where the proof for (2) differs.

Fix $\alpha > 0$ and introduce a constant m such that $1/n \ll 1/m \ll \eta \ll \varepsilon \ll \alpha$. Fix an n-vertex k-leaf oriented tree T and let G be a $((1+\alpha)n + k)$-vertex tournament which is $(\varepsilon n, 5\eta^{1/4} n, \ell)$-well-connected. We will show that G contains a copy of T, thus proving (1).[1]

Let $U_0 \subseteq V(G)$ be a random subset, with elements from $V(G)$ chosen independently at random with probability $2\sqrt{\eta}$, and let W_0 be the set of vertices v in $V(G) \setminus U_0$ with $d^{\pm}(v, U_0) \geqslant 4\eta n$. By Proposition 2.14, we have that $|V(G) \setminus W_0| \leqslant 24\sqrt{\eta} n$ with high probability. $V(G) \setminus U_0$ may be regarded as a random subset of $V(G)$ with elements chosen independently at random with probability $1 - 2\sqrt{\eta}$, and so, by Lemma 5.4, we have that $G[V(G) \setminus U_0]$ is $(6\varepsilon n, 25\sqrt{\eta} n, \ell+6)$-well-connected with high probability. Therefore, we may proceed assuming that $|V(G) \setminus W_0| \leqslant 24\sqrt{\eta} n$, and, using Lemma 5.2 i), that $G[W_0]$ is $(6\varepsilon n, \sqrt{\eta} n, \ell+6)$-well-connected.

Let $U_1 \subseteq W_0$ be a random subset, with elements from W_0 chosen independently at random with probability $\alpha/36$, and let W_1 be the set of vertices v in $W_0 \setminus U_1$ with $d^{\pm}(v, U_1) \geqslant 36\varepsilon n$. By Proposition 2.14, we have

[1]For (2), fix $\alpha > 0$ and introduce a constant m such that $1/n \ll c \ll 1/m \ll \eta \ll \varepsilon \ll \alpha$. Fix an n-vertex oriented tree T with $\Delta(T) \leqslant cn$ and let G be a $(1+\alpha)n$-vertex tournament which is $(\varepsilon n, 5\eta^{1/4} n, \ell)$-well-connected. We will show that G contains a copy of T, thus proving (2).

that $|W_1 \setminus W_0| \leqslant \alpha n/3$ with high probability, and, by Lemma 5.4, we have that $G[U_1]$ is $(36\varepsilon n, \eta n, \ell+12)$-well-connected with high probability. Therefore, we may proceed assuming that $|W_1| \geqslant ((1+\alpha/2)n+k)$, and that $G[U_1]$ is $(36\varepsilon n, \eta n, \ell+12)$-well-connected. [2]

Let $q = \ell + 14$. By Lemma 2.1, there exist forests $T_0 \subseteq T_1 \subseteq T_2 \subseteq T_3 \subseteq T_4 = T$, such that T_3 is a tree and properties **A1**-**A5** hold. By Theorem 3.1, $G[W_1]$ contains a copy, R_1 say, of T_1.[3] By Theorem 2.2 applied iteratively to the components of $T_2 - V(T_1)$, $G[W_1] - V(R_1)$ then contains a copy of $T_2 - V(T_1)$, which, taken together with R_1, gives a copy, R_2 say, of T_2.

Let $P_1, \ldots, P_r$ be the paths of length $\ell + 14$ attached to T_2 to obtain T_3. For $i \in [r]$, let x_i, y_i be the endvertices of P_i, let $P'_i = P_i - x_i - y_i$ (so that P'_i has length $\ell + 12$), and let x'_i, y'_i be the images of x_i, y_i in R_2. For each $i \in [r]$ in turn, using Lemma 5.2 ii), there is a copy Q_i of P'_i in the unoccupied vertices of $G[U_1]$, with first vertex in $N^{\diamond_1}(x'_i, U_1)$ and last vertex in $N^{\diamond_2}(y'_i, U_1)$, where $\diamond_1, \diamond_2 \in \{+,-\}$ are taken so that $x'_i Q_i y'_i$ gives a copy of P_i. We remark that we may always proceed as, by **A5**, the total number of vertices being embedded into U_1 is at most $|T_3 \setminus T_2| \leqslant \eta n$, and $G[U_1]$ is $(36\varepsilon n, \eta n, \ell + 12)$-well-connected with $d^{\pm}(v, U_1) \geqslant 36\varepsilon n$ for every $v \in W_1$. Thus, we obtain a copy, R_3 say, of T_3 in $G[W_0]$. Finally, using Corollary 2.3, R_3 can be extended to an copy of $T_4 = T$ in G, with the vertices of $V(T_4) \setminus V(T_3)$ copied to U_0. □

To finish the proof of both our main results simultaneously we will use *superadditive set functions*. In the proof of each case, we define a function $f_T : \mathcal{P}(V(T)) \to \mathbb{N}_0$ on the power set of $V(T)$, where, for $A \subseteq V(T)$, $f_T(A)$ may be interpreted as representing a rough upper bound on the number of vertices required to guarantee a copy of $T[A]$, according to each of the theorems. The only limitation on f_T required for the proof to work is for it to be *superadditive*.

Definition 5.8. *Given a set X, we say that a set function $f : \mathcal{P}(X) \to \mathbb{N}_0$ is* superadditive *if $f(A \cup B) \geqslant f(A) + f(B)$ for any disjoint sets $A, B \subseteq X$.*

In particular, we will use the property that, if $f : \mathcal{P}(X) \to \mathbb{N}_0$ is a superadditive set function and $X = A_1 \cup \ldots \cup A_r$ is a partition, then

$$f(X) \geqslant \sum_{i \in [r]} f(A_i). \tag{18}$$

We also remark that superadditive set functions are increasing, in the sense that if $A \subseteq B$, then $f(A) \leqslant f(B)$.

Given a tree T, we will have one particular superadditive set function $f_T : \mathcal{P}(V(T)) \to \mathbb{N}_0$ for each main theorem. For Theorem 1.1, we take $f_T(A) = |A| + k(A) - 2s(A)$, where $k(A)$ denotes the number of leaves of the forest $T[A]$ (and isolated vertices count as two leaves), and $s(A)$ denotes the number of components of the forest $T[A]$. For Theorem 1.2, we take $f_T(A) = |A|$. To see that the first function is superadditive, let $A, B \subseteq V(T)$ be disjoint sets, and compare the forest $T[A \cup B]$ to the forest $T[A] \cup T[B]$. $T[A] \cup T[B]$ can be reached from $T[A \cup B]$ by removing the edges with one endpoint in each of A and B one at a time. Each time an edge is removed, the total number of vertices remains the same, the total number of leaves increases by at most 2, and the total number of components increases by 1. Thus we find that $f_T(A \cup B) \geqslant f_T(A) + f_T(B)$.

We are now ready to prove Theorems 1.1 and 1.2 using Lemma 5.7. The proof for each theorem is nearly identical, so we will present a proof for Theorem 1.1, and explain in the footnotes any places where the proof for Theorem 1.2 differs. In each case, using Lemma 5.6, we will have a partition of most of the vertex set of the tournament G into sets $W_1, \ldots, W_r$ such that, if $u \in W_i$ and $v \in W_j$ with $i < j$, then $uv \in E(G)$, and such that each $G[W_i]$ is well-connected if W_i is not too small (i.e., **K1** and **K2** hold). It remains to find a good way to partition the tree T to embed it across this decomposition.

For each $i \in [r]$, if $|W_i| \geqslant \sqrt{\varepsilon} n$, then let $w_i = (1 - \alpha/4)|W_i|$, and otherwise let $w_i = |W_i|$. We want to assign a set U_i of w_i vertices of the tree to embed in W_i, for each $i \in [r]$. First, if $|W_r| \geqslant \sqrt{\varepsilon} n$, then we order the vertices of T as $v_1, \ldots, v_n$ so that all edges of T go forwards in this ordering, and let U_r be the set of the last w_r vertices in this ordering (in this case, we will be able to embed $T[U_r]$ into $G[W_r]$ by **K2** and Lemma 5.7). If $|W_r| < \sqrt{\varepsilon} n$, then, if possible, we let U_r be w_r out-leaves of T, and if it is not possible then we stop. If we have not stopped, then we remove U_r from T and repeat this procedure to find U_{r-1}, and so on. Note that if this stops then either we have assigned vertices for each W_i, or the remaining forest

[2] For (2), we may proceed assuming that $|W_1| \geqslant (1 + \alpha/2)n$, and that $G[U_1]$ is $(36\varepsilon n, \eta n, \ell + 12)$-well-connected.

[3] For (2), as $\Delta(T) \leqslant cn$, each tree S_v of **A2** satisfies $|S_v| \leqslant cmn + 1 \leqslant \eta n$. By Theorem 4.1, $G[W_1]$ contains a copy, R_1 say, of T_1.

has at most $\sqrt{\varepsilon}n$ out-leaves. In the latter case we carry out a similar assignment for $W_1, W_2, \ldots$. When this stops either all the vertices have been assigned or the remaining forest has at most $\sqrt{\varepsilon}n$ in-leaves as well as at most $\sqrt{\varepsilon}n$ out-leaves — such a forest we can embed with $O(\sqrt{\varepsilon}n)$ spare vertices using Theorem 2.4.

Proof of Theorem 1.1 (with appropriate alterations for Theorem 1.2 indicated). Fix $\alpha > 0$, and note that we may additionally assume that $\alpha \leqslant 1$. Introduce constants ε, η, n_0 such that $1/n_0 \ll \eta \ll \varepsilon \ll \alpha$. Given a tree T, let $f_T : \mathcal{P}(V(T)) \to \mathbb{N}_0$ be the superadditive set function defined by $f_T(A) = |A| + k(A) - 2s(A)$, where $k(A)$ denotes the number of leaves of the forest $T[A]$ (and isolated vertices count as two leaves), and $s(A)$ denotes the number of components of the forest $T[A]$.[4]

Let $n \geqslant n_0$. Fix an n-vertex k-leaf oriented tree T and let G be a $((1+\alpha)n + k)$-vertex tournament, so that $f_T(V(T)) + \alpha n \leqslant |G| \leqslant 3n$. We will show that G contains a copy of T, thus proving the theorem.[5]

By Lemma 5.6, there is a partition $V(G) = B \cup W_1 \cup \ldots \cup W_r$ so that $|B| \leqslant \varepsilon n$ and the properties **K1** and **K2** hold. For each $i \in [r]$, if $|W_i| \geqslant \sqrt{\varepsilon}n$, then let $w_i = (1-\alpha/4)|W_i|$, and otherwise let $w_i = |W_i|$. Partition $[r]$ into intervals I^-, I, I^+ (in that order), so that I is minimal subject to there being disjoint sets $U_i \subseteq V(T)$, $i \in I^- \cup I^+$, for which the following hold.

L1 For each $i \in I^- \cup I^+$, $(1-\varepsilon)w_i \leqslant f_T(U_i) \leqslant w_i$.

L2 There are no edges from $\cup_{i \in I^+} U_i$ to $V(T) \setminus (\cup_{i \in I^+} U_i)$ in T.

L3 There are no edges from $V(T) \setminus (\cup_{i \in I^-} U_i)$ to $\cup_{i \in I^-} U_i$ in T.

L4 If $|W_i| < \sqrt{\varepsilon}n$, then there are no edges in $T[U_i]$.

L5 If $i, j \in I^- \cup I^+$ with $i < j$, then there are no edges from U_j to U_i in T.

Note that this is possible as $I = [r]$ is a valid partition. Let $T' = T - \cup_{i \in I^+ \cup I^-} U_i$ and $W = \cup_{i \in I} W_i$. We will show that, for each $i \in I^- \cup I^+$, $G[W_i]$ contains a copy of $T[U_i]$, and $G[W]$ contains a copy of T'. Putting these together then gives a copy of T, by **L2**, **L3**, **L5**, and **K1**.

For each $i \in I^- \cup I^+$, if $|W_i| \geqslant \sqrt{\varepsilon}n$, then

$$|U_i| \geqslant f_T(U_i)/2 \overset{\mathbf{L1}}{\geqslant} (1-\varepsilon)w_i/2 \geqslant \sqrt{\varepsilon}n/4,$$

and

$$f_T(U_i) + (\alpha/4) \cdot |U_i| \leqslant (1+\alpha/4) \cdot f_T(U_i) \overset{\mathbf{L1}}{\leqslant} (1+\alpha/4)w_i \leqslant |W_i|,$$

and so $G[W_i]$ contains a copy of $T[U_i]$ by **K2** and Lemma 5.7. On the other hand, if $|W_i| < \sqrt{\varepsilon}n$, then by **L4**, $G[W_i]$ contains a copy of $T[U_i]$, noting that we have $f_T(U_i) = |U_i|$ in this case.

It is left to show that $G[W]$ contains a copy of T'. Note that this is trivial if $I = \emptyset$, and so we can assume $I \neq \emptyset$ and label $j_1, j_2 \in [r]$ so that I is the interval from j_1 to j_2. Also, note that, because $\sum_{i \in I^- \cup I^+} f_T(U_i) \leqslant 2n$,

$$\begin{aligned} |W| = |G| - |B| - \sum_{i \in I^- \cup I^+} |W_i| &\overset{\mathbf{L1}}{\geqslant} f_T(V(T)) + \alpha n - \varepsilon n - (1-\alpha/4)^{-1}(1-\varepsilon)^{-1} \sum_{i \in I^- \cup I^+} f_T(U_i) \\ &\geqslant f_T(V(T)) - \sum_{i \in I^- \cup I^+} f_T(U_i) + \alpha n/4 \overset{(18)}{\geqslant} f_T(V(T')) + \alpha n/4 \geqslant |T'| + \alpha n/4. \end{aligned}$$

If $|W_{j_2}| \geqslant \sqrt{\varepsilon}n$, then we must have $f_T(V(T')) < (1-\varepsilon)w_{j_2}$, otherwise we could order $V(T')$ as $v_1, \ldots, v_{|T'|}$ so that all edges of T' go forwards in this ordering, and define $U_{j_2} = \{v_s, \ldots, v_{|T'|}\}$ for some s chosen such that $(1-\varepsilon)w_{j_2} \leqslant f_T(U_{j_2}) \leqslant w_{j_2}$, a contradiction to the minimality of I. Thus, if $|W_{j_2}| \geqslant \sqrt{\varepsilon}n$, then $G[W_{j_2}]$, and hence $G[W]$, contains a copy of T' by Lemma 5.7. Similarly, if $|W_{j_1}| \geqslant \sqrt{\varepsilon}n$, then $G[W_{j_1}]$, and hence $G[W]$, contains a copy of T'. We must have then that $|W_{j_1}| < \sqrt{\varepsilon}n$ and $|W_{j_2}| < \sqrt{\varepsilon}n$. Thus, by the minimality of I, T' has at most $w_{j_2} \leqslant \sqrt{\varepsilon}n$ out-leaves and at most $w_{j_1} \leqslant \sqrt{\varepsilon}n$ in-leaves. As $|W| \geqslant |T'| + \alpha n/4$, $G[W]$ then contains a copy of T' by Theorem 2.4, as required. □

[4] For (2), fix $\alpha > 0$ and introduce constants $\varepsilon, \eta, c, n_0$ such that $1/n_0 \ll c \ll \eta \ll \varepsilon \ll \alpha$. Given a tree T, let $f_T : \mathcal{P}(V(T)) \to \mathbb{N}_0$ be the superadditive set function defined by $f_T(A) = |A|$.

[5] For (2), fix an n-vertex oriented tree T with $\Delta(T) \leqslant cn$ and let G be a $(1+\alpha)n$-vertex tournament, so that $f_T(V(T)) + \alpha n = |G| \leqslant 3n$. We will show that G contains a copy of T, thus proving the theorem.

6 Proof of Theorem 3.4

In this section, we prove Theorem 3.4, which finds an index j_t for a regularity cluster for the core T_0 of a tree, and a random homomorphism of a fixed digraph H with vertex weight function β representing an average component of $T_1 - V(T_0)$ (using the notation in Section 2.2). For convenience, we restate the definition of H (see Figure 3) and Theorem 3.4. Let H be the oriented forest with vertex and edge sets given by

$$V(H) = \{x^+, y^+, z^+, u^+, w^+, \bar{x}^+, \bar{z}^+, \bar{u}^+, \bar{w}^+, x^-, y^-, z^-, u^-, w^-, \bar{x}^-, \bar{z}^-, \bar{u}^-, \bar{w}^-\},$$
$$E(H) = \{x^+y^+, z^+x^+, z^+u^+, w^+z^+, \bar{z}^+\bar{x}^+, \bar{z}^+\bar{u}^+, \bar{w}^+\bar{z}^+, y^-x^-, x^-z^-, u^-z^-, z^-w^-, \bar{x}^-\bar{z}^-, \bar{u}^-\bar{z}^-, \bar{z}^-\bar{w}^-\}.$$

For each $\diamond \in \{+, -\}$, let $X^\diamond = \{x^\diamond, \bar{x}^\diamond\}$. Let $X = X^+ \cup X^-$.

Theorem 3.4. *Let $1/r \ll \varepsilon \ll \mu \ll \alpha < 1$. Let $\beta : V(H) \to [0,1]$ be a function satisfying $\sum_{v \in V(H)} \beta(v) = 1$ with $\beta(y^+) \geqslant \beta(x^+)$ and $\beta(y^-) \geqslant \beta(x^-)$, and, for every $v \in V(H)$, $\beta(v) \geqslant \mu$. Let D be a complete looped digraph on vertex set $[r]$ with ε-complete edge weights $d(e)$, $e \in E(D)$. Let*

$$\gamma = \max\{\beta(x^+, \bar{x}^+), \beta(z^+, \bar{z}^+)\} + \max\{\beta(x^-, \bar{x}^-), \beta(z^-, \bar{z}^-)\}. \tag{2}$$

Then, there is some $j_t \in [r]$ and a random $\phi : H \to D$ such that the following hold.

E1 *With probability 1, ϕ is a homomorphism from H to D, and $j_t \notin \phi(\{x^+, \bar{x}^+, x^-, \bar{x}^-\})$.*

E2 *For each $j \in [r]$, $\mathbb{E}(\beta(\phi^{-1}(j))) \leqslant \frac{1+\gamma+\alpha}{r}$.*

E3 *For each $j \in [r]$, either*

E3.1 $\mathbb{E}(\beta(\phi^{-1}(j) \cap X^+)) \leqslant d(j_t, j) \cdot \frac{1+\gamma+\alpha}{r}$ *and* $\mathbb{E}(\beta(\phi^{-1}(j) \cap X)) \leqslant d(j, j_t) \cdot \frac{1+\gamma+\alpha}{r}$*, or*

E3.2 $\mathbb{E}(\beta(\phi^{-1}(j) \cap X^-)) \leqslant d(j, j_t) \cdot \frac{1+\gamma+\alpha}{r}$ *and* $\mathbb{E}(\beta(\phi^{-1}(j) \cap X)) \leqslant d(j_t, j) \cdot \frac{1+\gamma+\alpha}{r}$.

E4 *With probability 1, we have $|\phi(e)| = 2$ for every $e \in E(H)$.*

To prove Theorem 3.4, we first make two key simplifications before dividing into three critical cases. Our first simplification is to work only with the vertices of H representing components attached by an out-edge from T_0 (see Figure 3). Let H^+ and H^- be the subdigraphs of H induced on the vertices with $+$ and $-$ in the superscript, respectively (i.e., the right and the left parts of H in Figure 3). Considering each possible location $j \in [r]$ for j_t in Theorem 3.4, either a) j has enough weight on its out-edges that a random embedding of H^- can be extended relatively easily (with perhaps some modification) to one of H satisfying our requirements, or b) j has enough weight on its in-edges to similarly extend a random embedding of H^+. If many $j \in [r]$ satisfy a), then we may randomly embed H^- into the weighted looped digraph D induced on these j. If not, then enough $j \in [r]$ satisfy b), so that we may randomly embed H^+ into the weighted looped digraph D induced on these j. By appealing to directional duality if necessary, this allows us to prove a simplified version of Theorem 3.4 with weight only on H^+ (see Section 6.4).

Our second simplification to Theorem 3.4 is to drop the condition **E4**; we later show this condition can be recovered without undue difficulty. These two simplifications of Theorem 3.4 result in Theorem 6.5, which we state in Section 6.2 after introducing a notational framework of 'distillations' in Section 6.1 in order to have a concise and consistent language for the proofs in this section. The proof of Theorem 6.5 varies depending on the weight distribution β on the vertices in H. This falls into three main cases, which we also state in Section 6.2, in the form of Lemmas 6.6, 6.8 and 6.9, before deducing Theorem 6.5 from these cases. We then prove the lemma for each of these cases in Section 6.3, before finally deducing Theorem 3.4 from Theorem 6.5 in Section 6.4.

6.1 Distillations

We prove Theorem 3.4 from three specific cases where, roughly speaking, H is replaced by simpler subgraphs of H. In order to have a concise and consistent language for proving these cases, we will use the notion of a *distillation*, as follows.

Definition 6.1. *A* distillation *is a triple $\mathcal{F} = (F, X, \beta)$, where F is an oriented forest, $X \subseteq V(F)$ is a set containing precisely one vertex in each component of F, and $\beta : V(F) \to [0,1]$ satisfies $\sum_{v \in V(F)} \beta(v) = 1$.*

In Theorem 3.4, we have a distillation (H, X, β) which we used to represent the average component of $T - V(T_0)$ for Theorem 3.1. There is some flexibility in how we could have chosen this distillation — for example we could move all the weight from y^+ to x^+, or from u^+ to z^+ and still have a useful distillation of the average component if we can find a matching random homomorphism. However, H is the smallest digraph that records enough structure in the average component to allow every applicable distillation to have a matching random homomorphism. The construction of the random homomorphism falls into three cases depending on the distribution of the weight — in each case we can move weight off some vertices (different in each case) to simplify the digraph in the distillation for which we find a random homomorphism.

To describe which simplifications of distillations are valid in this way formally, and prove this validity, we will define a transitive relation $\hookrightarrow$ between distillations. Very roughly, given two distillations $\mathcal{F}_0$ and $\mathcal{F}_1$, if $\mathcal{F}_0 \hookrightarrow \mathcal{F}_1$, then we can move weight in $\mathcal{F}_1$ (and possibly delete vertices) to transform it into $\mathcal{F}_0$. Formally, we define the relation as follows.

Definition 6.2. *Given distillations $\mathcal{F}_i = (F_i, X_i, \beta_i)$ for $i \in \{0,1\}$, say $\mathcal{F}_0 \hookrightarrow \mathcal{F}_1$ if there is a random homomorphism $\rho : F_0 \to F_1$ with the following properties.*

- **M1** *With probability 1, $\rho(X_0) \subseteq X_1$.*
- **M2** $\mathbb{E}(\beta_0(\rho^{-1}(v))) = \beta_1(v)$ *for every $v \in V(F_1)$.*

Finally, we need a notion of which distillations are useful — i.e., which distillations have a matching random homomorphism with properties like those in Theorem 3.4. It will be convenient to consider a small collection of distillations and conclude that at least one has such a matching random homomorphism, and so we define the following notion of γ-goodness on sets of distillations. Roughly speaking, γ corresponds to the extra proportion of vertices we need to embed the tree, as in the use of Theorem 3.4.

Definition 6.3. *Given $\gamma \geqslant 0$, and distillations $\mathcal{F}_i = (F_i, X_i, \beta_i)$, $i \in [m]$, we say $\{\mathcal{F}_i\}_{i=1}^m$ is* γ-good *if the following holds for any fixed $\alpha > 0$: if $1/r \ll \varepsilon \ll \alpha$ and D is a complete looped digraph on vertex set $[r]$ with ε-complete edge weights $d(e)$, $e \in E(D)$, then there exists some $j_t \in [r]$ and a random $(\phi, i(\phi))$ with the following properties.*

- **N1** *With probability 1, we have that $i(\phi) \in [m]$, that ϕ is a homomorphism from $F_{i(\phi)}$ to D, and that $j_t \notin \phi(X_{i(\phi)})$.*
- **N2** *For each $j \in [r]$, $\mathbb{E}(\beta_{i(\phi)}(\phi^{-1}(j))) \leqslant \frac{1+\gamma+\alpha}{r}$.*
- **N3** *For each $j \in [r]$, $\mathbb{E}(\beta_{i(\phi)}(\phi^{-1}(j) \cap X_{i(\phi)})) \leqslant d(j_t, j) \cdot \frac{1+\gamma+\alpha}{r}$.*

Note that if $\{\mathcal{F}_i\}_{i=1}^m$ is γ-good for some $\gamma \geqslant 0$, then $\{\mathcal{F}_i\}_{i=1}^m$ is γ'-good for every $\gamma' \geqslant \gamma$. In addition, a set of distillations is γ-good if and only if it contains a non-empty subset which is γ-good.

Finally here, we prove the following key lemma that confirms that if a distillation can be simplified via the relation $\hookrightarrow$ to each one of a family of distillations which are collectively γ-good, then that original distillation is γ-good, as follows.

Lemma 6.4. *Let $\gamma \geqslant 0$, and suppose $\mathcal{F}$ and $\mathcal{G}_1, \ldots, \mathcal{G}_m$ are distillations such that $\mathcal{F} \hookrightarrow \mathcal{G}_i$ for every $i \in [m]$. If $\{\mathcal{G}_i\}_{i=1}^m$ is γ-good, then $\{\mathcal{F}\}$ is γ-good.*

Proof. Let $\mathcal{F} = (F, X, \beta)$ and $\mathcal{G}_i = (G_i, X_i, \beta_i)$ for each $i \in [m]$. For each $i \in [m]$, let $\rho_i : F \to G_i$ be a random homomorphism realising $\mathcal{F} \hookrightarrow \mathcal{G}_i$.

Take $1/r \ll \varepsilon \ll \alpha$, and let D be a complete looped digraph on vertex set $[r]$ with ε-complete edge weights $d(e)$, $e \in E(D)$. Let $j_t \in [r]$ and $(\phi, i(\phi))$ realise that $\{\mathcal{G}_i\}_{i=1}^m$ is γ-good in the case of D. Define $(\psi, k(\psi))$ as follows. First, sample $(\phi, i(\phi))$. Then, with $i(\phi)$ now fixed, sample $\rho_{i(\phi)}$, and set $\psi = \phi \circ \rho_{i(\phi)}$. Let $k(\psi) = 1$ with probability 1, and note that **N1** holds for $(\psi, k(\psi))$.

For $i \in [m]$ let A_i be the event $\{i(\phi) = i\}$. Then, by the law of total expectation,

$$\mathbb{E}(\beta(\psi^{-1}(j))) = \sum_{i\in[m]} \mathbb{P}(A_i) \cdot \mathbb{E}(\beta(\psi^{-1}(j)) \mid A_i) = \sum_{i\in[m]} \mathbb{P}(A_i) \cdot \mathbb{E}(\beta(\rho_i^{-1}(\phi^{-1}(j))) \mid A_i)$$
$$\overset{\mathbf{M2}}{=} \sum_{i\in[m]} \mathbb{P}(A_i) \cdot \mathbb{E}(\beta_i(\phi^{-1}(j)) \mid A_i) = \mathbb{E}(\beta_{i(\phi)}(\phi^{-1}(j))),$$

so **N2** holds for $(\psi, k(\psi))$, and

$$\mathbb{E}(\beta(\psi^{-1}(j) \cap X)) = \sum_{i\in[m]} \mathbb{P}(A_i) \cdot \mathbb{E}(\beta(\psi^{-1}(j) \cap X) \mid A_i) = \sum_{i\in[m]} \mathbb{P}(A_i) \cdot \mathbb{E}(\beta(\rho_i^{-1}(\phi^{-1}(j)) \cap X) \mid A_i)$$
$$\leqslant \sum_{i\in[m]} \mathbb{P}(A_i) \cdot \mathbb{E}(\beta(\rho_i^{-1}(\phi^{-1}(j) \cap \rho_i(X))) \mid A_i)$$
$$\overset{\mathbf{M1}}{\leqslant} \sum_{i\in[m]} \mathbb{P}(A_i) \cdot \mathbb{E}(\beta(\rho_i^{-1}(\phi^{-1}(j) \cap X_i)) \mid A_i) \overset{\mathbf{M2}}{=} \mathbb{E}(\beta_{i(\phi)}(\phi^{-1}(j) \cap X_{i(\phi)})),$$

so **N3** holds for $(\psi, k(\psi))$. □

6.2 Statement of overarching theorem and subcases

Let H_0 have vertex set $\{x, y, z, u, w, \bar{x}, \bar{z}, \bar{u}, \bar{w}\}$ and edge set $\{xy, zx, zu, wz, \bar{z}\bar{x}, \bar{z}\bar{u}, \bar{w}\bar{z}\}$, noting that this is the oriented forest which is the subdigraph of H defined at the start of this section restricted to the vertices with $+$ in the superscript, and with vertices labelled more concisely (see also Figure 3). As noted at the start of this section, we will first prove a version of Theorem 3.4 for this subdigraph of H, without the condition **E4**, before deducing Theorem 3.4 from this in Section 6.4. Having introduced our relevant notation, we can now state this version of Theorem 3.4 concisely, as follows.

Theorem 6.5. *Let $\beta_0 : V(H_0) \to [0,1]$ be a function with $\sum_{v\in V(H_0)} \beta_0(v) = 1$ and $\beta_0(y) \geqslant \beta_0(x)$, and set $X_0 = \{x, \bar{x}\}$. Set $\gamma = \max\{\beta_0(x, \bar{x}), \beta_0(z, \bar{z})\}$. Let $\mathcal{H}_0 = (H_0, X_0, \beta_0)$. Then, $\{\mathcal{H}_0\}$ is γ-good.*

As mentioned before, the proof of Theorem 6.5 depends on the weight distribution β on H_0. Dividing into cases, solving them, and showing they combine to prove this theorem is no easy task. Doing so while additionally motivating the choice of these cases is more difficult still. However, while we do concentrate on giving as clear and concise a proof of Theorem 6.5 as possible, we will give some motivation behind the cases by relating them to an embedding of a tree T into a tournament G.

In particular, our notation is designed to make the cases as efficient as possible to check, rather than explain the larger understanding that is necessary to produce these cases. To motivate the cases more directly, we now recall the discussion in Section 2.2. Our aim is to use Theorem 3.4 to embed a tree T with a small core T_0, where $T - V(T_0)$ is a collection of small components. To do this we take the tournament G and find a regularity partition $V_1 \cup \ldots \cup V_r$. We then want to make a careful choice of $j_t \in [r]$ and embed T_0 into V_{j_t} before distributing the components of $T - V(T_0)$ across the clusters of the regularity partition. The choice of j_t restricts which component any vertex in $v \in V(T) \setminus V(T_0)$ can be embedded to. For example, if the path from T_0 to v in T is a directed path towards v, and $v \in V_i$, then there must be a directed path from V_{j_t} to V_i of edges with positive weight in the reduced digraph obtained from the regularity partition. Fortunately, for our cases we need to consider at most the direction of the first three edges on the path from T_0 to v (and first edge will always be directed away from T_0).

Very roughly, we first divide into two cases corresponding to the following situations, where, for example, we use the notation of a $(++)$-path from u to v to be a length two path from u to v comprised of two edges directed forwards from u to v, with other notation used similarly.

- For most of the vertices $v \in V(T) \setminus V(T_0)$, if T_0 is embedded to V_{j_t}, $i \in [r]$, and there is a $(++)$-path from V_{j_t} to V_i in D, then we could embed v to V_i.
- For most of the vertices $v \in V(T) \setminus V(T_0)$, if T_0 is embedded to V_{j_t}, $i \in [r]$, and there is a $(+-)$-path from V_{j_t} to V_i in D then we could embed v to V_i.

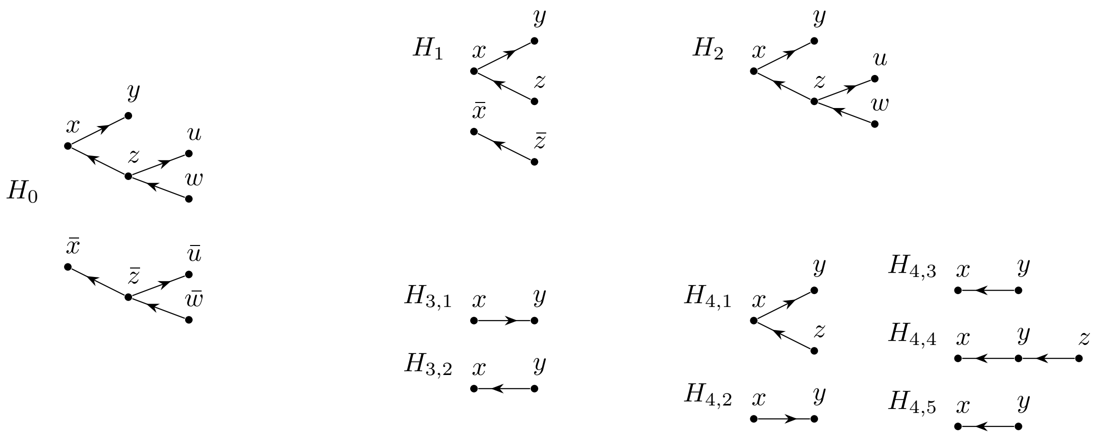

Figure 6: The underlying forests of the distillations described in this section.

These cases correspond roughly to Lemma 6.6 and Lemma 6.7, respectively. We then further subdivide the latter case into two cases, essentially replacing the $(+-)$-path with a $(+--+)$-path and a $(+---)$-path, respectively. This gives us cases 1, 2, and 3 which, in terms of the weight distribution β on H correspond to the following roughly-defined three cases:

1. Most of the weight not on $\{x, \bar{x}\}$ is on y.
2. Most of the weight not on $\{x, \bar{x}\}$ is on $\{y, u, \bar{u}\}$ (but Case 1 does not apply).
3. Most of the weight not on $\{x, \bar{x}\}$ is on $\{z, w, \bar{z}, \bar{w}\}$.

As previously described, H_0 is the minimal digraph which records all the structure we need. For example, we need two components for H_0 so that we can have $\beta(y) \geqslant \beta(x)$. The vertices $\bar{u}$ and $\bar{w}$ are not used explicitly in these cases below, as in the relevant case we are able to move the weight on $\bar{x}$ to x, $\bar{z}$ to z, and so on.

We now use our concept of distillations and the relation $\hookrightarrow$ to state the lemmas corresponding to these cases and combine them to prove Theorem 6.5. We first divide Theorem 6.5 into two lemmas – based on the distribution of β, we distill H_0 into H_1 or H_2, where for the former we remove the vertices $\{u, w, \bar{u}, \bar{w}\}$ and in the latter we remove $\{\bar{x}, \bar{z}, \bar{u}, \bar{w}\}$. This gives Lemma 6.6 (corresponding to Case 1 above) and Lemma 6.7. We then break Lemma 6.7 into two further lemmas, Lemma 6.8 and 6.9 which correspond respectively to Case 2 and 3 above, where in each case we have a set of distillations rather than simplifying to just one distillation. The structure of this division is depicted in Figure 7.

Let H_1 be the oriented forest with vertex set $\{x, y, z, \bar{x}, \bar{z}\}$ and edge set $\{xy, zx, \bar{z}\bar{x}\}$.

Lemma 6.6. *Let $\mathcal{H} = (H_1, \{x, \bar{x}\}, \beta)$ be a distillation with $\beta(y) \geqslant \beta(z, \bar{z}), \beta(x)$. Then, $\{\mathcal{H}\}$ is $\beta(x, \bar{x})$-good.*

Let H_2 be the oriented forest with vertex set $\{x, y, z, u, w\}$ and edge set $\{xy, zx, zu, wz\}$.

Lemma 6.7. *Let $\mathcal{H} = (H_2, \{x\}, \beta)$ be a distillation with $\beta(y) \leqslant \beta(z, u, w)$. Then, $\{\mathcal{H}\}$ is $(\max\{\beta(x), \beta(z)\})$-good.*

Let $H_{3,1}$ and $H_{3,2}$ be the oriented forests with vertex set $\{x, y\}$ and edge set $\{xy\}$ and $\{yx\}$ respectively.

Lemma 6.8. *Let $\beta_x \in [0, 1]$ and, for $i \in [2]$, set $\beta_i(x) = (1 + \beta_x)/2$ and $\beta_i(y) = (1 - \beta_x)/2$. For $i \in [2]$, set $\mathcal{H}_i = (H_{3,i}, \{x\}, \beta_i)$. Then, $\{\mathcal{H}_i\}_{i=1}^2$ is β_x-good.*

Let $H_{4,1}$ be the oriented forest with vertex set $\{x, y, z\}$ and edge set $\{xy, zx\}$. Let $H_{4,2}$ be the oriented forest with vertex set $\{x, y\}$ and edge set $\{xy\}$. Let $H_{4,3}$ be the oriented forest with vertex set $\{x, y\}$ and edge set $\{yx\}$. Let $H_{4,4}$ be the oriented forest with vertex set $\{x, y, z\}$ and edge set $\{yx, zy\}$. Let $H_{4,5} = H_{4,3}$.

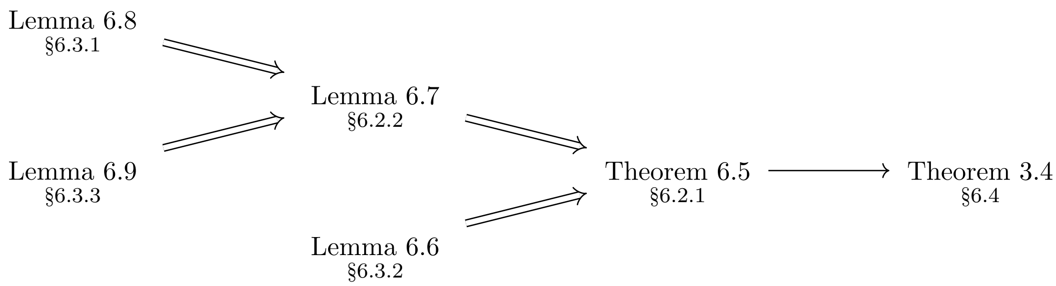

Figure 7: An overview of how the results of this section combine to prove Theorem 3.4. Each implication denoted by $\Longrightarrow$ indicates a suitable application of Lemma 6.4.

Lemma 6.9. *Let $\beta_x, \beta_y, \beta_z, \beta_u, \beta_w \in [0,1]$ have sum 1 and $\beta_y + \beta_u \leqslant \beta_z + \beta_w$. Let $\gamma = \max\{\beta_x, \beta_z\}$. Take the following weight functions $\beta_i : V(H_{4,i}) \to [0,1]$ for $i \in [5]$.*

$$\beta_1(y) = \max\{\beta_x + \beta_y - \gamma, 0\} \qquad \beta_1(z) = \min\{\beta_w + \beta_z, 1 - \beta_z\} \qquad \beta_1(x) = 1 - \beta_1(y) - \beta_1(z),$$

$$\beta_2(x) = \beta_3(x) = \beta_x + \beta_z + \beta_w \qquad \beta_2(y) = \beta_3(y) = \beta_y + \beta_u,$$

$$\beta_4(x) = \min\{\beta_x + \beta_y + \beta_u, \max\{\beta_x + \beta_y, \beta_z\}\} \qquad \beta_4(z) = \beta_w \qquad \beta_4(y) = 1 - \beta_4(x) - \beta_4(z),$$

$$\beta_5(x) = \beta_x + \beta_y \qquad \beta_5(y) = \beta_z + \beta_u + \beta_w.$$

For $i \in [5]$, set $\mathcal{H}_i = (H_{4,i}, \{x\}, \beta_i)$. Then, $\{\mathcal{H}_i\}_{i=1}^5$ is γ-good.

We remark that each set of weights defined in Lemma 6.9 sum to 1, either by the choice of $\beta_1(x)$ or $\beta_4(y)$ or as $\beta_x + \beta_y + \beta_z + \beta_u + \beta_w = 1$. Furthermore, from the choices they can all immediately be seen to be non-negative except for $\beta_1(x)$ and $\beta_4(y)$, but this we can also show, as follows.

First note that

$$\begin{aligned}\beta_4(y) &= 1 - \min\{\beta_x + \beta_y + \beta_u + \beta_w, \max\{\beta_x + \beta_y + \beta_w, \beta_z + \beta_w\}\} \\ &= \max\{\beta_z, \min\{\beta_u + \beta_z, \beta_x + \beta_y + \beta_u\}\}.\end{aligned} \tag{19}$$

Therefore, $\beta_4(y) \geqslant 0$. For use later, we will show that $\beta_1(x) \geqslant \beta_4(y)$, which also then confirms that $\beta_1(x) \geqslant 0$.

First suppose that $\beta_z \leqslant \beta_x + \beta_y$. Then, $1 - \beta_z = \beta_u + \beta_y + \beta_x + \beta_w \geqslant \beta_w + \beta_z$, so that $\beta_1(z) = \beta_w + \beta_z$, and hence

$$\beta_1(x) = 1 - \beta_1(y) - \beta_1(z) = 1 - (\beta_x + \beta_y - \gamma) - (\beta_w + \beta_z) = \beta_u + \gamma \geqslant \beta_u + \beta_z.$$

On the other hand, if $\beta_z > \beta_x + \beta_y$, then $\beta_1(y) = 0$, so that

$$\beta_1(x) = 1 - \beta_1(z) = \max\{1 - (\beta_w + \beta_z), \beta_z\} = \max\{\beta_u + \beta_y + \beta_x, \beta_z\}.$$

Therefore, in both cases, we have $\beta_1(x) \geqslant \beta_z$ and either $\beta_1(x) \geqslant \beta_u + \beta_z$ or $\beta_1(x) \geqslant \beta_u + \beta_y + \beta_x$. Thus, by (19), we have

$$\beta_1(x) \geqslant \beta_4(y). \tag{20}$$

We will now outline how Lemmas 6.6, 6.7, 6.8, and 6.9 together imply Theorem 6.5 (see Figure 7). First, we will show that Theorem 6.5 follows from Lemmas 6.6 and 6.7. We will then show that Lemma 6.7 follows from Lemmas 6.8 and 6.9. The proof of each of these implications is a straightforward case of verifying certain $\hookrightarrow$ relations between the relevant distillations hold according to the different values β may take, and applying Lemma 6.4 to deduce γ-goodness. All that will remain then is to prove Lemmas 6.6, 6.8, and 6.9, which we do in Section 6.3, and then deduce Theorem 3.4 from Theorem 6.5, which we do in Section 6.4.

6.2.1 Proof of Theorem 6.5 using Lemmas 6.4, 6.6, and 6.7

Using Lemma 6.4, it is simple to deduce Theorem 6.5 from Lemmas 6.6 and 6.7.

Proof of Theorem 6.5. Define $\rho_1 : H_0 \to H_1$ by setting $\rho_1(x) = x$, $\rho_1(y) = y$, $\rho_1(z, u, w) = z$, $\rho_1(\bar{x}) = \bar{x}$ and $\rho_1(\bar{z}, \bar{u}, \bar{w}) = \bar{z}$. Define $\rho_2 : H_0 \to H_2$ by setting $\rho_2(x, \bar{x}) = x$, $\rho_2(y) = y$, $\rho_2(z, \bar{z}) = z$, $\rho_2(u, \bar{u}) = u$ and $\rho_2(w, \bar{w}) = w$.

For each $i \in [2]$, let $\beta_i : V(H_i) \to [0, 1]$ be given by $\beta_i(v) = \beta_0(\rho_i^{-1}(v))$. Let $\mathcal{H}_1 = (H_1, \{x, \bar{x}\}, \beta_1)$ and $\mathcal{H}_2 = (H_2, \{x\}, \beta_2)$. Note that, for each $i \in [2]$, the homomorphism $\rho_i : H_0 \to H_i$ realises $\mathcal{H}_0 \hookrightarrow \mathcal{H}_i$.

If $\beta_0(y) \geqslant \beta_0(z, u, w, \bar{z}, \bar{u}, \bar{w})$, then, by Lemma 6.6, $\{\mathcal{H}_1\}$ is $\beta_0(x, \bar{x})$-good. On the other hand, if $\beta_0(y) \leqslant \beta_0(z, u, w, \bar{z}, \bar{u}, \bar{w})$, then, by Lemma 6.7, $\{\mathcal{H}_2\}$ is $(\max\{\beta_0(x, \bar{x}), \beta_0(z, \bar{z})\})$-good. In either of these cases we find $\{\mathcal{H}_0\}$ is γ-good, by Lemma 6.4. □

6.2.2 Proof of Lemma 6.7 using Lemmas 6.4, 6.8, and 6.9

To prove Lemma 6.7 follows from Lemmas 6.8 and 6.9 using Lemma 6.4 requires more checking due to the larger sets of distillations, but this is straightforward, as follows.

Proof of Lemma 6.7. For $v \in V(H_2)$, let $\beta_v = \beta(v)$. Note that we have $\beta_y \leqslant \beta_z + \beta_u + \beta_w$. Define $\mathcal{H}_{3,i}$, $i \in [2]$ and $\mathcal{H}_{4,i}$, $i \in [5]$ as described in the statements of Lemmas 6.8 and 6.9.

We will later prove the following two claims.

Claim 6.10. *If $\beta_y + \beta_u \geqslant \beta_z + \beta_w$, then $\mathcal{H} \hookrightarrow \mathcal{H}_{3,i}$ for $i \in [2]$.*

Claim 6.11. *If $\beta_y + \beta_u \leqslant \beta_z + \beta_w$, then $\mathcal{H} \hookrightarrow \mathcal{H}_{4,i}$ for $i \in [5]$.*

If $\beta_y + \beta_u \geqslant \beta_z + \beta_w$, then $\mathcal{H}$ is γ-good by Lemma 6.8, Lemma 6.4 and Claim 6.10. Otherwise, if $\beta_y + \beta_u \leqslant \beta_z + \beta_w$, then $\mathcal{H}$ is γ-good by Lemma 6.9, Lemma 6.4 and Claim 6.11. Therefore, it remains only to prove Claims 6.10 and 6.11.

Proof of Claim 6.10. For $i \in [2]$, let $\beta_i : V(H_{3,i}) \to [0, 1]$ be defined as in Lemma 6.8.

To realise $\mathcal{H} \hookrightarrow \mathcal{H}_{3,1}$: If $\beta_y + \beta_u = 0$ (and hence, $\beta_1(y) = 0$) then let $p_1 = 0$, and otherwise let

$$p_1 = \frac{\beta_1(y)}{\beta_y + \beta_u} = \frac{1 - \beta_x}{2(\beta_y + \beta_u)} = \frac{\beta_y + \beta_u + \beta_z + \beta_w}{2(\beta_y + \beta_u)} \leqslant 1.$$

Define $\rho_1 : H_2 \to H_{3,1}$ by $\rho_1(x, z, w) = x$ and setting $\rho_1(y, u) = y$ with probability p_1, and otherwise setting $\rho_1(x, y, z, u, w) = x$.

To realise $\mathcal{H} \hookrightarrow \mathcal{H}_{3,2}$: If $\beta_z + \beta_u + \beta_w = 0$ (and hence, $\beta_2(y) = 0$) then let $p_2 = 0$, and otherwise let

$$p_2 = \frac{\beta_2(y)}{\beta_z + \beta_u + \beta_w} = \frac{1 - \beta_x}{2(\beta_z + \beta_u + \beta_w)} = \frac{\beta_y + \beta_z + \beta_u + \beta_w}{2(\beta_z + \beta_u + \beta_w)} \leqslant 1.$$

Define $\rho_2 : H_2 \to H_{3,2}$ by $\rho_2(x, y) = x$ and setting $\rho_2(z, u, w) = y$ with probability p_2, and otherwise setting $\rho_2(x, y, z, u, w) = x$. ⊡

Proof of Claim 6.11. For $i \in [5]$, let $\beta_i : V(H_{4,i}) \to [0, 1]$ be defined as in Lemma 6.9.

To realise $\mathcal{H} \hookrightarrow \mathcal{H}_{4,1}$: If $\beta_y = 0$ (and hence, $\beta_1(y) = 0$) then let $p_1 = 0$, and otherwise let

$$p_1 = \frac{\beta_1(y)}{\beta_y} = \frac{\max\{\beta_x + \beta_y - \max\{\beta_x, \beta_z\}, 0\}}{\beta_y} \leqslant \frac{\beta_y}{\beta_y} = 1.$$

If $\beta_z + \beta_u + \beta_w = 0$ (and hence, $\beta_1(z) = 0$) then let $p_1' = 0$, and otherwise let

$$p_1' = \frac{\beta_1(z)}{\beta_z + \beta_u + \beta_w} \leqslant \frac{\beta_w + \beta_z}{\beta_z + \beta_u + \beta_w} \leqslant 1.$$

Define $\rho_1 : H_0 \to H_{4,1}$ by $\rho_1(x) = x$, and independently at random with probability p_1 setting $\rho_1(y) = y$ and otherwise setting $\rho_1(y) = x$, and independently at random with probability p_1' setting $\rho_1(z, u, w) = z$ and otherwise setting $\rho_1(z, u, w) = x$.

To realise $\mathcal{H} \hookrightarrow \mathcal{H}_{4,2}$: Define $\rho_2 : H_2 \to H_{4,2}$ by $\rho_2(x, z, w) = x$ and $\rho_2(y, u) = y$.

To realise $\mathcal{H} \hookrightarrow \mathcal{H}_{4,3}$: If $\beta_z + \beta_w + \beta_u = 0$ (and hence, $\beta_3(y) = 0$) then let $p_3 = 0$, and otherwise let

$$p_3 = \frac{\beta_3(y)}{\beta_z + \beta_w + \beta_u} = \frac{\beta_y + \beta_u}{\beta_z + \beta_w + \beta_u} \leqslant 1.$$

Define $\rho_3 : H_2 \to H_{4,3}$ by $\rho_3(x, y) = x$ and setting $\rho_3(z, w, u) = y$ with probability p_3, and otherwise setting $\rho_3(z, w, u) = x$.

To realise $\mathcal{H} \hookrightarrow \mathcal{H}_{4,4}$: From (19) we have $\beta_z \leqslant \beta_4(y) \leqslant \beta_z + \beta_u$. Using this, if $\beta_u = 0$ (and hence, $\beta_4(y) = \beta_z$) then let $p_4 = 0$, and otherwise let

$$p_4 = \frac{\beta_4(y) - \beta_z}{\beta_u},$$

so that $0 \leqslant p_4 \leqslant 1$ and $p_4\beta_u + \beta_z = \beta_4(y)$. Define $\rho_4 : H_2 \to H_{4,4}$ by $\rho_4(x, y) = x$, $\rho_4(z) = y$, $\rho_4(w) = z$, and setting $\rho_4(u) = y$ with probability p_4, and otherwise setting $\rho_4(u) = x$.

To realise $\mathcal{H} \hookrightarrow \mathcal{H}_{4,5}$: Define $\rho_5 : H_2 \to H_{4,5}$ by $\rho_5(x, y) = x$ and $\rho_5(z, u, w) = y$. ⊡ □

6.3 Proofs of the three cases

We are now ready to prove Lemmas 6.6, 6.8, and 6.9, thus completing the proof of Theorem 6.5. We give these in order of difficulty, first proving Lemma 6.8, followed by Lemma 6.6 and finally Lemma 6.9, with an informal motivating discussion preceding each proof.

6.3.1 Proof of Lemma 6.8

In the following proof of Lemma 6.8, we describe the random $(\phi, i(\phi))$ realising the β_x-goodness of the set $\{\mathcal{H}_i\}_{i=1}^2$. We assume (by relabelling) that $j_t = r$ has at least average out-edge weight, and describe a simple random homomorphism based on these edge weights. As this proof is relatively easy to check we do not motivate this further, but comment on it in the motivation for the other two cases. In the proof, we will use $N_D(j)$ to denote the set of $j' \in V(D)$ with $d(j, j') + d(j', j) > 0$.

Proof of Lemma 6.8. Let $\gamma = \beta_x$. Let $1/r \ll \varepsilon \ll \alpha$, and let D be a complete looped digraph on vertex set $[r]$ with ε-complete edge weights $d(e)$, $e \in E(D)$. We will find a random $(\phi, i(\phi))$ satisfying **N1**-**N3**. By relabelling, we can assume that

$$\sum_{j\in[r-1]} d(r, j) \geqslant (\tfrac{1}{2} - 2\varepsilon) \cdot r.$$

For $j \in [r - 1]$, choose $0 \leqslant d_j \leqslant d(r, j)$ so that $\sum_{j\in[r-1]} d_j = (\frac{1}{2} - 2\varepsilon) \cdot r$.

Define $(\phi, i(\phi))$ randomly as follows. First, choose $\phi(x) \in [r - 1]$ at random, so that $\phi(x) = j$ with probability $d_j/(\frac{1}{2} - 2\varepsilon) \cdot r$. Then choose $\phi(y) \in N_D(\phi(x))$ at random so that $\phi(y) = j$ with probability $(1 - d_j)/\sum_{j\in N_D(\phi(x))}(1 - d_j)$. Set $i(\phi) = 1$ if $d(\phi(x), \phi(y)) > 0$, and set $i(\phi) = 2$ otherwise, so that, with probability 1, ϕ is a homomorphism from $H_{3,i(\phi)}$ to D. By identifying $j_t = r$, **N1** holds. We now note that, for $j \in [r - 1]$,

$$\mathbb{E}(\beta_{i(\phi)}(\phi^{-1}(j) \cap \{x\})) = \frac{(1 + \beta_x)}{2} \cdot \frac{d_j}{(\frac{1}{2} - 2\varepsilon) \cdot r} \leqslant d_j \cdot \frac{1 + \gamma + \alpha}{r},$$

and, because $\sum_{j\in N_D(j')}(1 - d_j) \geqslant (\frac{1}{2} - 3\varepsilon) \cdot r$ for any $j' \in [r - 1]$,

$$\mathbb{E}(\beta_{i(\phi)}(\phi^{-1}(j) \cap \{y\})) \leqslant \frac{(1 - \beta_x)}{2} \cdot \frac{(1 - d_j)}{(\frac{1}{2} - 3\varepsilon) \cdot r} \leqslant (1 - d_j) \cdot \frac{1 + \gamma + \alpha}{r}.$$

Thus, we deduce **N2** and **N3** hold.

□

6.3.2 Proof of Lemma 6.6

Describing our proofs of the two remaining cases directly is difficult, in part because we are finding homomorphisms to a weighted digraph so we can apply our results to a regularity partition. What we do instead is describe an embedding of a tree T into a tournament G that follows all the major steps in our proof in an analogous, but more comprehensible, way. The embedding we describe is plausible but lacks detail and ignores several subtleties that influence the formal proof – our aim is to give a step by step embedding (see steps 1 to 8 below, and also Figure 8) in a simplified set-up that, by comparison, makes the proof of Lemma 6.6 easier to follow.

For this simplified set-up, assume that we have a tree T containing a vertex t with only out-neighbours in T, such that $T - \{t\}$ consists of small components. Assume further we have a tournament G, which is larger than T, into which we are attempting to find an embedding ψ of T. As t has many out-neighbours in T, an obvious choice for $\psi(t)$ is a vertex in G with maximal out-degree – say v_t is such a vertex and set $\psi(t) = v_t$. Let $A = N^+_G(v_t)$ and $B = V(G) \setminus (A \cup \{v_t\})$. Any component of $T - \{t\}$ can be embedded into $G[A]$ to extend the embedding to cover that component (using, for example, Theorem 2.2), but there is not necessarily enough room in A to embed all of the components at once, and so the challenge is to embed components so that many vertices in B are also used.

Let $\hat{H}_1$ have vertex set $\{x, y, z\}$ and edge set $\{xy, zx\}$. For each component of $T - \{t\}$, map the out-neighbour of t to x, any vertex whose path in T from t begins with two out-edges to y, and any vertex whose path in T from t begins with an out-edge then an in-edge to z. Thus, all vertices are mapped to $V(\hat{H}_1)$. We always want to embed the vertices of T mapped to x into A, so that they are out-neighbours of v_t. If all the edges between A and B in G are directed from A to B then, given a component of $T - \{t\}$ we could embed vertices mapped to y into B and vertices mapped to z into A. On the other hand, if all the edges between A and B in G are directed from B to A then, given a component, we could embed vertices mapped to z into B and vertices mapped to y into A. In practice, we expect the edges between A and B to meet neither extreme, and that some components can be embedded with vertices in B using edges directed from A to B and some using edges directed from B to A.

It may be be that all, or almost all, the edges between A and B in G are directed from A to B. Here, it is crucial that Lemma 6.6 covers the case corresponding to components having enough vertices mapped to y in order to place plenty of vertices into B. The other extreme, where almost all the edges between A and B in G are directed from B to A, cannot occur unless B is very small, otherwise we can find a vertex in B with high enough out-degree, both in $G[B]$ and into A, so that it has higher out-degree than v_t.

In trees corresponding to Case 1 (i.e., those for which we use Lemma 6.6), more vertices are mapped to y than z, so we prefer to embed components with the vertices mapped to y embedded in B. Our embedding is then via the following steps (where we first recap the embedding of t), and also sketched in Figure 8.

1. Embed t to a vertex v_t with a largest out-neighbourhood in G, and let $A = N^+_G(v)$ and $B = V(G) \setminus (A \cup \{v_t\})$.
2. Embed as many components of $T - \{t\}$ mapped to $\{x, y, z\} \subseteq V(H_1)$ as possible with the vertex mapped to $x \in V(H_1)$ in A and either a) all other vertices embedded into B or b) the vertices mapped to y embedded into B and the vertices mapped to z embedded into A.
3. Subject to this, choose the embedding maximising the number of components satisfying a).
4. Let A_0 be the unused vertices in A and B_0 be the unused vertices in B.
5. If B_0 is small (smaller than the number of leaves of T, say), then we do not need to use these vertices and can greedily find the remaining components within A_0. Thus, we assume B_0 is not too small and that we have not found all our components.
6. A_0 then cannot be too small as we always embedded enough vertices in B compared to A.
7. There cannot be many edges directed from A_0 to B_0, for otherwise another component could be embedded with vertices mapped to y embedded in B_0.

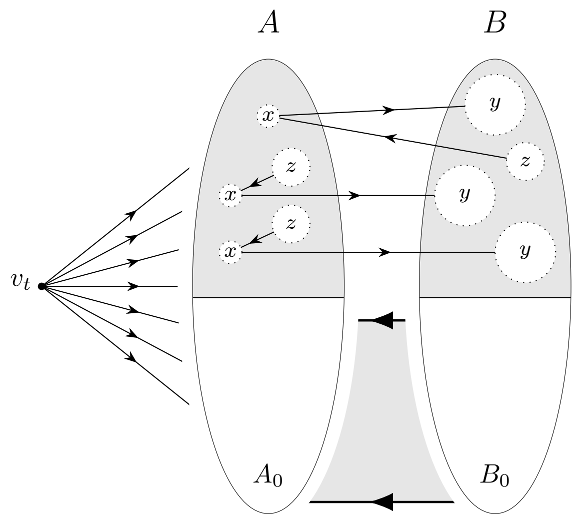

Figure 8: An example of how we aim to embed components of $T-\{t\}$ into the sets A and B in the simplified set-up. For this case, where more vertices in the components are mapped to y, we at first aim to embed these vertices mapped to y in B. Once this is no longer possible, we find that edges between the leftover vertices are mostly directed from B_0 towards A_0.

8. Using a sequence of deductions from the maximality of our component embeddings, we can then find large sets A_1 and B_1 with $A_0 \subseteq A_1 \subseteq A$ and $B_0 \subseteq B_1 \subseteq B$ so that the edges in G are mostly directed from B_1 into A_1, before concluding there is some vertex in B_1 with out-degree approximately $|A_1| + |B_1|/2$, which will be higher than the out-degree of v_t, $|A|$, giving a contradiction.

An example deduction is the following: for components satisfying b), the image of the vertex mapped to x must have few in-neighbours in B_0, for otherwise the vertices mapped to z could be embedded into B_0 to increase the number of embedded components satisfying a). Consequently, there cannot be many edges from A_0 to the images of vertices mapped to y in components satisfying b), else we could move the images of such vertices into B_0 and embed a new component using the freed up space. Thus we can add the images of vertices mapped to y in components satisfying b) to B_1.

We now describe the proof of Lemma 6.6 in comparison to these steps. For the lemma, T contains a small core $T_0 \subseteq T$ (corresponding to t above) and we have a distillation $(H_1, \{x, \bar{x}\}, \beta)$ representing an average component of $T - V(T_0)$, where H_1 is the oriented forest with vertex set $\{x, y, z, \bar{x}, \bar{z}\}$ and edge set $\{xy, zx, \bar{z}\bar{x}\}$. We have a complete looped weighted digraph D which represents a regularity partition, and choose j_t to maximise the weight on the out-edges from j_t in D (cf. 1. above). By relabelling, we assume $j_t = r$. Instead of having a partition $A \cup B$ of the other vertices $j \in V(D)$, each vertex is duplicated and lies in both A and B with a weight, representing the proportion of that vertex that is in the out- and in-neighbourhood of j_t respectively (i.e., the proportions $d(j_t, j)$ and $d(j, j_t)$).

Instead of embedding components, we find homomorphisms $\phi_1, \ldots, \phi_s$ (for some appropriate s) from H_1, before ultimately picking at random from these homomorphisms to get our required random homomorphism. These homomorphisms effectively allocate space within regularity clusters (represented by vertices of D) to embed a batch of components of $T - V(T_0)$, and we similarly aim to allocate as much space in B as possible. To do this, we first find as many as many homomorphisms $\hat{\phi}_1, \ldots, \hat{\phi}_{s_0}$ from $\hat{H}_1 = H_1[\{x, y, z\}]$ as possible so that, ideally, y and z are both embedded into B (see condition **O1** in the proof), and, failing this, at least y is embedded into B (see **O2**), while x is always embedded to A (cf. 2. and 3. above). In doing so, we always ensure that the total weight assigned to any one vertex of D is not too much (see **O3**). Maximising the number of such homomorphisms (s_0), and then the number for which **O1** is relevant, in fact will allow us to find the remaining homomorphisms $\hat{\phi}_{s_0+1}, \ldots, \hat{\phi}_s$ before extending them to homomorphisms from H_1 (cf. 5. above). We prove this by assuming it cannot be done and steadily deducing a series of claimed properties of D that ultimately allow us to find a vertex with more weight on its out-edges in D than j_t, a contradiction.

Proof of Lemma 6.6. Let $\gamma = \beta(x, \bar{x})$. Let $1/r \ll \varepsilon \ll \alpha$. We remark that $\gamma \leqslant 1$, and we may also assume that $\alpha \leqslant 1$, so we have $(1 + \gamma + \alpha) \leqslant 3$.

Let D be a complete looped digraph on vertex set $[r]$ with ε-complete edge weights $d(e)$, $e \in E(D)$. We will find a random $(\phi, i(\phi))$ satisfying **N1-N3**. By relabelling, we can assume that

$$\sum_{j\in[r-1]} d(r,j) = \max_{i\in[r]} \sum_{j\in[r]\setminus\{i\}} d(i,j) \geqslant (\tfrac{1}{2} - 2\varepsilon) \cdot r. \tag{21}$$

Take two new disjoint vertex sets $A = \{a_1, \dots, a_{r-1}\}$ and $B = \{b_1, \dots, b_{r-1}\}$. Let $\bar{D}$ be the weighted complete looped digraph on $A \cup B$ in which the edges a_ib_j, a_ia_j, b_ib_j and b_ia_j have weight $d(i, j)$.

Let s be such that $1/s \ll 1/r$. For each $i \in [r-1]$, let $w_{a_i} = d(r, i) \cdot (1 + \gamma + \alpha) \cdot s/r$ and $w_{b_i} = (1 - d(r, i)) \cdot (1 + \gamma + \alpha) \cdot s/r$. Note that

$$\sum_{v\in A\cup B} w_v \geqslant (1 + \beta(x, \bar{x}) + 7\alpha/8) \cdot s, \tag{22}$$

and

$$\sum_{v\in B} w_v = \sum_{i\in[r-1]} (1 - d(r, i)) \cdot (1 + \gamma + \alpha) \cdot s/r \overset{(21)}{<} (\tfrac{1}{2} + \tfrac{1}{2}\beta(x, \bar{x}) + 3\alpha/4) \cdot s. \tag{23}$$

We aim to find homomorphisms $\phi_1, \dots, \phi_s : H_1 \to \bar{D}$, with $\phi_i(x), \phi_i(\bar{x}) \in A$ for each $i \in [s]$ and $\sum_{i\in[s]} \beta(\phi_i^{-1}(v)) \leqslant w_v$ for every $v \in V(\bar{D})$. Then if $\phi : H_1 \to D$ is the natural homomorphism induced by a uniform random selection from $\{\phi_1, \dots, \phi_s\}$, the conclusion of the theorem will hold.

Let $\hat{H}_1 = H_1[\{x, y, z\}]$. Let $s_0 \leqslant s$ be the largest integer for which there exist homomorphisms $\hat{\phi}_1, \dots, \hat{\phi}_{s_0} : \hat{H}_1 \to \bar{D}$ and indicators $j_1, \dots, j_{s_0} \in [2]$ such that the following properties hold.

O1 For each $i \in [s_0]$ with $j_i = 1$, we have $\hat{\phi}_i(x) \in A$, $\hat{\phi}_i(y) \in B$, and $\hat{\phi}_i(z) \in B$.

O2 For each $i \in [s_0]$ with $j_i = 2$, we have $\hat{\phi}_i(x) \in A$, $\hat{\phi}_i(y) \in B$, and $\hat{\phi}_i(z) \in A$.

O3 For each $v \in A \cup B$, $\sum_{i\in[s_0]} \beta(\hat{\phi}_i^{-1}(v)) \leqslant w_v$.

Subject to this, maximise the number of $i \in [s_0]$ with $j_i = 1$. Let I_1 be the set of $i \in [s_0]$ with $j_i = 1$, and let I_2 be the set of $i \in [s_0]$ with $j_i = 2$. For each $v \in A \cup B$, let $\hat{w}_v = \sum_{i\in[s_0]} \beta(\hat{\phi}_i^{-1}(v))$, so that, by **O3**, we have $\hat{w}_v \leqslant w_v$.

Note that

$$\sum_{v\in A\cup B} \hat{w}_v = \sum_{v\in A\cup B} \sum_{i\in[s_0]} \beta(\hat{\phi}_i^{-1}(v)) = \sum_{i\in[s_0]} \beta(\hat{\phi}_i^{-1}(A \cup B)) = \beta(x, y, z) \cdot s_0. \tag{24}$$

Let B_0 be the set of $v \in B$ with $w_v - \hat{w}_v \geqslant 1$. Let A_0 be the set of $v \in A$ with $w_v - \hat{w}_v \geqslant 2$, noting that we are placing a slightly stronger condition on the definition of A_0 to enable a switching argument later on (see the end of the proof of Claim 6.16).

We now show that we are done, unless $\sum_{v\in B}(w_v - \hat{w}_v)$ is not too small.

Claim 6.12. *Either there exists a random $(\phi, i(\phi))$ satisfying* **N1-N3**, *or*

$$\sum_{v\in B} (w_v - \hat{w}_v) > (\beta(x, \bar{x}) + 3\alpha/4) \cdot s. \tag{25}$$

Proof of Claim 6.12. Suppose that $\sum_{v\in B}(w_v - \bar{w}_v) \leqslant (\beta(x, \bar{x}) + 3\alpha/4) \cdot s$. Then

$$\begin{aligned}
\sum_{v\in A} (w_v - \hat{w}_v) &= \sum_{v\in A\cup B} (w_v - \hat{w}_v) - \sum_{v\in B} (w_v - \hat{w}_v) \\
&\overset{(22),(24)}{\geqslant} (1 + \beta(x, \bar{x}) + 7\alpha/8) \cdot s - \beta(x, y, z) \cdot s_0 - (\beta(x, \bar{x}) + 3\alpha/4) \cdot s \\
&= (1 + \alpha/8) \cdot s - \beta(x, y, z) \cdot s_0 \\
&= \beta(\bar{x}, \bar{z}) \cdot s_0 + \beta(x, y, z, \bar{x}, \bar{z}) \cdot (s - s_0) + (\alpha/8) \cdot s.
\end{aligned} \tag{26}$$

Greedily extend the homomorphisms $\hat{\phi}_1, \ldots, \hat{\phi}_{s_0} : \hat{H}_1 \to \bar{D}$ to homomorphisms $\phi_1, \ldots, \phi_{s_0} : H_1 \to \bar{D}$, so that, for every $i \in [s_0]$, $\phi_i|_{\{x,y,x\}} = \hat{\phi}_i$, and $\phi_i(\bar{x}), \phi_i(\bar{z}) \in A$. Then, greedily choose homomorphisms $\phi_{s_0+1}, \ldots, \phi_s$ so that, for every $i \in [s]\setminus[s_0]$, $\phi_i(V(H_1)) = \{a_j\}$ for some $a_j \in A$. These steps are possible, while also ensuring that $\sum_{i\in[s]} \beta(\phi_i^{-1}(v)) \leqslant w_v$ for every $v \in V(\bar{D})$, due to (26). Then, by defining $(\phi, i(\phi))$ by sampling ϕ from $\phi_1, \ldots, \phi_s$ uniformly at random (identifying the result as a map $V(H_1) \to V(D)$ in the natural way) and setting $i(\phi) = 1$, we obtain a random $(\phi, i(\phi))$ satisfying **N1-N3**. ⊡

Thus, we may now assume that (25) holds, and hence also $|B_0| \geqslant 2\varepsilon r$. In particular, as $\hat{\phi}_i(y) \in B$ for each $i \in [s_0]$, we have

$$\beta(y) \cdot s_0 \leqslant \sum_{v\in B} \hat{w}_v \overset{(25)}{<} \sum_{v\in B} w_v - (\beta(x,\bar{x}) + 3\alpha/4) \cdot s \overset{(23)}{<} \frac{1}{2}(1 - \beta(x,\bar{x})) \cdot s \leqslant \beta(y) \cdot s,$$

and so we have that $s_0 < s$.

Claim 6.13. *If $i \in I_2$ and $v \in B_0$, then $d(v, \hat{\phi}_i(x)) = 0$. Hence, by* **D***, given $i \in I_2$, there is some $v \in B_0$ with $d(\hat{\phi}_i(x), v) = 1$.*

Proof of Claim 6.13. Let $i \in I_2$, $v \in B_0$, and suppose that $d(v, \hat{\phi}_i(x)) > 0$. Then we may instead set $\hat{\phi}_i(z) = v$ and $j_i = 1$ and observe that **O1**–**O3** still hold. As this increases $|I_1|$, this is a contradiction. ⊡

Given $v \in B$, let $\bar{w}_v = \beta(y) \cdot |\{i \in I_2 : \hat{\phi}_i(y) = v\}|$. We remark that $\bar{w}_v \leqslant \hat{w}_v$, and

$$\sum_{j\in[r-1]} \bar{w}_{b_j} \overset{\mathbf{O1},\mathbf{O2}}{\geqslant} \sum_{j\in[r-1]} \hat{w}_{a_j} - \beta(x) \cdot s. \tag{27}$$

Let B_y be the set of $v \in B$ for which $\bar{w}_v \geqslant 2$.

Claim 6.14. *If $i, i' \in I_2$ are such that $i \neq i'$, then $d(\hat{\phi}_{i'}(y), \hat{\phi}_i(x)) = 0$. Hence, $d(v, \hat{\phi}_i(x)) = 0$ whenever $i \in I_2$ and $v \in B_y$.*

Proof of Claim 6.14. Let $i, i' \in I_2$ be such that $i \neq i'$, and suppose that $d(\hat{\phi}_{i'}(y), \hat{\phi}_i(x)) > 0$. Let $v' = \hat{\phi}_{i'}(y)$. By Claim 6.13, there is some $v \in B_0$ such that $d(\hat{\phi}_{i'}(x), v) = 1$. Then, because $\beta(y) \geqslant \beta(z)$, we may instead set $\hat{\phi}_i(z) = v'$, $\hat{\phi}_{i'}(y) = v$, and $j_i = 1$, increasing $|I_1|$, a contradiction. ⊡

Claim 6.15. *If $i \in I_2$, then $d(v, \hat{\phi}_i(x)) > 0$ for at least εr many $v \in A_0$.*

Proof of Claim 6.15. Let $i \in I_2$. Suppose that there are fewer than εr many $v \in A_0$ for which $d(v, \hat{\phi}_i(x)) > 0$. So, using Claim 6.13 and Claim 6.14, $d(v, \hat{\phi}_i(x)) = 0$ for all but at most εr many $v \in A_0 \cup B_0 \cup B_y$. Then, by **D**, for all but at most $2\varepsilon r$ many $v \in A_0 \cup B_0 \cup B_y$, we have $d(\hat{\phi}_i(x), v) = 1$.

Let j' be such that $\hat{\phi}_i(x) = a_{j'}$. Let J be the set of $j \in [r-1]$ such that $(w_{a_j} - \hat{w}_{a_j}) + (w_{b_j} - \hat{w}_{b_j}) + \bar{w}_{b_j} \geqslant 5$. If $j \in J$, then either $a_j \in A_0$ or $b_j \in B_0 \cup B_y$. So $d(j', j) = 1$ for all but at most $2\varepsilon r$ many $j \in J$, and hence $\sum_{j\in[r]\setminus\{j'\}} d(j', j) \geqslant |J| - 3\varepsilon r$.

We remark that, for any $j \in [r-1]$, we have $(w_{a_j} - \hat{w}_{a_j}) + (w_{b_j} - \hat{w}_{b_j}) + \bar{w}_{b_j} \leqslant w_{a_j} + w_{b_j} = (1+\gamma+\alpha) \cdot s/r$. Therefore,

$$\begin{aligned}
\sum_{j\in[r]\setminus\{j'\}} d(j', j) \cdot (1 + \gamma + \alpha) \cdot s/r &\geqslant |J| \cdot (1 + \gamma + \alpha) \cdot s/r - 3\varepsilon(1 + \gamma + \alpha) \cdot s \\
&\geqslant \sum_{j\in[r-1]} [(w_{a_j} - \hat{w}_{a_j}) + (w_{b_j} - \hat{w}_{b_j}) + \bar{w}_{b_j}] - 5r - 9\varepsilon s \\
&\overset{(25)}{\geqslant} \sum_{j\in[r-1]} w_{a_j} + \sum_{j\in[r-1]} (\bar{w}_{b_j} - \hat{w}_{a_j}) + (\beta(x,\bar{x}) + \alpha/2) \cdot s \\
&\overset{(27)}{\geqslant} \sum_{j\in[r-1]} w_{a_j} + (\beta(\bar{x}) + \alpha/2) \cdot s > \sum_{j\in[r-1]} d(r, j) \cdot (1 + \gamma + \alpha) \cdot s/r,
\end{aligned}$$

contradicting (21). ⊡

Let I_Y be the set of $i \in [s_0]$ such that there exist distinct $i_0, \dots, i_\ell \in [s_0]$ with $i_0 = i$, such that

- $d(\hat{\phi}_{i_{k-1}}(x), \hat{\phi}_{i_k}(y)) > 0$ for $k \in [\ell]$, and
- $d(\hat{\phi}_{i_\ell}(x), v) > 0$ for some $v \in B_0$.

We remark that $d(\hat{\phi}_i(x), v) = 0$ whenever $i \notin I_Y, v \in B_0$, and also that $d(\hat{\phi}_i(x), \hat{\phi}_{i'}(y)) = 0$ whenever $i \notin I_Y, i' \in I_Y$.

Let A_1 be the set of $v \in A$ with $w_v - \sum_{i \in I_Y} \beta(\hat{\phi}_i^{-1}(v) \cap \{x\}) \geqslant 1$, and let B_1 be the set of $v \in B$ with $w_v - \hat{w}_v + \sum_{i \in I_Y} \beta(\hat{\phi}_i^{-1}(v) \cap \{y\}) \geqslant 1$.

Claim 6.16. *If $u \in A_1$ and $v \in B_1$, then $d(u, v) = 0$.*

Proof of Claim 6.16. Let $u \in A_1$ and $v \in B_1$, and suppose for contradiction that $d(u, v) > 0$.

If $\hat{\phi}_i(x) = u$ for some $i \notin I_Y$, then we must have $d(u, v') = 0$ for every $v' \in B_0$, else $i \in I_Y$. So in particular, we would have $v \in B_1 \setminus B_0$, and hence $\hat{\phi}_{i'}(y) = v$ for some $i' \in I_Y$. But then $d(\hat{\phi}_i(x), \hat{\phi}_{i'}(y)) > 0$ for some $i \notin I_Y$, $i' \in I_Y$, a contradiction.

Therefore, we may assume that $\sum_{i \in [s_0] \setminus I_Y} \beta(\hat{\phi}_i^{-1}(u) \cap \{x\}) = 0$, and hence

$$w_u - \hat{w}_u + \sum_{i \in I_2} \beta(\hat{\phi}_i^{-1}(u) \cap \{z\}) = w_u - \sum_{i \in [s_0]} \beta(\hat{\phi}_i^{-1}(u) \cap \{x\}) \geqslant 1. \tag{28}$$

Set $\hat{\phi}_{s_0+1}(x), \hat{\phi}_{s_0+1}(z) = u$ and $\hat{\phi}_{s_0+1}(y) = v$, and set $j_{s_0+1} = 2$.

Let $I_Z \subseteq I_2$ be a minimal set such that $\hat{\phi}_i(z) = u$ for $i \in I_Z$ and $w_u - \hat{w}_u + \sum_{i \in I_Z} \beta(\hat{\phi}_i^{-1}(u) \cap \{z\}) \geqslant 1$, noting that (28) shows such a choice is possible. By minimality of I_Z, we have $\sum_{i \in I_Z} \beta(\hat{\phi}_i^{-1}(u) \cap \{z\}) \leqslant 2$. Using Claim 6.15, choose $u_i \in A_0 \setminus \{u\}$ for $i \in I_Z$ with $d(u_i, u) > 0$.

If $w_v - \hat{w}_v \geqslant \beta(y)$, then set $\ell = 0$, $i_\ell = s_0 + 1$, and $v^\star = v$. Otherwise, we find there is some $i \in I_Y$ with $\hat{\phi}_i(y) = v$, and so let $i_0, \dots, i_\ell \in [s_0]$ be distinct with $i_0 = i$, and $v^\star \in B_0$ be such that $d(\hat{\phi}_{i_{k-1}}(x), \hat{\phi}_{i_k}(y)) > 0$ for $k \in [\ell]$ and $d(\hat{\phi}_{i_\ell}(x), v^\star) > 0$. In either case, set $v_{i_{k-1}} = \hat{\phi}_{i_k}(y)$ for $k \in [\ell]$, and set $v_{i_\ell} = v^\star$.

Now, setting $\hat{\phi}_i(z) = u_i$ for $i \in I_Z$ and $\hat{\phi}_{i_k}(y) = v_{i_k}$ for $k \in \{0\} \cup [\ell]$ yields a contradiction to the maximality of s_0, proving the claim. ⊡

Let J_A be the set of $j \in [r-1]$ with $a_j \in A_1$ and J_B be the set of $j \in [r-1]$ with $b_j \in B_1$. By Claim 6.16, J_A and J_B are disjoint. Let $j' \in J_B$ be such that $\sum_{j \in J_B} d(j', j)$ is maximised. So by Claim 6.16,

$$\sum_{j \in [r] \setminus \{j'\}} d(j', j) \geqslant |J_A| + \frac{1}{2}|J_B| - 2\varepsilon r. \tag{29}$$

Also, because $\beta(y) \geqslant \beta(x)$,

$$\begin{aligned}\sum_{v \in B} \left(w_v - \hat{w}_v + \sum_{i \in I_Y} \beta(\hat{\phi}_i^{-1}(v) \cap \{y\}) \right) &\overset{(25)}{>} (\beta(x, \bar{x}) + 3\alpha/4) \cdot s + \beta(y) \cdot |I_Y| \\ &\geqslant 2\beta(x) \cdot |I_Y| + (3\alpha/4) \cdot s = 2 \sum_{v \in A} \sum_{i \in I_Y} \beta(\hat{\phi}_i^{-1}(v) \cap \{x\}) + (3\alpha/4) \cdot s. \end{aligned} \tag{30}$$

Therefore,

$$\begin{aligned}\sum_{j \in [r] \setminus \{j'\}} d(j', j) \cdot (1 + \gamma + \alpha) \cdot s/r &\overset{(29)}{\geqslant} |J_A| \cdot (1 + \gamma + \alpha) \cdot s/r + \frac{1}{2}|J_B| \cdot (1 + \gamma + \alpha) \cdot s/r - 2\varepsilon(1 + \gamma + \alpha) \cdot s \\ &\geqslant \sum_{v \in A} (w_v - \sum_{i \in I_Y} \beta(\hat{\phi}_i^{-1}(v) \cap \{x\})) + \frac{1}{2} \sum_{v \in B} \left(w_v - \hat{w}_v + \sum_{i \in I_Y} \beta(\hat{\phi}_i^{-1}(v) \cap \{y\}) \right) - 10\varepsilon s \\ &\overset{(30)}{>} \sum_{v \in A} w_v + (3\alpha/8) \cdot s > \sum_{j \in [r-1]} d(r, j) \cdot (1 + \gamma + \alpha) \cdot s/r,\end{aligned}$$

contradicting (21).

□

6.3.3 Proof of Lemma 6.9

In this section, we will prove Lemma 6.9. Similarly as in Section 6.3.2, we will outline our strategy in a simplified setting, along with a depiction in Figure 9, so that this outline may guide the reader through the technical proof. For this, let T again be a tree containing a vertex t with only out-neighbours in T, such that $T - \{t\}$ consists of small components. Furthermore, let G be a tournament with more vertices than T into which we are attempting to find an embedding ψ of T. We proceed initially with a very similar strategy to that described at the start of Section 6.3.2, as follows.

Let v_t be a vertex in G with maximal out-degree and set $\psi(t) = v_t$. Let $A = N^+(v_t)$ and $B = V(G) \setminus (A \cup \{v_t\})$. Just as before, we now aim to embed components of $T - \{t\}$ so that they can be attached correctly to v_t but so that as many vertices as possible lie in B. For the trees relevant for Lemma 6.6, if we carefully maximised the number of components we embedded, then we covered enough vertices in B that we were able to finish the embedding by embedding the remaining components into the unused vertices in A. The problem here is that Lemma 6.9 covers trees for which this might not be possible. To see this, consider again $\hat{H}_1$ with vertex set $\{x, y, z\}$ and edge set $\{xy, zx\}$, and, for each component of $T - \{t\}$, map the out-neighbour of t to x, and map the other vertices to y or z according to the direction of the first edge of their path from t in the component as before. If all edges between A and B in G are directed from A to B then the only vertices we can embed into B are those mapped to y, and in the trees relevant to Lemma 6.9 there may be few or even none of these! That is, it simply may not be the case that we can embed T with t embedded to v_t as before.

Nevertheless, with t embedded to v_t, we attempt to embed as many components as possible subject to a careful maximisation as before. In particular, we always have the vertex mapped to x embedded into A and vertices mapped to z embedded into B. Previously, we then made a sequence of deductions that led us to find sets $A_1 \subseteq A$ and $B_1 \subseteq B$ such that almost all the edges of G were directed from B_1 to A_1, and this led to a contradiction (see step 8 in Section 6.3.2). For Lemma 6.9, we again make a (here, simpler) sequence of deductions. Roughly, if A_0 and B_0 are the vertices in A and B respectively which are not in the image of any component we have embedded, then G must have almost all the possible edges directed from A_0 into B_0 as well as certain other properties. The vertices in B_0 are then the vertices we struggled to cover when embedding components of the tree. However, if we pick a typical vertex v'_t in A_0 and try instead to embed the tree starting with embedding t to v'_t, then it is easy to cover vertices in B_0 with components in $T - \{t\}$ as most of these vertices are in the out-neighbourhood of v'_t. Even better, the vertices in $A_0 \setminus \{v'_t\}$ can also be used relatively easily as there are many edges from A_0 to B_0 and, now the embedding of t is changed, components of $T - \{t\}$ can be embedded with vertices mapped to $\{x, y\}$ embedded into B_0 and vertices mapped to z embedded into A_0 (and for the trees covered by Lemma 6.9 significantly many vertices are mapped to z). Of course, the remaining vertices of G – those which had components embedded to them – may now be hard to cover, but, from our maximised component embedding and subsequent deductions, we will have information about how we can adjust the embedded components to attach them instead to the new image, v'_t, of t while still using roughly the same vertices for that component. This is not always simple, and we allocate some additional vertices in B_0 to each embedded component before finding a new version of that component using the old vertices for that component and possibly also the additional vertices allocated from B_0. That we can allocate some additional vertices in this fashion relies on the fact that G is larger than T, so we can afford not to use every vertex in G in the final embedding.

The portioning of the vertices in B_0 to allow the rearrangement and reattachment of the components requires quite some delicacy, resulting ultimately in the choice of the functions $\beta_1, \dots, \beta_5$ in Lemma 6.9. This choice can be motivated, but only at some length and difficulty, so instead we concentrate on writing a proof that can be directly verified. As in Section 6.3.2, in the actual setting, with a tree T with small core T_0 in mind, given a distillation of an average component, we find homomorphisms to a weighted looped digraph which will represent a regularity partition. We then pick from these homomorphisms randomly to get a random homomorphism. These homomorphisms again essentially allocate room for components of $T - V(T_0)$ in the regularity clusters corresponding to the weighted digraph D.

To find these homomorphisms, we begin by selecting a vertex $j_t \in V(D)$ with at least the average amount of weight on its out-edges. By relabelling, we assume $j_t = r$. Similarly to the proof of Lemma 6.6, we again duplicate vertices of D to get weighted vertex sets A and B representing the proportion of each vertex that is in the out- and in-neighbourhood of j_t respectively, and aim to find homomorphisms which allocate plenty of space in B. To do this, we find maximally many homomorphisms from $H_{4,1}$ to D satisfying certain conditions (**P1**–**P4**), and further optimising over two additional conditions (stated just after **P4**), to get

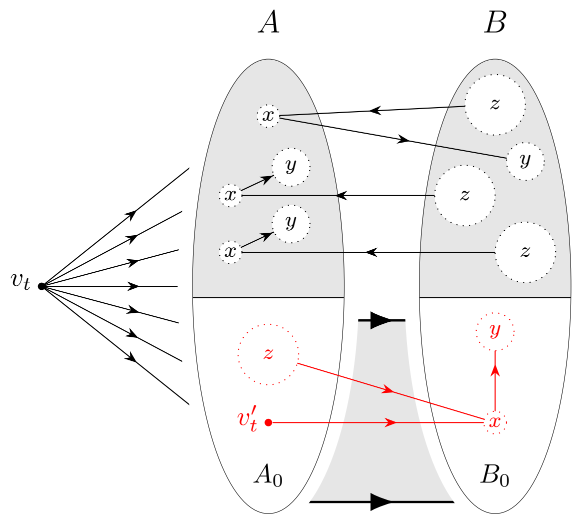

Figure 9: In our simplified set-up, for the case corresponding to Lemma 6.9, we may instead begin by aiming to embed vertices of $T - \{t\}$ mapped to z in B. The component in the lower half of the diagram illustrates how leftover space in A_0 and B_0 may be used efficiently via a new embedding v'_t for t.

homomorphisms $\phi_1, \ldots, \phi_{s_0}$. Assuming we do not have enough homomorphisms to easily find the remaining ones we wanted (see Claim 6.18), we make a sequence of deductions based on this assumption and the maximality (Claim 6.17 and Claim 6.19) about the weight distribution on D. Letting $A_0 \subseteq A$ and $B_0 \subseteq B$ be the sets of vertices with unallocated weight, an example deduction is that almost all of the weight on the edges of D between A_0 and B_0 is on edges directed from A_0 to B_0 (for otherwise we would have found at least one more homomorphism). We then choose a new vertex $j'_t \in [r]$, with $a_{j'_t} \in A_0$, to use instead of $j_t = r$ by selecting it so that $a_{j'_t}$ has sufficient out-weight to certain vertices in the image of the homomorphisms we have found (just before Claim 6.20) – this will allow us later to rearrange these homomorphisms to embed x into the out-neighbourhood of j'_t instead of j_t. While it may not necessarily be possible to ensure the desired conditions on j'_t hold with respect to every homomorphism chosen so far, Claim 6.20 confirms that there is a choice of j'_t with the required conditions for rearrangement holding with respect to most of them – by relabelling, these will be the homomorphisms $\phi_1, \ldots, \phi_{s_1}$. We add a dummy vertex q with weight $\beta_1(q) = \max\{\beta_x, \beta_z\}$ to $H_{4,1}$, and then extend $\phi_1, \ldots, \phi_{s_1}$ to homomorphisms $\psi_1, \ldots, \psi_{s_1}$ also covering q (see **S1**, **S2** and Claim 6.21). The role played by q is to reserve additional weight in the out-neighbourhood of j'_t, which may be required for the desired rearrangement. Each ψ_i then represents some allocated space, for which we then find homomorphisms $\phi_{i,j}$ from $H_{4,2}, H_{4,3}$ or $H_{4,4}$ which use a proportion of this space (in either Claim 6.22 or Claim 6.23) – a proportion matching the space a tree must cover in the tournament following an eventual application of the lemma (that is, the proportion $|T|/|G|$). The homomorphisms $\phi_{i,j}$ potentially leave plenty of weight remaining on vertices in A_0 and B_0, but, as noted above in the simplified setting, these vertices are easiest to use for new homomorphisms as there is a lot of weight on edges from j'_t to B_0 and from B_0 to A_0 in D. We therefore find more homomorphisms, this time from $H_{4,5}$, to use enough of this weight (that is, so that the proportion of the weight used is at least the matching proportion $|T|/|G|$). Finally, then, we have a collection of homomorphisms from which to pick our random homomorphism to complete the proof.

Proof of Lemma 6.9. Let $\gamma = \max\{\beta_x, \beta_z\}$. Let $1/r \ll \varepsilon \ll \alpha$. We remark that $\gamma \leqslant 1$, and we may also assume that $\alpha \leqslant 1$, so we have $(1 + \gamma + \alpha) \leqslant 3$.

Let D be a complete looped digraph on vertex set $[r]$ with ε-complete edge weights $d(e)$, $e \in E(D)$. We will find a random $(\phi, i(\phi))$ satisfying **N1-N3**. By relabelling, we can assume that

$$\sum_{j \in [r-1]} d(r, j) \geqslant (\tfrac{1}{2} - 2\varepsilon) \cdot r. \tag{31}$$

Take two new disjoint vertex sets $A = \{a_1, \dots, a_{r-1}\}$ and $B = \{b_1, \dots, b_{r-1}\}$. Let $\bar{D}$ be the weighted complete looped digraph on $A \cup B$ in which the edges a_ib_j, a_ia_j, b_ib_j and b_ia_j have weight $d(i, j)$.

Let s be such that $1/s \ll 1/r$. For each $i \in [r-1]$, let $w_{a_i} = d(r, i) \cdot (1 + \gamma + \alpha) \cdot s/r$ and $w_{b_i} = (1 - d(r, i))(1 + \gamma + \alpha) \cdot s/r$. Note that

$$\sum_{v \in A \cup B} w_v \geqslant (1 + \gamma + \alpha/2) \cdot s, \tag{32}$$

and

$$\sum_{v \in A} w_v = \sum_{i \in [r-1]} d(r, i) \cdot (1 + \gamma + \alpha) \cdot s/r \overset{(31)}{\geqslant} \frac{1}{2}(1 + \gamma + \alpha/2) \cdot s. \tag{33}$$

Let $s_0 \leqslant s$ be the largest integer for which there exist homomorphisms $\phi_1, \dots, \phi_{s_0} : H_{4,1} \to \bar{D}$ and indicators $j_1, \dots, j_{s_0} \in [2]$ such that the following properties hold.

P1 For each $i \in [s_0]$, $d(\phi_i(y), \phi_i(z)) + d(\phi_i(z), \phi_i(y)) > 0$.

P2 For each $i \in [s_0]$ with $j_i = 1$, we have $\phi_i(x) \in A$, $\phi_i(y) \in B$, and $\phi_i(z) \in B$.

P3 For each $i \in [s_0]$ with $j_i = 2$, we have $\phi_i(x) \in A$, $\phi_i(y) \in A$, and $\phi_i(z) \in B$.

P4 For each $v \in A \cup B$, $\sum_{i \in [s_0]} \beta_1(\phi_i^{-1}(v)) \leqslant w_v$.

Given s_0, take $\phi_1, \dots \phi_{s_0}$ and $j_1, \dots, j_{s_0}$ such that the number of $i \in [s_0]$ with $j_i = 1$ is maximised. For each $v \in A \cup B$, let $\tilde{w}_v = \sum_{i \in [s_0]} \beta_1(\phi_i^{-1}(v))$, so that, by **P4**, we have $\tilde{w}_v \leqslant w_v$.

Note that

$$\sum_{v \in A \cup B} \tilde{w}_v = \sum_{v \in A \cup B} \sum_{i \in [s_0]} \beta_1(\phi_i^{-1}(v)) = \sum_{i \in [s_0]} \beta_1(\phi_i^{-1}(A \cup B)) = s_0. \tag{34}$$

Let A_0 be the set of $v \in A$ with $w_v - \tilde{w}_v \geqslant 1$. Let B_0 be the set of $v \in B$ with $w_v - \tilde{w}_v \geqslant 1$. Subject to the value of s_0, and the value of $|\{i \in [s_0] : j_i = 1\}|$, assume that

$$|\{(i, v) : i \in [s_0], v \in A_0, d(\phi_i(y), v) > 0\}| \tag{35}$$

is minimised.

Suppose that $s_0 < s$, for otherwise we are done by letting $i(\phi) = 1$ and picking ϕ from $\{\phi_i : i \in [s]\}$ uniformly at random. Note that, as $1 - \gamma \leqslant 1 - \beta_x = \beta_y + \beta_u + \beta_z + \beta_w \leqslant 2(\beta_z + \beta_w)$, we have

$$\begin{aligned} \frac{1+\gamma}{2} - \beta_1(x, y) &= \frac{1+\gamma}{2} - (1 - \beta_1(z)) = \min\{\beta_w + \beta_z, 1 - \beta_z\} - \frac{1-\gamma}{2} \\ &\geqslant \min\left\{\frac{1-\gamma}{2}, 1 - \gamma\right\} - \frac{1-\gamma}{2} = 0. \end{aligned} \tag{36}$$

Therefore,

$$\begin{aligned} \sum_{v \in A_0} (w_v - \tilde{w}_v) &\geqslant \sum_{v \in A} (w_v - \tilde{w}_v) - r = \sum_{v \in A} w_v - r - \sum_{i \in I_1} \beta_1(x) - \sum_{i \in I_2} \beta_1(x, y) \\ &\geqslant \sum_{v \in A} w_v - r - \beta_1(x, y) \cdot s_0 \overset{(33)}{\geqslant} \frac{1}{2}(1 + \gamma + \alpha/4) \cdot s - \beta_1(x, y) \cdot s_0 \\ &= \left(\frac{1+\gamma}{2}\right) \cdot (s - s_0) + \left(\frac{1+\gamma}{2} - \beta_1(x, y)\right) \cdot s_0 + (\alpha/8) \cdot s \\ &\overset{(36)}{\geqslant} \left(\frac{1+\gamma}{2}\right) \cdot (s - s_0) + (\alpha/8) \cdot s, \end{aligned} \tag{37}$$

and hence, $|A_0| \geqslant (\alpha/32) \cdot r$. In addition, we have

$$\begin{aligned} \sum_{v \in A_0 \cup B_0} (w_v - \tilde{w}_v) &\geqslant \sum_{v \in A \cup B} (w_v - \tilde{w}_v) - 2r \overset{(32),(34)}{\geqslant} (1 + \gamma + \alpha/2) \cdot s - s_0 - 2r \\ &\geqslant (s - s_0) + (\gamma + \alpha/4) \cdot s. \end{aligned} \tag{38}$$

We now show there are few pairs from B_0 to A_0 with positive weight.

Claim 6.17. *If $u \in A_0$ and $v \in B_0$, then $d(v,u) = 0$.*

Proof of Claim 6.17. If $u \in A_0$, $v \in B_0$, are such that $d(v,u) > 0$, then we may set $\phi_{s_0+1}(x), \phi_{s_0+1}(y) = u$, $\phi_{s_0+1}(z) = v$, and $i_{s_0+1} = 2$, contradicting the maximality of s_0. ⊡

We now show that we are done unless $\sum_{v\in B_0}(w_v - \tilde{w}_v)$ is not too small.

Claim 6.18. *Either there exists a random $(\phi, i(\phi))$ satisfying* **N1**-**N3**, *or*

$$\sum_{v\in B_0}(w_v - \tilde{w}_v) \geqslant (\gamma + \alpha/16)\cdot s + \beta_1(y)\cdot(s - s_0). \tag{39}$$

Proof of Claim 6.18. Suppose that (39) does not hold. If $\sum_{v\in B_0}(w_v - \tilde{w}_v) \leqslant (\gamma+\alpha/8)\cdot s$, then set $s_1 = s_0$. Otherwise, let $s_1 \in [s]\setminus[s_0]$ be maximal such that $\sum_{v\in B_0}(w_v - \tilde{w}_v) \geqslant (\gamma+\alpha/16)\cdot s + \beta_1(y)\cdot(s_1 - s_0)$. Note that $s_1 < s$, else (39) holds. By considering the cases $s_1 = s_0$ and $s_1 > s_0$ separately (using the maximality of s_1 in the latter), we deduce that

$$\sum_{v\in B_0}(w_v - \tilde{w}_v) \leqslant (\gamma + \alpha/8)\cdot s + \beta_1(y)\cdot(s_1 - s_0), \tag{40}$$

and hence,

$$\sum_{v\in A_0}(w_v - \tilde{w}_v) \overset{(38),(40)}{\geqslant} \beta_1(x,z)\cdot(s_1 - s_0) + (s - s_1) + (\alpha/8)\cdot s.$$

Therefore, using Claim 6.17, we may greedily choose homomorphisms $\phi_{s_0+1},\dots,\phi_s : H_{4,1} \to \bar{D}$ such that the following properties hold.

Q1 For each $i \in [s_1]\setminus[s_0]$, we have $\phi_i(x) \in A$, $\phi_i(y) \in B$, and $\phi_i(z) \in A$.

Q2 For each $i \in [s]\setminus[s_1]$ we have $\phi_i(x) \in A$, $\phi_i(y) \in A$, and $\phi_i(z) \in A$.

Q3 For each $v \in A\cup B$, $\sum_{i\in[s]}\beta(\phi_i^{-1}(v)) \leqslant w_v$.

Thus, by defining $(\phi, i(\phi))$ by sampling ϕ from $\phi_1,\dots,\phi_s$ uniformly at random (identifying the result as a map $V(H_{4,1}) \to V(D)$ in the natural way) and setting $i(\phi) = 1$, we obtain a random $(\phi, i(\phi))$ satisfying **N1**-**N3**. ⊡

Thus, we may now assume that (39) holds, and hence also $|B_0| \geqslant (\alpha/64)\cdot r$. Next we show that each $i \in [s_0]$ satisfies (at least) one of two properties, **R1** or **R2**.

Claim 6.19. *For each $i \in [s_0]$, at least one of the following holds.*

R1 *$d(\phi_i(z), v) = 0$ for every $v \in A_0$.*

R2 *$d(v, \phi_i(x)) = 0$ for every $v \in B_0$, $d(\phi_i(y), v) = 0$ for every $v \in A_0$, and $j_i = 1$.*

Proof of Claim 6.19. Let $i \in [s_0]$. Assume there is some $u \in A_0$ with $d(\phi_i(z), u) > 0$, for otherwise **R1** holds.

Now, if there is some $v \in B_0$ with $d(v, \phi_i(x)) > 0$, set $\phi_{s_0+1}(x), \phi_{s_0+1}(y) = u$ and $\phi_{s_0+1}(z) = \phi_i(z)$, and then switch $\phi_i(z) = v$. This contradicts the maximality of s_0.

Thus, $d(v, \phi_i(x)) = 0$ for every $v \in B_0$. As $|B_0| \geqslant (\alpha/64)\cdot r$, we may now fix some $v \in B_0$ with $d(\phi_i(x), v) > 0$ and $d(v, \phi_i(z)) + d(\phi_i(z), v) > 0$, by **D**. Then we must have that $j_i = 1$, else we could switch $\phi_i(y) = v$ and $j_i = 1$, contradicting the maximality of $|\{i \in [s_0] : j_i = 1\}|$.

Now, note that, by Claim 6.17, $|\{u' \in A_0 : d(v,u') > 0\}| = 0$. Therefore, if $|\{u' \in A_0 : d(\phi_i(y), u') > 0\}| > 0$, we could switch $\phi_i(y) = v$ to reduce $\sum_{i'\in[s_0]}|\{u' \in A_0 : d(\phi_{i'}(y), u') > 0\}|$ while leaving A_0 unmodified, a contradiction to the minimisation of (35). Thus, $|\{u' \in A_0 : d(\phi_i(y), u') > 0\}| = 0$, and hence **R2** holds. ⊡

Given $u \in A_0$, let $I_1(u)$ be the set of $i \in [s_0]$ which satisfy **R1** and are such that $d(u, \phi_i(z)) = 1$ and $\phi_i^{-1}(u) = \emptyset$, and let $I_2(u)$ be the set of $i \in [s_0]\setminus I_1(u)$ which satisfy **R2** and are such that $d(u, \phi_i(y)) = 1$ and $\phi_i^{-1}(u) = \emptyset$. Pick $j_t' \in \{j \in [r-1] : a_j \in A_0\}$ such that $|I_1(a_{j_t'})| + |I_2(a_{j_t'})|$ is maximised, and set $I_1 = I_1(a_{j_t'})$, $I_2 = I_2(a_{j_t'})$. By relabelling, we may assume that $I_1 \cup I_2 = [s_1]$ for some $s_1 \leqslant s_0$.

Claim 6.20. $s_1 \geqslant (1-\sqrt{\varepsilon})s_0$.

Proof of Claim 6.20. If $i \in [s_0]$ satisfies **R1**, then, by **D**, $d(u, \phi_i(z)) = 1$ holds for all but at most εr many $u \in A_0$. Similarly, if $i \in [s_0]$ satisfies **R2**, then $d(u, \phi_i(y)) = 1$ holds for all but at most εr many $u \in A_0$. In addition, for each $i \in [s_0]$, we have $\phi_i^{-1}(u) \neq \emptyset$ for at most three $u \in A_0$. Therefore, for every $i \in [s_0]$, we have $i \in I_1(u) \cup I_2(u)$ for all but at most $2\varepsilon r + 3$ many $u \in A_0$. In particular,

$$\sum_{u \in A_0} (|I_1(u)| + |I_2(u)|) \geqslant s_0 \cdot (|A_0| - 3\varepsilon r).$$

Noting that $|I_1| + |I_2| \geqslant \frac{1}{|A_0|} \sum_{u \in A_0}(|I_1(u)| + |I_2(u)|)$ and $|A_0| \geqslant (\alpha/32) \cdot r$, we deduce $s_0 - |I_1| - |I_2| \leqslant \sqrt{\varepsilon} s_0$, and hence the claim. ⊡

Let $B_1 \subseteq B_0$ be the subset of vertices $v \in B_0$ with $d(a_{j'_t}, v) = 1$, and note, by Claim 6.17 and **D**, that $|B_0 \setminus B_1| \leqslant \varepsilon r$. Let q be a new vertex disjoint from $V(H_{4,1})$ and add it to $H_{4,1}$ to get $H'_{4,1}$. Let $\beta_1(q) = \gamma$. Let $s'_1 \leqslant s_1$ be maximal for which, for each $i \in [s'_1]$, we can define $\psi_i(q) \in B_1$ such that the following hold.

S1 For each $i \in [s'_1]$ and $u \in \{x, y, z\}$, $d(\psi_i(q), \phi_i(u)) + d(\phi_i(u), \psi_i(q)) > 0$.

S2 For each $v \in B$,
$$\tilde{w}_v + \gamma \cdot |\{i \in [s'_1] : \psi_i(q) = v\}| \leqslant w_v. \tag{41}$$

Let $\psi_i : H'_{4,1} \to \bar{D}$ be defined by this choice of $\psi_i(q)$ and by setting $\psi_i(v) = \phi_i(v)$ for each $v \in V(H_{4,1})$.

Claim 6.21. $s'_1 = s_1$.

Proof of Claim 6.21. Suppose for contradiction that $s'_1 < s_1$. Let B'_1 be the set of $v \in B_1$ such that $d(v, \phi_{s'_1+1}(u)) + d(\phi_{s'_1+1}(u), v) > 0$ for every $u \in \{x, y, z\}$. Because the edge weights $d(e)$, $e \in E(D)$, are ε-complete and $|B_0 \setminus B_1| \leqslant \varepsilon r$, we have $|B_0 \setminus B'_1| \leqslant 4\varepsilon r$. Therefore, we have

$$\sum_{v \in B'_1} (w_v - \tilde{w}_v - \gamma \cdot |\{i \in [s'_1] : \psi_i(q) = v\}|) \geqslant \sum_{v \in B_0} (w_v - \tilde{w}_v) - \sum_{v \in B'_1 \setminus B_0} w_v - \gamma \cdot s'_1$$
$$\overset{(39)}{\geqslant} (\gamma + \alpha/16) \cdot s - 12\varepsilon \cdot s - \gamma \cdot s'_1 \geqslant (\alpha/32) \cdot s \geqslant \gamma \cdot r$$

Thus, there is some $v \in B'_1$ such that $\tilde{w}_v + \gamma \cdot |\{i \in [s'_1] : \psi_i(q) = v\}| + \gamma \leqslant w_v$. But then setting $\psi_{s'_1+1}(q) = v$ contradicts the maximality of s'_1. ⊡

Let m satisfy $\varepsilon \ll 1/m \ll \alpha$.

Claim 6.22. *For each $i \in I_1$, there are homomorphisms $\phi_{i,j}$ and indicators $k_{i,j} \in \{2, 3\}$, $j \in [m-1]$, such that, for each $j \in [m-1]$, $\phi_{i,j}$ is a homomorphism from $H_{4,k_{i,j}}$ to $\bar{D}[\psi_i(V(H'_{4,1}))]$, and the following hold.*

T1 $d(a_{j'_t}, \phi_{i,j}(x)) = 1$.

T2 *For each $v \in \psi_i(V(H'_{4,1}))$, $\beta_1(\psi_i^{-1}(v)) \geqslant \sum_{j \in [m-1]} \beta_{k_{i,j}}(\phi_{i,j}^{-1}(v))/m$.*

Proof of Claim 6.22. Fixing $i \in I_1$, for each $j' = 1, \ldots, m-1$ in turn, choose a homomorphism $\phi_{i,j'}$ from $H_{4,2}$ or $H_{4,3}$ to $\bar{D}[\psi_i(V(H'_{4,1}))]$ such that $\beta_1(\psi_i^{-1}(v)) \geqslant \sum_{j \in [j']} \beta_{k_{i,j}}(\phi_{i,j}^{-1}(v))/m$ for each $v \in \psi_i(H'_{4,1})$, $\phi_{i,j'}(x) \in \{\psi_i(z), \psi_i(q)\}$ and $\phi_{i,j'}(y) \in \{\psi_i(x), \psi_i(y)\}$. Then **T1** holds, through the definition of either I_1 (if $\phi_{i,j'}(x) = \psi_i(z)$) or B_1 (if $\phi_{i,j'}(x) = \psi_i(q)$).

Note that this is possible as $\beta_2 = \beta_3$, so that, as

$$\beta_1(z, q) = \min\{\beta_w + \beta_z, 1 - \beta_z\} + \max\{\beta_x, \beta_z\} \geqslant \min\{\beta_w + \beta_z + \beta_x, 1\} = \beta_2(x),$$

there is enough room in $\{\psi_i(z), \psi_i(q)\}$ for $\phi_{i,j'}(x)$, and, as

$$\beta_1(x, y) = 1 - \beta_1(z) \geqslant 1 - (\beta_w + \beta_z) = \beta_x + \beta_u + \beta_y \geqslant \beta_2(y).$$

there is enough room in $\{\psi_i(x), \psi_i(y)\}$ for $\phi_{i,j'}(y)$. We also use that there is weight in at least one direction between any pair from $\{\psi_i(z), \psi_i(q)\}$ and $\{\psi_i(x), \psi_i(y)\}$, where the direction of the edge gives whether we embed $H_{4,2}$ or $H_{4,3}$. ⊡

Claim 6.23. *For each $i \in I_2$, there are homomorphisms $\phi_{i,j}$, $j \in [m-1]$, such that, for each $j \in [m-1]$, $\phi_{i,j}$ is a homomorphism from $H_{4,4}$ to $\bar{D}[\psi_i(V(H'_{4,1}))]$, and the following holds.*

U1 $d(a_{j'_t}, \phi_{i,j}(x)) = 1$.

U2 *For each $v \in \psi_i(V(H'_{4,1}))$, $\beta_1(\psi_i^{-1}(v)) \geqslant \sum_{j\in[m-1]} \beta_4(\phi_{i,j}^{-1}(v))/m$.*

Proof of Claim 6.23. Fixing $i \in I_2$, for each $j' = 1, \ldots, m-1$ in turn, choose a homomorphism $\phi_{i,j'}$ from $H_{4,4}$ to $\bar{D}[\psi_i(V(H'_{4,1}))]$ such that $\beta_1(\psi_i^{-1}(v)) \geqslant \sum_{j\in[j']} \beta_4(\phi_{i,j}^{-1}(v))/m$ for each $v \in \psi_i(H'_{4,1})$, $\phi_{i,j'}(x) \in \{\psi_i(y), \psi_i(q)\}$, $\phi_{i,j'}(y) = \psi_i(x)$ and $\phi_{i,j'}(z) = \psi_i(z)$. Then **U1** holds, through the definition of either I_2 (if $\phi_{i,j'}(x) = \psi_i(y)$) or B_1 (if $\phi_{i,j'}(x) = \psi_i(q)$).

Note that this is possible using the following. As $\beta_1(y)+\beta_1(q) = \beta_1(y)+\gamma = \max\{\beta_x+\beta_y, \gamma\} = \max\{\beta_x+\beta_y, \beta_z\} \geqslant \beta_4(x)$, there is enough room in $\{\psi_i(y), \psi_i(q)\}$ for $\phi_{i,j'}(x)$. As $\beta_1(z) \geqslant \min\{\beta_w, 1-\beta_z\} \geqslant \beta_w = \beta_4(z)$, there is enough room in $\psi_i(z)$ for $\phi_{i,j'}(z)$. Finally, recall from (20), that $\beta_4(y) \leqslant \beta_1(x)$, so there is enough room in $\psi_i(x)$ for $\phi_{i,j'}(y)$. Because $d(\psi_i(x), \psi_i(y)), d(\psi_i(z), \psi_i(x)) > 0$ (as $\psi_i : H'_{4,1} \to \bar{D}$ is a homomorphism) and $d(\psi_i(x), \psi_i(q)) > 0$ (by **R2** and **S1**), we also have that $\phi_{i,j'} : H_{4,4} \to \bar{D}[\psi_i(V(H'_{4,1}))]$ is a homomorphism, as required. ⊡

For each $i \in I_2$ and $j \in [m-1]$, let $k_{i,j} = 4$. For each $v \in A \cup B$, let

$$\begin{aligned} \hat{w}_v &= \frac{1}{m} \sum_{i\in[s_1]} \sum_{j\in[m-1]} \beta_{k_{i,j}}(\phi_{i,j}^{-1}(v)) \overset{\textbf{T2},\textbf{U2}}{\leqslant} \sum_{i\in[s_1]} \beta_1(\psi_i^{-1}(v)) \\ &= \sum_{i\in[s_1]} \beta_1(\phi_i^{-1}(v)) + \gamma \cdot |\{i \in [s_1] : \psi_i(q) = v\}| = \tilde{w}_v + \gamma \cdot |\{i \in [s_1] : \psi_i(q) = v\}| \overset{(41)}{\leqslant} w_v. \end{aligned} \tag{42}$$

We note that

$$\begin{aligned} \sum_{v\in A_0\cup B_1} (w_v - \hat{w}_v) &\geqslant \sum_{v\in A_0\cup B_0} (w_v - \hat{w}_v) - 3\varepsilon \cdot s \\ &\overset{(42)}{\geqslant} \sum_{v\in A_0\cup B_0} (w_v - \tilde{w}_v) - 3\varepsilon \cdot s - \sum_{v\in A_0\cup B_0} \gamma \cdot |\{i \in [s_1] : \psi_i(q) = v\}| \\ &\overset{(38)}{\geqslant} s - s_0 + (\gamma + \alpha/8) \cdot s - \gamma \cdot s_1 \overset{\text{Claim 6.20}}{\geqslant} (1+\gamma) \cdot (s - s_1) + (\alpha/16) \cdot s. \end{aligned} \tag{43}$$

Furthermore,

$$\begin{aligned} \sum_{v\in B_1} (w_v - \hat{w}_v) \geqslant \sum_{v\in B_0} (w_v - \hat{w}_v) - 3\varepsilon \cdot s &\overset{(42)}{\geqslant} \sum_{v\in B_0} (w_v - \tilde{w}_v) - 3\varepsilon \cdot s - \sum_{v\in B_0} \gamma \cdot |\{i \in [s_1] : \psi_i(q) = v\}| \\ &\overset{(39)}{\geqslant} \beta_1(y) \cdot (s - s_0) + (\gamma + \alpha/32) \cdot s - \gamma \cdot s_1 \overset{\text{Claim 6.20}}{\geqslant} (\gamma + \beta_1(y)) \cdot (s - s_1) + (\alpha/64)s. \end{aligned} \tag{44}$$

Take a maximal set $J \subseteq ([s] \times [m]) \setminus ([s_1] \times [m-1])$ for which there are homomorphisms $\phi_{i,j} : H_{4,5} \to \bar{D}[A_0 \cup B_1]$, $(i,j) \in J$ such that the following hold.

V1 For each $(i,j) \in J$, $\phi_{i,j}(x) \in B_1$.

V2 For each $v \in A \cup B$, $\sum_{(i,j)\in J} \beta_5(\phi_{i,j}^{-1}(v))/m \leqslant w_v - \hat{w}_v$.

Subject this choice of J, maximise

$$|\{(i,j) \in J : \phi_{i,j}(y) \in A_0\}|. \tag{45}$$

Claim 6.24. $J = ([s] \times [m]) \setminus ([s_1] \times [m-1])$.

Proof of Claim 6.24. Suppose, for contradiction, that there is some $(i', j') \in (([s] \times [m]) \setminus ([s_1] \times [m-1])) \setminus J$. We must then have, for each $v \in B_1$, that $\sum_{(i,j)\in J} \beta_5(\phi_{i,j}^{-1}(v))/m \geqslant w_v - \hat{w}_v - 1/m$, for otherwise we can take the homomorphism $\phi_{i',j'}$ sending $H_{4,5}$ to v.

Therefore, we have

$$\sum_{v\in A_0}\left(w_v-\hat{w}_v-\tfrac{1}{m}-\sum_{(i,j)\in J}\beta_5(\phi_{i,j}^{-1}(v))/m\right)\overset{(43)}{\geqslant}\gamma\cdot(s-s_1)+(\alpha/16)\cdot s-|A_0|/m-|B_1|/m\geqslant(\alpha/32)\cdot s.$$

Thus, there must be at least $2\varepsilon r$ vertices $v\in A_0$ with $\sum_{(i,j)\in J}\beta_5(\phi_{i,j}^{-1}(v))/m\leqslant w_v-\hat{w}_v-1/m$. Therefore, by Claim 6.17, if there is some $(i,j)\in J$ with $\phi_{i,j}(y)\notin A_0$, then we could move $\phi_{i,j}(y)$ into A_0, and increase the value of (45). Thus, we must have $\phi_{i,j}(y)\in A_0$ for each $(i,j)\in J\setminus([s_1]\times[m-1])$.

Therefore, using that $\beta_5(x)\leqslant\gamma+\beta_1(y)$ and $|J|\leqslant(s-s_1)m+s$,

$$\begin{aligned}0\geqslant\sum_{v\in B_1}\left(w_v-\hat{w}_v-\tfrac{1}{m}-\sum_{(i,j)\in J}\beta_5(\phi_{i,j}^{-1}(v))/m\right)&\geqslant\sum_{v\in B_1}(w_v-\hat{w}_v)-|B_1|/m-((s-s_1)m+s)\cdot\beta_5(x)/m\\&\geqslant\sum_{v\in B_1}(w_v-\hat{w}_v)-(s-s_1)\cdot(\gamma+\beta_1(y))-s/m-|B_1|/m\overset{(44)}{>}0,\end{aligned}$$

a contradiction. Therefore, $J=([s]\times[m])\setminus([s_1]\times[m-1])$. ⊡

For each $(i,j)\in([s]\times[m])\setminus([s_1]\times[m-1])$, let $k_{i,j}=5$. Select $(\phi,i(\phi))$ uniformly at random from $(\phi_{i,j},k_{i,j})$, $(i,j)\in[s]\times[m]$. □

6.4 Proof of Theorem 3.4

We are now ready to complete this section by proving Theorem 3.4. To give a brief overview of this proof, we again turn to our simplified situation: assume we are trying to embed a tree T in a tournament G and suppose we have $t\in V(T)$ so that $T-\{t\}$ consists of small components. Unlike in Sections 6.3.2 and 6.3.3, t can have both in- and out-neighbours in T. Let T^+ and T^- be the trees covering the edges of T, intersecting only on t, so that t has only out-neighbours in T^+ and only in-neighbours in T^-. As G is a tournament with distinctly more vertices than T, each vertex $v\in V(G)$ either has enough out-neighbours in G that we can embed the components of $T^+-\{t\}$ greedily into $N_G^+(v)$ or the components of $T^--\{t\}$ greedily into $N_G^-(v)$. If we partition $V(G)=V^+\cup V^-$ so that the former holds for vertices in V^+ and the latter holds for vertices in V^-, then, from this partition, either $G[V^+]$ will be large enough to embed T^- (using our previous methods) or $G[V^-]$ will be large enough to embed T^+. By directional duality, we can assume that the latter case holds. This allows us to find a copy of T^+ in G with t embedded to v_t, a vertex of G which has enough in-neighbours for us to greedily embed the components of $T^--\{t\}$. Of course, some of these in-neighbours may be occupied already by the embedding of T^+, but, by embedding T^+ in such a way to cover minimally many of these in-neighbours we will have there are enough in-neighbours to embed the components of $T^--\{t\}$, and complete the embedding.

For Theorem 3.4, we do this in the setting of distillations, random homomorphisms and a weighted looped digraph D. We ultimately wish to find a random homorphism from H, where H is the oriented forest with vertex and edge sets given by

$$\begin{aligned}V(H)&=\{x^+,y^+,z^+,u^+,w^+,\bar{x}^+,\bar{z}^+,\bar{u}^+,\bar{w}^+,x^-,y^-,z^-,u^-,w^-,\bar{x}^-,\bar{z}^-,\bar{u}^-,\bar{w}^-\},\\E(H)&=\{x^+y^+,z^+x^+,z^+u^+,w^+z^+,\bar{z}^+\bar{x}^+,\bar{z}^+\bar{u}^+,\bar{w}^+\bar{z}^+,y^-x^-,x^-z^-,u^-z^-,z^-w^-,\bar{x}^-\bar{z}^-,\bar{u}^-\bar{z}^-,\bar{z}^-\bar{w}^-\}.\end{aligned}$$

For each $\diamond\in\{+,-\}$, let $X^\diamond=\{x^\diamond,\bar{x}^\diamond\}$, and let $X=X^+\cup X^-$.

Note that $H_0\cong H[\{x^+,y^+,z^+,u^+,w^+,\bar{x}^+,\bar{z}^+,\bar{u}^+,\bar{w}^+\}]$. Therefore, we will assume equality here by letting, for example, $x^+=x$. In addition, let $H_0'=H-V(H_0)$, and note that H_0' is itself isomorphic to a copy of H_0 with all edges reversed. Here H_0 and H_0' correspond to T^+ and T^- in the sketch above.

Instead of partitioning $V(G)$ as $V^+\cup V^-$ we partition $V(D)$ as $J^+\cup J^-$ in a similar manner, and assume, by directional duality, that J^- is large enough that we can apply Theorem 6.5 to get a random homomorphism of H_0 into $D[J^-]$ satisfying **W1**–**W3** below (comparable to the embedding of T^+ into $G[V^-]$ in the sketch above) before minimising a certain property (comparable to the embedding of T^+ using minimally many in-neighbours of v_t above). We then use the minimisation of the random homomorphism to extend it to cover H_0', so that we have a random homorphism of H into D. Finally, we adjust this random homomorphism to get the additional condition **E4** which we dropped for Theorem 6.5, completing the proof.

Proof of Theorem 3.4. For $\diamond \in \{+,-\}$, let

$$\lambda^\diamond = \beta(x^\diamond, y^\diamond, z^\diamond, u^\diamond, w^\diamond, \bar{x}^\diamond, \bar{z}^\diamond, \bar{u}^\diamond, \bar{w}^\diamond)$$

so that $\lambda^+ + \lambda^- = 1$, and let

$$\gamma^\diamond = \max\{\beta(x^\diamond, \bar{x}^\diamond), \beta(z^\diamond, \bar{z}^\diamond)\}/\lambda^\diamond.$$

For $\diamond \in \{+,-\}$, define

$$r^\diamond = \left\lceil \frac{\lambda^\diamond(1+\gamma^\diamond+\alpha/16)}{1+\gamma+\alpha/4} \cdot r \right\rceil,$$

so that $r^+ + r^- \leqslant (1-\varepsilon)\cdot r$.

Let K be an ε-almost tournament with vertex set $[r]$, such that $d(i,j) \geqslant 1/2$ whenever $i \to_K j$. Partition $[r]$ as $J^+ \cup J^-$ such that $d_K^\diamond(j) \geqslant r^\diamond$ whenever $j \in J^\diamond$. Note that we either have $|J^+| \geqslant r^-$ or $|J^-| \geqslant r^+$. By directional duality, we may assume that $|J^-| \geqslant r^+$.

Let $\beta_0 : V(H_0) \to [0,1]$ be given by $\beta_0(v) = \beta(v)/\lambda^+$, and note that $\sum_{v\in V(H_0)} \beta_0(v) = 1$ and $\beta_0(y^+) \geqslant \beta_0(x^+)$. By Theorem 6.5, the distillation $\mathcal{H}_0 = (H_0, X^+, \beta_0)$ is γ^+-good. Therefore, because $\beta = \lambda^+ \cdot \beta_0$ and $\lambda^+ \cdot \frac{1+\gamma^++\alpha/16}{r^+} \leqslant \frac{1+\gamma^++\alpha/4}{r}$, there exists some $j_t \in J^-$ and random $\psi : H_0 \to D[J^-]$ such that the following hold.

W1 With probability 1, we have that ψ is a homomorphism from H_0 to D, and that $j_t \notin \psi(X^+)$.

W2 For each $j \in [r]$, $\mathbb{E}(\beta(\psi^{-1}(j))) \leqslant \frac{1+\gamma+\alpha/4}{r}$.

W3 For each $j \in [r]$, $\mathbb{E}(\beta(\psi^{-1}(j) \cap X^+)) \leqslant d(j_t, j) \cdot \frac{1+\gamma+\alpha/4}{r}$.

Fix such a $j_t \in J^-$. Let $A = N_K^+(j_t)$ and $B = N_K^-(j_t)$. Take a random $\psi : H_0 \to D$ satisfying **W1-W3** so that $\mathbb{E}(\beta(\psi^{-1}(B)))$ is minimised.

Claim 6.25. $|B| \cdot \frac{1+\gamma+\alpha/4}{r} - \mathbb{E}(\beta(\psi^{-1}(B))) - \mathbb{E}(\beta(\psi^{-1}(B)\cap X^+)) \geqslant \lambda^- + \beta(X^-) + \alpha/8$.

Proof of Claim 6.25. First, if $|A| \cdot \frac{1+\gamma+\alpha/4}{r} - \mathbb{E}(\beta(\psi^{-1}(A)) - \mathbb{E}(\beta(\psi^{-1}(A)\cap X^+)) \leqslant 0$, then

$$\begin{aligned} |B| \cdot \tfrac{1+\gamma+\alpha/4}{r} - \mathbb{E}(\beta(\psi^{-1}(B))) - \mathbb{E}(\beta(\psi^{-1}(B)\cap X^+)) &\geqslant |D| \cdot \tfrac{1+\gamma+\alpha/4}{r} - 3\varepsilon - \lambda^+ - \beta(X^+) \\ &\geqslant 1+\gamma+\alpha/8 - \lambda^+ - \beta(X^+) \geqslant \lambda^- + \beta(X^-) + \alpha/8. \end{aligned}$$

So we may assume that $|A| \cdot \frac{1+\gamma+\alpha/4}{r} - \mathbb{E}(\beta(\psi^{-1}(A))) - \mathbb{E}(\beta(\psi^{-1}(A)\cap X^+)) > 0$, else the claim is proven. In particular, we may assume there is some $j \in A$ such that, if

$$\begin{aligned} p &:= \frac{1}{2}\left(\frac{1+\gamma+\alpha/4}{r} - \mathbb{E}(\beta(\psi^{-1}(j))) - \mathbb{E}(\beta(\psi^{-1}(j)\cap X^+))\right) \\ &\leqslant d(j_t, j) \cdot \frac{1+\gamma+\alpha/4}{r} - \mathbb{E}(\beta(\psi^{-1}(j)\cap X^+)), \end{aligned}$$

then $0 < p < 1$. In addition, we may assume that $\mathbb{E}(\beta(\psi^{-1}(B))) > 0$, else the claim follows immediately from the definition of J^-. Then, however, if we define $\hat{\psi} : H_0 \to D$ by setting $\hat{\psi}(V(H_0)) = j$ with probability p and sampling ψ otherwise, we find $\hat{\psi}$ satisfies **W1-W3**, but $\mathbb{E}(\beta(\hat{\psi}^{-1}(B))) = (1-p)\cdot\mathbb{E}(\beta(\psi^{-1}(B)))$, a contradiction to the minimality of $\mathbb{E}(\beta(\psi^{-1}(B)))$. □

Consider the random $\hat{\phi} : H \to D$ defined by sampling ψ to determine $\hat{\phi}|_{V(H_0)}$, and independently choosing $\hat{\phi}(V(H_0')) \in B$ at random so that

$$\mathbb{P}(\hat{\phi}(V(H_0')) = j) = p_j := \frac{\frac{1+\gamma+\alpha/4}{r} - \mathbb{E}(\beta(\psi^{-1}(j))) - \mathbb{E}(\beta(\psi^{-1}(j)\cap X^+))}{|B| \cdot \frac{1+\gamma+\alpha/4}{r} - \mathbb{E}(\beta(\psi^{-1}(B))) - \mathbb{E}(\beta(\psi^{-1}(B)\cap X^+))}.$$

We remark that for every $j \in B$, we have

$$p_j \cdot \max\{\lambda^-, 2\beta(X^-)\} \leqslant p_j \cdot (\lambda^- + \beta(X^-)) \overset{\text{Claim 6.25}}{\leqslant} \tfrac{1+\gamma+\alpha/4}{r} - \mathbb{E}(\beta(\psi^{-1}(j))) - \mathbb{E}(\beta(\psi^{-1}(j)\cap X^+)). \quad (46)$$

Therefore, if $j \in B$, we have

$$\mathbb{E}(\beta(\hat{\phi}^{-1}(j))) = \mathbb{E}(\beta(\psi^{-1}(j))) + p_j \cdot \lambda^- \overset{(46)}{\leqslant} \frac{1+\gamma+\alpha/4}{r},$$
$$\mathbb{E}(\beta(\hat{\phi}^{-1}(j) \cap X^+)) \overset{\mathbf{W3}}{\leqslant} d(j_t, j) \cdot \frac{1+\gamma+\alpha/4}{r},$$
$$\mathbb{E}(\beta(\hat{\phi}^{-1}(j) \cap X)) \leqslant \mathbb{E}(\beta(\psi^{-1}(j) \cap X^+)) + p_j \cdot \beta(X^-) \overset{(46)}{\leqslant} \frac{1}{2}\left(\frac{1+\gamma+\alpha/4}{r}\right) \leqslant d(j, j_t) \cdot \frac{1+\gamma+\alpha/4}{r},$$

whereas if $j \in [r] \setminus B$, we have

$$\mathbb{E}(\beta(\hat{\phi}^{-1}(j))) = \mathbb{E}(\beta(\psi^{-1}(j))) \overset{\mathbf{W2}}{\leqslant} \frac{1+\gamma+\alpha/4}{r},$$
$$\mathbb{E}(\beta(\hat{\phi}^{-1}(j) \cap X^-)) = 0 \leqslant d(j, j_t) \cdot \frac{1+\gamma+\alpha/4}{r},$$
$$\mathbb{E}(\beta(\hat{\phi}^{-1}(j) \cap X)) = \mathbb{E}(\beta(\psi^{-1}(j) \cap X^+)) \overset{\mathbf{W3}}{\leqslant} d(j_t, j) \cdot \frac{1+\gamma+\alpha/4}{r}.$$

Take $\bar{\phi} : H \to D$ with $\sum_{e \in E(H)} \mathbb{P}(|\bar{\phi}(e)| = 1)$ minimal, such that the following properties hold.

X1 With probability 1, we have that $\bar{\phi}$ is a homomorphism from H to D, and that $j_t \notin \bar{\phi}(\{x^+, \bar{x}^+, x^-, \bar{x}^-\})$.

X2 For each $j \in [r]$, $\mathbb{E}(\beta(\bar{\phi}^{-1}(j))) \leqslant \frac{1+\gamma+\alpha/2}{r} - \frac{\alpha^2}{r} \cdot \sum_{e \in E(H)} \mathbb{P}(|\bar{\phi}(e)| = 1)$.

X3 For each $j \in [r]$, either

X3.1 $\mathbb{E}(\beta(\bar{\phi}^{-1}(j) \cap X^+)) \leqslant d(j_t, j) \cdot \frac{1+\gamma+\alpha/2}{r}$ and $\mathbb{E}(\beta(\bar{\phi}^{-1}(j) \cap X)) \leqslant d(j, j_t) \cdot \frac{1+\gamma+\alpha/2}{r}$, or

X3.2 $\mathbb{E}(\beta(\bar{\phi}^{-1}(j) \cap X^-)) \leqslant d(j, j_t) \cdot \frac{1+\gamma+\alpha/2}{r}$ and $\mathbb{E}(\beta(\bar{\phi}^{-1}(j) \cap X)) \leqslant d(j_t, j) \cdot \frac{1+\gamma+\alpha/2}{r}$.

$\bar{\phi}$ is well defined, as we may take $\bar{\phi} = \hat{\phi}$.

We will shortly prove the following claim.

Claim 6.26. $\sum_{e \in E(H)} \mathbb{P}(|\bar{\phi}(e)| = 1) \leqslant \varepsilon^{1/4}$.

Then, if we take ϕ to be $\bar{\phi}$ conditioned on the event $\{|\bar{\phi}(e)| = 2$ for every $e \in E(H)\}$, we have

$$\mathbb{E}(\beta(\phi^{-1}(j) \cap \{v\})) \leqslant (1 - \varepsilon^{1/4})^{-1} \cdot \mathbb{E}(\beta(\bar{\phi}^{-1}(j) \cap \{v\}))$$

for every $j \in [r], v \in V(H)$, thus ϕ satisfies the conclusion of the theorem. So it only remains to prove Claim 6.26.

Proof of Claim 6.26. Suppose for contradiction that $\sum_{e \in E(H)} \mathbb{P}(|\bar{\phi}(e)| = 1) \geqslant \varepsilon^{1/4}$, so in particular there is some $e \in E(H)$ with $\mathbb{P}(|\bar{\phi}(e)| = 1) \geqslant \varepsilon^{1/4}/|E(H)|$. Let $e = (v_1, v_2)$. For $i \in [2]$, let H_i be the component of $H - e$ containing v_i. Assume, by directional duality, that $X \cap V(H_2) = \emptyset$.

Note that, because $\beta(v_1, v_2) \geqslant 2\mu$, **X2** implies that $\mathbb{P}(\bar{\phi}(e) = \{j\}) \leqslant \frac{2}{\mu r}$ for every $j \in [r]$. So, if J is the set of $j \in [r]$ for which $\mathbb{P}(\bar{\phi}(e) = \{j\}) \geqslant \frac{\sqrt{\varepsilon}}{r}$, then $|J| \geqslant \sqrt{\varepsilon} \cdot r$, else we find $\mathbb{P}(|\bar{\phi}(e)| = 1) = \sum_{j \in [r]} \mathbb{P}(\bar{\phi}(e) = \{j\}) \leqslant \frac{2\sqrt{\varepsilon}}{\mu} + \sqrt{\varepsilon} < \varepsilon^{1/4}/|E(H)|$.

Let m be the number of possible homomorphisms $H \to D$. Choose homomorphisms $\phi_j, j \in J$ such that $\phi_j(e) = \{j\}$ and $\mathbb{P}(\bar{\phi} = \phi_j) \geqslant \frac{\sqrt{\varepsilon}}{mr}$ for every $j \in J$ (these can be found as the edge weights of D are ε-complete).

Set $s = \lceil 1/\alpha^3 \rceil$. Let $j_1, \ldots, j_{s+1} \in J$ be distinct such that $d(j_i, j_{i+1}) > 0$ for every $i \in [s]$.

Let $k \in [2]$ be random, with distribution coupled with $\bar{\phi}$ such that $\mathbb{P}(k = 2) = \frac{(s+1)\sqrt{\varepsilon}}{mr}$, and that $\mathbb{P}(\bar{\phi} = \phi_{j_i} \mid k = 2) = \frac{\sqrt{\varepsilon}}{mr}$ for every $i \in [s+1]$.

Define a random $\psi : H \to D$ as follows. Sample $(\bar{\phi}, k)$, choose $i \in [s+1]$ uniformly at random, and set

$$\psi(v) = \begin{cases} \bar{\phi}(v) & \text{if } k = 1, \\ \phi_{j_i}(v) & \text{if } k = 2,\ i = s+1, \\ \phi_{j_i}(v) & \text{if } k = 2,\ i \in [s],\ v \in V(H_1), \\ \phi_{j_{i+1}}(v) & \text{if } k = 2,\ i \in [s],\ v \in V(H_2). \end{cases}$$

We then find $\mathbb{P}(|\psi(e)| = 1) = \mathbb{P}(|\bar{\phi}(e)| = 1) - \frac{s\sqrt{\varepsilon}}{mr}$, yet $\mathbb{E}(\beta(\psi^{-1}(j) \cap X)) = \mathbb{E}(\beta(\bar{\phi}^{-1}(j) \cap X))$ and $\mathbb{E}(\beta(\psi^{-1}(j))) \leqslant \mathbb{E}(\beta(\bar{\phi}^{-1}(j))) + \frac{\sqrt{\varepsilon}}{mr}$ for every $j \in [r]$, a contradiction to the minimality of $\sum_{e \in E(H)} \mathbb{P}(|\bar{\phi}(e)| = 1)$. ⊡ □

References

[1] N. Alon and A. Shapira. Testing subgraphs in directed graphs. *Journal of Computer and System Sciences*, 69:354–382, 2004.

[2] A. Benford and R. Montgomery. Trees with few leaves in tournaments. *Journal of Combinatorial Theory, Series B*, 155:141–170, 2022.

[3] F. Dross and F. Havet. On the unavoidability of oriented trees. *Journal of Combinatorial Theory, Series B*, 151:83–110, 11 2021.

[4] A. El Sahili. Trees in tournaments. *Journal of Combinatorial Theory, Series B*, 92:183–187, 2004.

[5] A. Frieze and M. Karoński. **Introduction to Random Graphs**. Cambridge University Press, 2015.

[6] R. Häggkvist and A. Thomason. Trees in tournaments. *Combinatorica*, 11:123–130, 1991.

[7] F. Havet. Trees in tournaments. *Discrete Mathematics*, 243(1):121–134, 2002.

[8] F. Havet. On unavoidability of trees with k leaves. *Graphs and Combinatorics*, 19:101–110, 2003.

[9] F. Havet and S. Thomassé. Median orders of tournaments: A tool for the second neighborhood problem and Sumner's conjecture. *Journal of Graph Theory*, 35(4):244–256, 2000.

[10] F. Havet and S. Thomassé. Oriented Hamiltonian paths in tournaments: A proof of Rosenfeld's conjecture. *Journal of Combinatorial Theory, Series B*, 78(2):243–273, 2000.

[11] A. Kathapurkar and R. Montgomery. Spanning trees in dense directed graphs. *Journal of Combinatorial Theory, Series B*, 156:223–249, 2022.

[12] D. Kühn, R. Mycroft, and D. Osthus. An approximate version of Sumner's universal tournament conjecture. *Journal of Combinatorial Theory, Series B*, 101(6):415–447, 2011.

[13] D. Kühn, R. Mycroft, and D. Osthus. A proof of Sumner's universal tournament conjecture for large tournaments. *Proceedings of the London Mathematical Society*, 102(4):731–766, 2010.

[14] R. Mycroft and T. Naia. Unavoidable trees in tournaments. *Random Structures & Algorithms*, 53(2):352–385, 2018.

[15] K. Reid and N. Wormald. Embedding oriented n-trees in tournaments. *Studia Sci. Math. Hungar*, 18(2-4):377–387, 1983.

[16] A. Thomason. Paths and cycles in tournaments. *Transactions of the American Mathematical Society*, 296:167–180, 1986.