Remarks on the second neighborhood problem

Dror Fidler *   Raphael Yuster †

Abstract

The second neighborhood conjecture of Seymour asserts that for any orientation $G = (V, E)$, there exists a vertex $v \in V$ so that $|N^+(v)| \leq |N^{++}(v)|$. The conjecture was resolved by Fisher for tournaments. In this paper we prove the second neighborhood conjecture for several additional classes of dense orientations. We also prove some approximation results, and reduce an asymptotic version of the conjecture to a finite case.

1 Introduction

Throughout this paper all directed graphs (digraphs) are finite orientations. As usual, we denote by $N_G^+(v)$ (resp. $N_G^-(v)$) the set of out-neighbors (resp. in-neighbors) of a vertex v in the digraph G. We denote by $N_G^{++}(v)$ (resp. $N_G^{--}(v)$) the set of vertices at distance 2 from v (resp. to v). We also denote $d_G^+(v) = |N_G^+(v)|$, $d_G^-(v) = |N_G^-(v)|$, $d_G^{++}(v) = |N_G^{++}(v)|$, and $d_G^{--}(v) = |N_G^{--}(v)|$. We will omit the subscript if the digraph is clear from the context.

We say that a vertex v has the *second neighborhood property* (SNP) if $d^{++}(v) \geq d^+(v)$. Dean (see [1]) conjectured that every tournament has a vertex with the SNP. Seymour conjectured a more general statement (see [1]).

Conjecture 1.1 (The Second Neighborhood Conjecture (SNC)) *Every orientation has a vertex with the SNP.*

This conjecture, posed over a decade ago, is still open. In 1996, Fisher [2] solved Dean's conjecture, thus asserting the SNC for tournaments. Fisher's proof uses a certain probability distribution on the vertices, and a basic Markovian argument. Another proof of Dean's conjecture was given in 2000 by Havet and Thomassé [4]. Their proof uses a tool called *median orders*. Furthermore, they proved that if a tournament has no sink then there are at least two vertices with the SNP. In Section 2 we give a brief summary of the proof from [4] and show a more general result: In every

*Department of Mathematics, University of Haifa, Haifa 31905, Israel.
E–mail: dror_f@hotmail.com
†Department of Mathematics, University of Haifa, Haifa 31905, Israel.
E–mail: raphy@research.haifa.ac.il

tournament where each vertex has a non-negative real weight there is a vertex with the *weighted* SNP. This weighted version is needed in the proof of one of our theorems.

For orientations other than tournaments, only partial results have been achieved. Chen, Shen and Yuster have shown in [3], that in every orientation there is a vertex v such that $d^{++}(v) \geq \gamma d^{+}(v)$, where $\gamma \approx \frac{2}{3}$. Kaneko and Locke [5] proved the SNC for orientations with minimum out-degree at most 6.

In Section 3 we obtain several results that assert the SNC for other non-trivial classes of dense orientations. We can summarize our results in the following theorem:

Theorem 1.2 *The SNC holds for the following classes of orientations.*
(i) Orientations of n-vertex graphs with minimum degree $n-2$.
(ii) Orientations of complete graphs missing a star.
(iii) Orientations of pseudocliques (graphs obtained from the complete graph by deleting the set of edges of a smaller clique).

Although all of the results mentioned here are for orientations that are, in a sense, close to being a tournament, one should not consider them as simple extensions of Fisher's Theorem. In fact, even proving the SNC for orientations obtained from a tournament by deleting two arcs is *not* an easy exercise, even given Fisher's Theorem as a *black box*.

In Section 4 we prove an extension of the approximation result from [3]. We obtain better approximation ratios in the case of orientations of K_{k+1}-free graphs.

Theorem 1.3 *Let $G = (V, E)$ be an orientation of a K_{k+1}-free graph. Then there exists $v \in V$ such that $d^{++}(v) \geq \gamma d^{+}(v)$, where γ is the real root of $f(x) = \frac{2k-2}{k}x^3 + \frac{k-2}{k}x^2 - 1$.*

For example, in the case $k = 3$ we have $\gamma \geq 0.8324$.

In Section 5 we show how to reduce a general asymptotic version of the SNC to a finite case.

Theorem 1.4 *For every $\epsilon > 0$, there is a finite set $F = F(\epsilon)$ of orientations, so that if the SNC holds for all elements in F, then* every *orientation with n vertices has a vertex v with $d^{++}(v) \geq d^{+}(v)(1-\epsilon) - \epsilon n$.*

Theorem 1.4 is proved using probabilistic arguments.

2 Weighted median orders

Given an orientation $G = (V, E)$, a *median order* of G, is an order of V such that the number $|\{(v_i, v_j) \in E, i < j\}|$ is maximized. A *local median order* of a tournament T, is an order that satisfies the *feedback property*, which means that for every interval $[v_i, v_j]$, $|N^{+}_{T|_{[v_i,v_j]}}(v_i)| \geq |N^{-}_{T|_{[v_i,v_j]}}(v_i)|$, $|N^{+}_{T|_{[v_i,v_j]}}(v_j)| \leq |N^{-}_{T|_{[v_i,v_j]}}(v_j)|$. It is easy to see that a median order is also a local median order.

The last vertex in a median order is called the *feed vertex.* Median orders provide an inductive tool that Havet and Thomassé [4] utilized to simplify and strengthen Fisher's result. As we shall require median orders in our proof of Theorem 1.2, we now present their short proof of the SNC for tournaments in a more general setting.

Given an orientation with a non-negative weight function $w : V \to \Re_+$, we say that a vertex v has the *weighted SNP* if

$$\sum_{u \in N^+(v)} w(u) \quad \le \sum_{u \in N^{++}(v)} w(u).$$

An order τ of V is a *weighted local median order* if for every interval $[u, v]$ of τ

$$\sum_{x \in N^+_D(u) \cap [u,v]} w(x) \quad \ge \sum_{y \in N^-_D(u) \cap [u,v]} w(y) \tag{1}$$

$$\sum_{x \in N^-_D(v) \cap [u,v]} w(x) \quad \ge \sum_{y \in N^+_D(v) \cap [u,v]} w(y). \tag{2}$$

Every tournament has a weighted local median order. Indeed, given an order of the vertices, we define the weight of the order as follows. Each arc (u, v) for which u precedes v in the order contributes $w(u)$ to the weight of the order. Clearly, an order with maximum weight is a weighted local median order.

Proposition 2.1 *In every tournament with non-negative vertex weights there is a vertex with the weighted SNP.*

Proof: Let $L = (x_1, \ldots, x_n)$ be a weighted local median order of the tournament T. Distinguish between two types of vertices of $N^-(x_n)$. A vertex $x_j \in N^-(x_n)$ is *good*, if there exists $x_i \in N^+(x_n)$, with $i < j$, such that $(x_i, x_j) \in E$. Otherwise x_j is *bad.* Denote the set of good vertices of (T, L) by G_L. We show by induction on n that x_n satisfies $\sum_{v \in N^+(x_n)} w(v) \le \sum_{v \in G_L} w(v)$. The case $n = 1$ holds vacuously. Assume now that $n > 1$. If there is no bad vertex, we have $\sum_{v \in G_L} w(v) = \sum_{v \in N^-(x_n)} w(v)$. Moreover, the feedback property ensures that $\sum_{v \in N^+(x_n)} w(v) \le \sum_{v \in N^-(x_n)} w(v)$, so the conclusion holds. Now assume that there exists a bad vertex x_i. Choose it minimal with respect to its index i. Denote by G^u_L the set $G_L \cap [x_{i+1}, x_n]$, by G^d_L, the set $G_L \cap [x_1, x_i]$, by $N^+(x_n)^u$ the set $N^+(x_n) \cap [x_{i+1}, x_n]$, and by $N^+(x_n)^d$ the set $N^+(x_n) \cap [x_1, x_i]$. Applying the induction hypothesis to the restriction of (T, L) to $[x_{i+1}, x_n]$ gives directly that $\sum_{v \in G^u_L} w(v) \ge \sum_{v \in N^+(x_n)^u} w(v)$, since every good vertex of this restriction is, a fortiori, a good vertex of (T, L). By the minimality i, every vertex of $[x_1, x_{i-1}]$ is either in G^d_L or in $N^+(x_n)^d$. Furthermore, since x_i is bad, we have $N^+(x_n)^d \subseteq N^+(x_i) \cap [x_1, x_i]$ and equivalently, $G^d_L \supseteq N^-(x_i) \cap [x_1, x_i]$. The feedback property applied to $[x_1, x_i]$ gives:

$$\sum_{v \in G^d_L} w(v) \ge \sum_{v \in N^-(x_i) \cap [x_1, x_i]} w(v) \ge \sum_{v \in N^+(x_i) \cap [x_1, x_i]} w(v) \ge \sum_{v \in N^+(x_n)^d} w(v).$$

Thus, $\sum_{v \in G_L^d} w(v) \geq \sum_{v \in N^+(x_n)^d} w(v)$ and $\sum_{v \in G_L^u} w(v) \geq \sum_{v \in N^+(x_n)^u} w(v)$, so the induction hypothesis holds for n. Since $G_L \subset N^{++}(x_n)$, the result follows. ■

In fact, using a more complicated argument, the following stronger result is proved in [4].

Theorem 2.2 *In every tournament without a sink there are at least two vertices with the SNP.*

We note that there is an *expensive* reduction from the non-weighted version proved in [4] to our weighted version proved in Proposition 2.1. We may assume that all weights are positive (simply delete the vertices with zero weight). Since the rationals are dense in the reals, it suffices to prove the result for rational weights. In turn, by multiplying with a common denominator, we may assume that all weights are naturals. Given a tournament $T = (V, E)$ with a weight function $w : V \to N$, create a new *non-weighted* tournament T' as follows. Replace v with a *transitive* tournament on $w(v)$ vertices, denoted T_v. If $(u, v) \in E$ then orient all arcs from T_u to T_v. Now, suppose $x \in T_v$ is a vertex with the SNP in T'. Note that we may always assume that x is the sink of T_v. Then, v satisfies the weighted SNP in T.

3 Some classes of dense orientations

In this section, the restricted SNC to several classes of orientations is proved. There are many classes of orientations in which SNC is trivial. For instance, SNC for bipartite orientations is immediate by examining the vertex with minimum out-degree. On the other hand, already a tournament missing two arcs is not a trivial exercise, even *given* that the SNC holds for tournaments. Extensions of this class include tournaments missing the arcs of a sub-tournament, tournaments missing a matching, tournaments missing the arcs of a star, all are treated in this section. Note that the class of tournaments missing a matching consists of all orientations of undirected graphs with minimum degree $n-2$. SNC remains unproven for orientations of graphs with minimum degree $n-3$ (complete graphs missing vertex-disjoint cycles and paths).

Let $G = (V, E)$ be an orientation. Let C_G denote the set of all tournaments obtained by orienting the missing edges of G. Clearly, $|C_G| = 2^{\binom{n}{2} - |E|}$. Each of the cases in Theorem 1.2 is proved in a different subsection.

3.1 Orientations of graphs with minimum degree $n - 2$

An orientation of an n-vertex graph with minimum degree $n-2$ is a tournament missing a matching. Let, therefore, $G = (V, E)$ be a tournament missing a matching, and let M denote the set of missing edges. Thus, $xy \in M$ if $(x, y) \notin E$ and $(y, x) \notin E$. Notice that $|M| \leq n/2$. An *induced* 4-cycle (a_1, a_2, b_1, b_2) is called a *losing cycle* if $b_2 \notin N^{++}(a_1)$ and $a_2 \notin N^{++}(b_1)$. We say that $a_1 b_1 \in M$ *loses* to $a_2 b_2 \in M$. The *dependency digraph* of G, denoted Δ_G, is defined as follows. The vertex set of Δ_G is M. There is an arc (xy, wz) in Δ_G if and only if xy loses to wz. Notice that it is possible for Δ_G to have cycles of length 2 (trivial cycles) as well as isolated vertices (trivial paths). However, more can be said.

Lemma 3.1 *The maximum out-degree in Δ_G is 1, and the maximum in-degree in Δ_G is 1. Thus, Δ_G is composed of vertex-disjoint paths and cycles.*

Proof: Assume the contrary. Consider first the case of a vertex $wz \in M$ with two out-neighbors in Δ_G, say, x_1y_1 and x_2y_2. Notice that since M is a matching, all the six vertices x_1, y_1, x_2, y_2, w, z are distinct. By the definitions of losing cycles we may assume that $(w, x_1) \in E$, $(z, y_2) \in E$, $y_2 \notin N^{++}(w)$ and $x_1 \notin N^{++}(z)$. Consider the arc of G connecting x_1 and y_2. If its direction is (x_1, y_2) then $y_2 \in N^{++}(w)$, a contradiction. If its direction is (y_2, x_1) then $x_1 \in N^{++}(z)$, a contradiction. Consider next the case of a vertex $xy \in M$ with two in-neighbors in Δ_G, say, w_1z_1 and w_2z_2. Notice that since M is a matching, all the six vertices x, y, w_1, w_2, z_1, z_2 are distinct. By the definition of losing cycles we may assume that $(w_1, x) \in E$, $(z_2, y) \in E$, $y \notin N^{++}(w_1)$ and $x \notin N^{++}(z_2)$. Consider the arc of G connecting w_1 and z_2. If its direction is (w_1, z_2) then $y \in N^{++}(w_1)$, a contradiction. If its direction is (z_2, w_1) then $x \in N^{++}(z_2)$, a contradiction. ■

Let $C = (a_1b_1, \dots, a_kb_k)$ be a path or a cycle in Δ_G (possibly $k = 1$). Namely, $(a_i, a_{i+1}, b_i, b_{i+1})$ forms a losing cycle for $i = 1, \dots, k-1$. We shall orient the missing edges $a_1b_1, \dots, a_kb_k$ according to the following rule. In case C is a path we proceed as follows. Suppose there are two vertices $a_0, b_0 \in V \setminus \{a_1, b_1\}$ so that $(a_0, a_1) \in E$, $(b_0, b_1) \in E$, $b_1 \notin N^+(a_0) \cup N^{++}(a_0)$ and $a_1 \notin N^+(b_0) \cup N^{++}(b_0)$. This implies that $a_0b_0 \in M$, that (a_0, a_1, b_0, b_1) form a losing cycle and a_0b_0 loses to a_1b_1, contradicting the fact that a_1b_1 starts a path in Δ_G. Thus, we may assume, w.l.o.g., that for each $v \in V \setminus \{a_1, b_1\}$ with $(v, a_1) \in E$ we have $b_1 \in N^+(v) \cup N^{++}(v)$. In this case we orient the missing edge as (a_1, b_1) (otherwise we orient it (b_1, a_1)). If we oriented (a_1, b_1) we orient all others as (a_i, b_i) for $i = 2, \dots, k$. If we oriented (b_1, a_1) we orient all others as (b_i, a_i) for $i = 2, \dots, k$. If C is a cycle then we orient all missing edges as (a_i, b_i) for $i = 1, \dots, k$.

By performing the procedure detailed in the last paragraph for all paths and cycles in Δ_G we obtain a completion of G into a tournament $T_0 \in C_G$. We need two additional lemmas.

Lemma 3.2 *If C is an odd cycle then (a_k, a_1, b_k, b_1) is a losing cycle. If C is an even cycle then (a_k, b_1, b_k, a_1) is a losing cycle.*

Proof: Assume first that C is an odd cycle. Consider the set of vertices $a_1, b_2, a_3, b_4, \dots, a_k$. It induces a sub-tournament in G. It is a well-known classic fact that every tournament has a king; namely, a vertex that can reach all other vertices either in one step or two steps. Clearly a_1 cannot be the king since $b_2 \notin N^+(a_1) \cup N^{++}(a_1)$. Similarly, $b_2, a_3, ..., b_{k-1}$ cannot be kings. Thus, a_k must be a king. But if (a_k, b_1, b_k, a_1) were a losing cycle then $a_1 \notin N^+(a_k) \cup N^{++}(a_k)$, a contradiction. Thus, (a_k, a_1, b_k, b_1) is a losing cycle. If C is an even cycle consider the set $a_1, b_2, \dots, b_k$. Again,

b_k must be the king. But if (a_k, a_1, b_k, b_1) were a losing cycle then $a_1 \notin N^+(b_k) \cup N^{++}(b_k)$, a contradiction. Thus, (a_k, b_1, b_k, a_1) is a losing cycle. ■

Lemma 3.3 *If C is a cycle, let $K = \{a_1, b_1, \dots, a_k, b_k\}$. Then, for all $x, y \in K$ we have $N^+(x) \setminus K = N^+(y) \setminus K$ and $N^-(x) \setminus K = N^-(y) \setminus K$.*

Proof: We prove the lemma for even cycles. The proof for odd cycles is analogous. Indeed, assume C is an even cycle. We will prove that

$$N^+(a_1) \setminus K \subset N^+(b_2) \setminus K \subset N^+(a_3) \setminus K \subset \cdots \subset N^+(b_k) \setminus K \subset$$

$$\subset N^+(b_1) \setminus K \subset \cdots \subset N^+(a_k) \setminus K \subset N^+(a_1) \setminus K.$$

which implies that all inclusions are, in fact, equalities. By the cyclic structure of even cycles of Δ_G determined in Lemma 3.2 it suffices to prove, w.l.o.g., that $N^+(a_1) \setminus K \subset N^+(b_2) \setminus K$. Let $x \in N^+(a_1) \setminus K$. Thus, $(b_2, x) \in E$ since otherwise, $(x, b_2) \in E$ which implies that $b_2 \in N^{++}(a_1)$ contradicting the fact that (a_1, a_2, b_1, b_2) is a losing cycle. The claim on the incoming neighborhoods follows from the fact that each vertex outside of K is adjacent to each vertex of K. ■

We will now designate a vertex $f \in V$ that has the SNP in G. Let T^* be the tournament obtained from T_0 by contracting each set of vertices corresponding to even cycles of Δ_G into a single vertex. By Lemma 3.3 this contraction is well-defined. Namely, if C is such an even cycle and $K = \{a_1, b_1, \dots, a_k, b_k\}$ are its corresponding vertices then the newly contracted vertex has its incoming and outgoing neighbors well-defined. Assign weights to the vertices of T^* as follows. non-contracted vertices receive weight 1. A vertex contracted from an even cycle of length k in Δ_G receives weight $2k$ (the number of vertices contracted into it). Consider a weighted local median order of T^*, and let $v^* \in V(T^*)$ be the feed vertex. In case v^* is a non-contracted vertex we set $f = v^*$. If v^* is contracted from the even cycle $C = (a_1 b_1, \dots, a_k b_k)$ of Δ_G and $(a_i, b_i) \in T_0$ then we define f as follows. Consider the sub-tournament T_C of T_0 induced by $K = \{a_1, b_1, \dots, a_k, b_k\}$. Notice that T_C has no sink, and hence, by Theorem 2.2, T_C has two vertices with the (non-weighted) SNP (in T_C). Thus, at least one of these two vertices is not b_k. This other vertex will be f. Having designated f it remains to prove:

Lemma 3.4 *f has the SNP in G.*

Proof: We consider three cases. Consider first the case $f = v^*$, namely v^* is not a contracted vertex. From the fact that v^* has the weighted SNP in T^*, and from Lemma 3.3 we have that f has the (non-weighted) SNP in T_0. Namely, $d^+_{T_0}(f) \le d^{++}_{T_0}(f)$. There are two possibilities here. Either $d^+_{T_0}(f) = d^+(f)$ or $d^+_{T_0}(f) = d^+(f) + 1$. Consider first the case $d^+_{T_0}(f) = d^+(f)$. It suffices to prove that $d^{++}_{T_0}(f) = d^{++}(f)$. Assume that $(f, a_i) \in E$ and $a_i b_i$ belongs to some path or cycle

$C = (a_1b_1, \ldots, a_kb_k)$ of Δ_G that was oriented (a_i, b_i). We must prove that b_i is not a new second neighbor of f, namely, that $b_i \in N^+(f) \cup N^{++}(f)$. Consider first the case $i > 1$. If $f = b_{i-1}$ then, since $(b_{i-1}, b_i) \in E$ (recall the definition of losing cycles) we have $b_i \in N^+(f)$. Thus, we may assume that $f \neq b_{i-1}$. Furthermore, $f \neq a_{i-1}$ since otherwise we would have $d^+_{T_0}(f) = d^+(f) + 1$ which is not the current case. Consider, therefore, the arc connecting f and b_{i-1} in E. If its direction is (b_{i-1}, f) then $a_i \in N^{++}(b_{i-1})$ contradicting the fact that $(a_{i-1}, a_i, b_{i-1}, b_i)$ is a losing cycle. If its direction is (f, b_{i-1}) then, since $(b_{i-1}, b_i) \in E$ we have $b_i \in N^+(f) \cup N^{++}(f)$, as required. The case $i = 1$ also holds similarly when C is a cycle. In this case replace $i-1$ with k in the argument above and possibly change the roles of a_k and b_k in case C is an even cycle, as shown in Lemma 3.2. In case $i = 1$ and C is a path then the orientation (a_1, b_1) was decided upon *because* $b_1 \in N^+(v) \cup N^{++}(v)$ for each $v \in V \setminus \{a_1, b_1\}$ with $(v, a_1) \in E$, and, in particular, for $v = f$.

Consider next the case $d^+_{T_0}(f) = d^+(f) + 1$. Namely, there exists some path or odd cycle $C = (a_1b_1, \ldots, a_kb_k)$ of Δ_G that was oriented (a_i, b_i) and $f = a_i$. It suffices to prove that $d^{++}_{T_0}(f) \leq d^{++}(f) + 1$. Consider first the case where $i < k$. In fact, the proof in this case is identical to the proof in the previous paragraph, except that now b_{i+1} becomes a new second neighbor of $a_i = f$. We must also make sure that no outgoing neighbor of b_i becomes a new second neighbor of $a_i = f$ (due to the fact that b_i is a new first neighbor of $a_i = f$). Indeed, assume $t \in N^+(b_i)$. If $(t, a_{i+1}) \in E$ then, since $(a_i, a_{i+1}) \in E$ we have $a_{i+1} \in N^{++}(b_i)$ contradicting the fact that $(a_i, a_{i+1}, b_i, b_{i+1})$ is a losing cycle. Thus, we must have $(a_{i+1}, t) \in E$ and so t does not become a new second neighbor of a_i after adding (a_i, b_i). It follows that $d^{++}_{T_0}(f) \leq d^{++}(f) + 1$. The case $i = k$ also holds in case C is an odd cycle. In this case replace $i+1$ with 1 in the argument above. In case $i = k$ and C is a path there is no b_{i+1} to worry about, but, on the other hand, outgoing neighbors of b_k may become new second neighbors of a_k. Modify T_0 and T^* by replacing (a_k, b_k) with (b_k, a_k) and notice that the same weighted local median order of the original T^* is also a weighted local median order of the modified one. Now apply exactly the same argument as in the case $d^+_{T_0}(f) = d^+(f)$.

Finally, consider the case that v^* is a contracted vertex. Thus, there exists some even cycle $C = (a_1b_1, \ldots, a_kb_k)$ of Δ_G that was oriented (a_i, b_i) and that was contracted to v^*. Consider the sub-tournament T_C of T_0 induced by $K = \{a_1, b_1, \ldots, a_k, b_k\}$. Then, f was chosen to satisfy the SNP inside T_C, and, furthermore, $f \neq b_k$. Thus,

$$|N^+_{T_0}(f) \cap K| \leq |N^{++}_{T_0}(f) \cap K|. \tag{3}$$

As shown in the preceding paragraphs, we have

$$N^+_{T_0}(f) \setminus K = N^+(f) \setminus K\ , \quad N^{++}_{T_0}(f) \setminus K = N^{++}(f) \setminus K. \tag{4}$$

From the fact that v^* has the SNP in T^* and from Lemma 3.3 it follows that

$$|N^+_{T_0}(f) \setminus K| \leq |N^{++}_{T_0}(f) \setminus K|. \tag{5}$$

If $f = b_i$ for some $i = 1, \ldots, k-1$ we have

$$|N^+_{T_0}(f) \cap K| = |N^+(f) \cap K| \ , \quad |N^{++}_{T_0}(f) \cap K| = |N^{++}(f) \cap K| \tag{6}$$

(here we use the fact that $f \neq b_k$ since for b_k we have that b_1 is a new second neighbor of b_k since (a_k, b_k, b_1, a_1) is a losing cycle). If $f = a_i$ for some $i = 1, \ldots, k$ we have

$$|N^+_{T_0}(f) \cap K| = |N^+(f) \cap K| + 1 \ , \quad |N^{++}_{T_0}(f) \cap K| \leq |N^{++}(f) \cap K| + 1. \tag{7}$$

The right part of (7) follows from the fact that b_{i+1} becomes a new second neighbor of a_i for $i = 1, \ldots, k-1$. It now follows from (3)-(7) that

$$|N^+(f)| \leq |N^{++}(f)|.$$

■

3.2 Removing a star

Theorem 3.5 *Let $G = (V, E)$ be an orientation obtained from a tournament by deleting a set of arcs incident with the same vertex. Then the SNC holds for G.*

Proof: Let x denote the vertex incident with all the missing edges. We will prove that there exists a completion $T \in C_G$ so that for all $v \in V - x$, $N^{++}(v) = N^{++}_T(v)$. We will prove this by induction on the number of missing edges. If G is a tournament, the claim holds vacuously. Otherwise, let $\{x, y\}$ be a missing edge. Let $Q \subset V$ be the set of vertices with $q \in Q$ if and only if $(y, q) \in E$, $(q, x) \in E$ and $y \notin N^{++}(q)$. Let $R \subset V$ be the set of vertices with $r \in R$ if and only if $(r, y) \in E$, $(x, r) \notin E$ and $x \notin N^{++}(r)$. Clearly, $Q \cap R = \emptyset$. Furthermore, at least one of Q or R is an empty set. Indeed, assume $q \in Q$ and $r \in R$. Since q and r are adjacent there are two possibilities. If $(r, q) \in E$ then $x \in N^{++}(r)$, a contradiction. If $(q, r) \in E$ then $y \in N^{++}(q)$, a contradiction. Assume first that $Q = \emptyset$. In this case, add the arc (x, y). No vertex, except, possibly, x, received a new vertex in its second outgoing neighborhood. Assume next that $R = \emptyset$. In this case, add the arc (y, x). In this case, no vertex at all received a new vertex in its second outgoing neighborhood.

Having proved the existence of T, consider a feed vertex v of a local median order of T. If $v \neq x$ then, since v satisfies the SNP in T, and since $N^{++}(v) = N^{++}_T(v)$, then v also satisfies the SNP in G. If x is the feed vertex, then we can modify T so that all arcs that have been added, enter into x. Clearly, x satisfies the SNP in the modified T as well. Since all the outgoing neighbors of x in G are whole vertices (vertices adjacent to every other vertex in G) we have in the modified T that $N^+(v) = N^+_T(v)$ and also $N^{++}(v) = N^{++}_T(v)$. Since x satisfies the SNP in the modified T, it also satisfies the SNP in G. ■

3.3 Pseudocliques

Lemma 3.6 *Let G be an orientation. If v is the feed vertex of a local median order ψ_T of $T \in C_G$, then v is also a feed vertex of a local median order of some $T' \in C_G$ with $N^+(v) = N^+_{T'}(v)$.*

Proof: Let v be a vertex as stated. Choose T' to be the same as T, except that all arcs of $E(T) \setminus E(G)$ that emanate from v are inverted. Clearly, we now have $N^+(v) = N^+_{T'}(v)$ and the same local median order is also a local median order of T'. ■

A vertex of an orientation is called *whole* if it is adjacent to each other vertex.

Lemma 3.7 *Let G be an orientation of a pseudoclique. If v is a feed vertex of a local median order of some $T \in C_G$ and v is not whole, then v has the SNP in G.*

Proof: By Lemma 3.6, v is also a feed vertex in a local median order of $T' \in C_G$, and every non-neighbor of v in G is an incoming neighbor in T'. Since every neighbor of v in G is a whole vertex, we have that v is unaffected by T' (the first and second outgoing neighborhoods of v do not change). Thus, the SNP holds for v in G. ■

Theorem 3.8 *Let G be an orientation of a pseudoclique. Then, the SNC holds for G.*

Proof: Let Z be the set of whole vertices of G. If xy is a missing edge in G then either adding (x, y) does not create new second neighbors for the vertices of Z or else adding (y, x) does not create new second neighbors for the vertices of Z. Thus, orient xy so that whole vertices do not get new second neighbors (in case both (x, y) and (y, x) do not introduce new second neighborhoods, choose an arbitrary direction). We do this for each missing edge independently. In the obtained tournament T, every whole vertex of G is unaffected by T. let v be the feed vertex of a local median order of T. If v is whole, then the fact that v is unaffected by T implies that v has the SNP in G as well. If v is not whole then Lemma 3.7 implies that v has the SNP in G. ■

We note that a similar proof (using weighted local median orders) shows that the vertex-weighted version of the SNC holds for pseudocliques as well.

4 Approximations of second neighborhoods

In [3], it is shown that in every orientation there exists a vertex v with $d^{++}(v) \geq \gamma d^+(v)$, where $\gamma = 0.657298....$ In this section we prove a better approximation holds for orientations of K_{k+1}-free graphs. Notice that this class of orientations includes all k-partite orientations. We note the the SNC is not known to hold even for k-partite tournaments for all $k \geq 3$ (the case $k = 2$ is trivial).

Proof of Theorem 1.3: The proof is by induction on the number of vertices. The theorem is trivial for orientations with 1 or 2 vertices. Suppose that G is an orientation of a K_{k+1}-free graph on n vertices. Assume, to the contrary, that G does not contain a vertex v such that $d^{++}(v) \geq \gamma d^+(v)$. Let u be a vertex of G with minimum out-degree. For convenience, let $A = N^+(u)$, $B = N^{++}(u)$, $a = |A|$, and $b = |B|$. By our assumption, the following inequality holds,

$$b = d^{++}(u) < \gamma d^+(u) = \gamma a. \tag{8}$$

It follows that $d^+_A(x) + d^+_B(x) = d^+(x) \geq d^+(u) = a$ for every vertex $x \in A$. Notice that the underlying graph induced by A is K_k-free, and hence it contains at most $\frac{k-2}{2(k-1)}a^2$ edges. We therefore have that $\sum_{x\in A} d^+(x) \leq \frac{k-2}{2(k-1)}a^2$. It follows that:

$$e(A,B) = \sum_{x\in A} d^+_B(x) \geq \sum_{x\in A}(a - d^+_A(x)) \geq \frac{k}{2(k-1)}a^2. \tag{9}$$

Since $|A| = a < n$, by the induction hypothesis, there is a vertex $x \in N^+(u)$ such that $|N^{++}_A(x)| \geq \gamma|N^+_A(x)|$. Let $X = N^+_A(x)$, $Y = N^+(x) - A = N^+(x) \cap B$, and $d = |Y|$. Since $|A - X| \geq |N^{++}_A(x)| \geq \gamma|X|$, we have $(1+\gamma)|X| \leq a$. Thus,

$$|X| \leq \frac{1}{1+\gamma}a \leq \frac{9a}{10},$$

where the last inequality follows since $\gamma \geq 1/9$. Since $d^+(x) \geq d^+(u)$,

$$d = |Y| = |N^+(x)| - |X| \geq a - \frac{9a}{10} = \frac{a}{10}.$$

For every $y \in Y$, since $d^{++}(x) < \gamma d^+(x)$ and $d^{++}_A(x) \geq \gamma d^+_A(x)$, we have

$$d^+_{V-A-Y}(y) \leq d^{++}(x) - d^{++}_A(x) < \gamma d^+(x) - \gamma d^+_A(x) = \gamma|Y| = \gamma d.$$

Using $d^+(y) \geq d^+(u) = a$ and the fact that

$$\sum_{y\in Y} d^+_Y(y) \leq \frac{k-2}{2(k-1)}d^2$$

we obtain:

$$e(Y,A) = \sum_{y\in Y} d^+_A(y) \geq \sum_{y\in Y}(a - d^+_{V-A-Y}(y) - d^+_Y(y)) \geq (a-\gamma d)d - \sum_{y\in Y} d^+_Y(y)$$
$$\geq (a-\gamma d)d - \frac{k-2}{2(k-1)}d^2 > (a - \gamma d - \frac{k-2}{2(k-1)}d)d.$$

Combining (8), (9) and the last inequality we obtain that

$$\gamma a^2 > ab \geq e(A,B) + e(B,A) \geq e(A,B) + e(Y,A) \geq \frac{k}{2(k-1)}a^2 + (a - \gamma d - \frac{k-2}{2(k-1)}d)d$$

k	γ	k	γ
3	0.8324	7	0.7172
4	0.7754	8	0.7088
5	0.7465	9	0.7024
6	0.7290	10	0.6975

Table 1: The approximation constant for various small k

where $a/10 \leq d \leq \gamma a$. Let

$$f(z) = \frac{k}{2(k-1)}a^2 + (a - \gamma z - \frac{k-2}{2(k-1)}z)z = -(\gamma + \frac{k-2}{2(k-1)})z^2 + az + \frac{k}{2(k-1)}a^2.$$

Since $f(z)$ is a quadratic function with a negative leading coefficient, the following inequality holds.

$$f(z) > \min\{f(a/10), f(\gamma a)\} \text{ for all } z \in (a/10, \gamma a).$$

Thus, $\gamma a^2 > \min\{f(a/10), f(\gamma a)\}$. A simple calculation gives us that

$$f(a/10) = \frac{a^2(119k - 18 - \gamma(2k-2))}{200k - 200}.$$

Solving $\gamma a^2 > \frac{a^2(119k-18-\gamma(2k-2)}{200k-200}$, we obtain that $\gamma > \frac{119k-18}{202k-202}$. By assigning $t = \frac{119k-18}{202k-202}$ in $(2k-2)x^3 + (k-2)x^2 - k = q(x,k)$, and assigning $k = 3$, we obtain $q(\frac{339}{404}, 3) > 0$. Since $\gamma(k) > t$ and $q(x,3)$ is increasing in $[0,1]$, we reach a contradiction to the fact that γ is the real root of $q(x,3)$. For any k other than 3, observe that

$$\frac{\partial q(x,k)}{\partial k} = 2x^3 + x^2 - 1.$$

Since $2x^3 + x^2 - 1$ is positive within $(\gamma(\infty), 1]$, it follows that $q(t,k) > 0$ for all $k > 3$. Also, a simple calculation gives us that

$$f(\gamma a) = a^2(-\gamma^3 - \frac{k-2}{2(k-1)}\gamma^2 + \gamma + \frac{k}{2(k-1)}).$$

Simplifying the inequality

$$\gamma a^2 > a^2(-\gamma^3 - \frac{k-2}{2(k-1)}\gamma^2 + \gamma + \frac{k}{2(k-1)}),$$

we obtain that $\frac{2k-2}{k}\gamma^3 + \frac{k-2}{k}\gamma^2 - 1 > 0$, which contradicts that γ is the unique real root of the equation $\frac{2k-2}{k}x^3 + \frac{k-2}{k}x^2 - 1 = 0$. ■

Table 1 lists the values of γ for a small values of k.

5 A conditional result

Proof of Theorem 1.4: Let $\epsilon > 0$. Assume the theorem is false for ϵ. Thus, we may assume that there exists an orientation $G = (V, E)$ with n vertices (and we may assume, in the sequel, that n is sufficiently large) such that for each $v \in V$, $d^{++}(v) < d^{+}(v)(1-\epsilon) - \epsilon n$. In particular, the minimum out-degree of G is greater than ϵn. Let $k = k(\epsilon)$ be an integer to be chosen later. We may, and will, assume that n is a multiple of k. Consider a random partition of V into n/k subsets of size k each. For $v \in V$, let K_v be the induced sub-orientation of G on the vertex set to which v belongs in the random partition, and notice that K_v has k vertices. Let $x(v) = d^{+}_{K_v}(v)$ and let $y(v) = d^{++}_{K_v}(v)$. Clearly, $x(v)$ and $y(v)$ are random variables and their respective expectations are

$$E[x(v)] = \frac{k-1}{n-1} d^{+}(v) \quad , \quad E[y(v)] = \frac{k-1}{n-1} d^{++}(v).$$

Let, therefore, $z(v) = x(v) - y(v)$, then

$$E[z(v)] = \frac{k-1}{n-1}(d^{+}(v) - d^{++}(v)) > \frac{k-1}{n-1}\, \epsilon\, d^{+}(v) > \frac{k-1}{n-1}\epsilon^2 n > (k-1)\epsilon^2 \gg 0.$$

By taking $k = k(\epsilon)$ sufficiently large and $n = n(\epsilon)$ sufficiently large, we guarantee that the probability that $x(v)$ or $y(v)$ deviate from their means by more half the mean is smaller than $1/(2k)$ (in fact, the probability for such deviation is exponentially small in k). In particular, the probability that $z(v) \le 0$ is smaller than $1/k$. Hence, there *exists* a partition into n/k induced orientations in which more than $n(1-1/k)$ of the vertices v have $z(v) > 0$. Hence, some k-vertex orientation K has all of its vertices with $z(v) > 0$. But this simply means that in this sub-orientation, every vertex has $d^{+}_{K}(v) > d^{++}_{K}(v)$. By letting $F = F(\epsilon)$ be the family of all orientations with at least k vertices, the result follows. ■

We note that the dependency of k on ϵ is quite moderate. Already for $k = 50$ the obtained ϵ is quite small (much smaller than the approximation ratio in [3]). However, verifying the SNC for all orientations with, say, 50 vertices is currently beyond the computational frontier. It is known that every orientation with less than 16 vertices satisfies the SNC as such orientations are either tournaments or else have minimal out-degree at most 6, where the latter satisfy the SNC by [5].

References

[1] N. Dean and B.J. Latka, *Squaring the tournament: an open problem*, Congressus Numerantium 109 (1995), 73-80.

[2] D. Fisher, *Squaring a tournament: a proof of Dean's conjecture*, J. Graph Theory 23 (1996), 43-48.

[3] G. Chen, J. Shen and R. Yuster, *Second neighborhood via first neighborhood in digraphs*, Annals of Combinatorics 7 (2003), 15–20.

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

[5] Y. Kaneko and S.C. Locke, *The minimum degree approach for Paul Seymour's distance* 2 *conjecture*, Congressus Numerantium 148 (2001), 201–206.