On the length of directed paths in digraphs

Yangyang Cheng*,† Peter Keevash‡§

[a]Mathematical Institute, University of Oxford, UK

Abstract

Thomassé conjectured the following strengthening of the well-known Caccetta-Haggkvist Conjecture: any digraph with minimum out-degree δ and girth g contains a directed path of length $\delta(g-1)$. Bai and Manoussakis gave counterexamples to Thomassé's conjecture for every even $g \geq 4$. In this note, we first generalize their counterexamples to show that Thomassé's conjecture is false for every $g \geq 4$. We also obtain the positive result that any digraph with minimum out-degree δ and girth g contains a directed path of $2\delta(1-\frac{2}{g})$. For small g we obtain better bounds, e.g. for $g = 3$ we show that oriented graph with minimum out-degree δ contains a directed path of length 1.5δ. Furthermore, we show that each d-regular digraph with girth g contains a directed path of length $\Omega(dg/\log d)$. Our results give the first non-trivial bounds for these problems.

1 Introduction

The Caccetta-Haggkvist Conjecture [1] states that any digraph on n vertices with minimum out-degree δ contains a directed cycle of length at most $\lceil n/\delta \rceil$; it remains largely open (see the survey [2]). A stronger conjecture proposed by Thomassé (see [3],[2]) states that any digraph with minimum out-degree δ and girth g contains a directed path of length $\delta(g-1)$. Bai and Manoussakis gave counterexamples to Thomassé's conjecture for every even $g \geq 4$. The conjecture remains open for $g = 3$, which in itself was highlighted as an unsolved problem in the textbook [4].

Conjecture 1. *Any oriented graph with minimum out-degree δ contains a directed path of length 2δ.*

In this note, we first generalize the counterexamples to show that Thomassé's conjecture is false for every $g \geq 4$.

Proposition 2. *For every $g \geq 2$ and $\delta \geq 1$ there exists a digraph D with girth g and $\delta^+(D) \geq \delta$ such that any directed path has length at most $\frac{g\delta}{2}$ if g is even or $\frac{(g+1)\delta}{2}$ if g is odd.*

In the positive direction, when g is large we can find a directed path of length close to 2δ.

Theorem 3. *Every digraph D with girth g and $\delta^+(D) \geq \delta$ contains a directed path of length $2\delta(1-\frac{1}{g})$.*

* *E-mail address:* Yangyang.Cheng@maths.ox.ac.uk

† Supported by a PhD studentship of ERC Advanced Grant 883810.

‡ *E-mail address:* Peter.Keevash@maths.ox.ac.uk

§ Supported by ERC Advanced Grant 883810.

For the cases $g = 3$ or $g = 4$, we have the following better bounds.

Theorem 4. *Every oriented graph D with $\delta^+(D) \geq \delta$ contains a directed path of length 1.5δ. Every digraph D with $\delta^+(D) \geq \delta$ and girth $g \geq 4$ contains a directed path of length 1.6535δ.*

Finally, we consider the additional assumption of approximate regularity, under which a standard application of the Lovász Local Lemma gives much better bounds, We call a digraph (C, d)-*regular* if $d^+(v) \geq d$ and $d^-(v) \leq Cd$ for each vertex v.

Theorem 5. *For every $C > 0$ there exists $c > 0$ such that if D is a (C, d)-regular digraph with girth g then D contains a directed path of length at least $cdg/\log d$.*

1.1 Notation

We adopt standard notation as in [3]. A *digraph* D is defined by a vertex set $V(D)$ and arc set $A(D)$, which is a set of ordered pairs in $V(D)$. An *oriented graph* is a digraph where we do not allow 2-cycles $\{(x, y), (y, x)\}$, i.e. it is obtained from a simple graph by assigning directions to the edges. For each vertex $v \in D$ and any vertex set $S \subseteq V(D)$, let $N^+(v, S)$ be the set of out-neighbours of v in S and let $d^+(v, S) = |N^+(v, S)|$. If $S = V(D)$, then we simply denote $d^+(v, S)$ by $d^+(v)$. If H is an induced subgraph of D, then we define $d^+(v, H) = d^+(v, V(H))$ for short. We let $\delta^+(D) = \min_v d^+(v)$ be the minimum out-degree of D. Indegree notation is similar, replacing $+$ by $-$.

For every vertex set $X \subseteq V(D)$, let $N^+(X)$ be the set of vertices that are not in X but are out-neighbours of some vertex in X. For every two vertex sets A, B of $V(D)$, let $E(A, B)$ be the set of arcs in $A(D)$ with tail in A and head in B. A digraph D is *strongly-connected* if for every ordered pair of vertices $u, v \in V(D)$ there exists a directed path from u to v.

The *girth* $g(D)$ of D is the minimum length of a directed cycle in D (if D is acyclic we define $g(D) = \infty$). We write $\ell(D)$ for the maximum length of a directed path in D.

2 Construction

We start by constructing counterexamples to Thomassé's conjecture for every $g \geq 4$, as stated in Proposition 2. Suppose that D is a digraph with $d^+(v) = \delta$ for each vertex $v \in V(D)$. For each $k \geq 1$, we define the *k-lift* operation on some fixed vertex v as follows: we delete all arcs with tail v, add $k - 1$ disjoint sets of δ new vertices $U_{v,1}, ..., U_{v,k-1}$ to D, write $U_{v,0} := \{v\}$, $U_{v,k} := N^+(v)$ and add arcs so that $U_{v,i-1}$ is completely directed to $U_{v,i}$ for $1 \leq i \leq k$. (For example, a 1-lift does not change the digraph.) We note that any lift preserves the property that all out-degrees are δ.

Write $\overrightarrow{K}_{\delta+1}$ for the complete directed graph on $\delta + 1$ vertices. Our construction is $D_{a,b} := \overrightarrow{K}^{\uparrow}_{\delta+1}(a, b, \ldots, b)$ for some integer $1 \leq a \leq b$, meaning that starting from $\overrightarrow{K}_{\delta+1}$, we a-lift some vertex v_1 and b-lift all the other vertices.

Claim 6. *The girth of $D_{a,b}$ is $a + b$ and the longest path has length δb.*

Proof. Let C be any directed cycle in $D_{a,b}$. By construction, we can decompose $E(C)$ into directed paths of the form $v_i u_1 \cdots u_t v_j$ such that $u_j \in U_{v_i,j}$ for $1 \leq j \leq t$ where $t = a$ if $i = 1$ and $t = b$ if $i \geq 2$. A directed cycle contains at least two such pieces, so its length is at least $a + b$ since $a \leq b$. It is also easy to see $D_{a,b}$ does contain a directed cycle of length $a + b$.

Now suppose P is a directed path in $D_{a,b}$ of maximum length. Similarly, we can decompose $E(P)$ into directed paths of the form $v_i u_1 \cdots u_t v_j$ such that $u_j \in U_{v_i,j}$ for $1 \leq j \leq t$ where $t = a$ if $i = 1$

and $t = b$ if $i \geq 2$ (the first and last pieces could be shorter). Each piece has length at most b, so P has length at most δb. □

Proposition 2 follows from Claim 6 by taking $a = b = \frac{g}{2}$ for g even or $a = \frac{g-1}{2}$ and $b = \frac{g+1}{2}$ for g odd.

3 The key lemma

Here we show that Theorems 3 and 4 follow directly from known results on the Caccetta-Häggkvist conjecture and the following key lemma.

Lemma 7. *If D is an oriented graph with $\delta^+(D) \geq \delta$ then D either contains a directed path of length 2δ or an induced subgraph S such that $|V(S)| \leq \delta$ and $\delta^+(S) \geq 2\delta - \ell(D)$.*

We use the following bounds on Caccetta-Häggkvist in general by Chvátal and Szemerédi [5] and in the case of directed triangles by Hladký, Král, and Norin [6].

Theorem 8. *Every digraph D with order n and $\delta^+(D) \geq \delta$ contains a directed cycle of length at most $\lceil \frac{2n}{\delta+1} \rceil$.*

Theorem 9. *Every oriented graph with order n and minimum out-degree $0.3465n$ contains a directed triangle.*

Now we deduce Theorems 3 and 4, assuming the key lemma.

Proof of Theorem 3. Suppose that D is an oriented graph with $\delta^+(D) \geq \delta$ and girth g. By Lemma 7, D contains an induced subgraph S with $|S| \leq \delta$ and $\delta^+(S) \geq 2\delta - \ell(D)$. According to Theorem 8, S contains a directed cycle of length at most $\frac{2\delta}{2\delta - \ell(D)+1}$. Therefore, $g \leq \frac{2\delta}{2\delta - \ell(D)+1}$, so $\ell(D) \geq 2\delta(1 - \frac{1}{g}) + 1$. □

Proof of Theorem 4. First, suppose that D is an oriented graph with $\delta^+(D) \geq \delta$. By Lemma 7, either D contains a directed path of length 2δ or D contains an induced subgraph S such that $|S| \leq \delta$ and $\delta^+(S) \geq 2\delta - \ell(D)$. Since D is oriented, for some vertex $b \in S$, we have $d^+(b, S) \leq \frac{|S|-1}{2}$, which means that $\delta^+(S) \leq \frac{|S|-1}{2} \leq \frac{\delta-1}{2}$ and so $\ell(D) \geq 2\delta + 1 - \delta^+(S) \geq \frac{3}{2}\delta$. Similarly, if D has girth at least 4 then substituting the bound $\delta^+(S) < 0.3465\delta$ from Theorem 9 we obtain $\ell(D) > 1.6535\delta$. □

In fact, by Lemma 7, any improved bound towards the Caccetta-Häggkvist conjecture can be used to get a better bound for $\ell(D)$ when $\delta^+(D) \geq \delta$ and girth g. The Caccetta-Häggkvist conjecture itself would imply $\ell(D) \geq (2 - \frac{1}{g})\delta$.

4 Proof of the key lemma

Suppose that D is an oriented graph with $\delta^+(D) \geq \delta$ and no directed path of length 2δ. We can assume that D is strongly-connected, as there is a strong component of D with minimum out-degree at least δ. By deleting arcs, we can also assume that all out-degrees are exactly δ. Note that $|V(D)| \geq 2\delta + 1$, since D is oriented and $\delta^+(D) \geq \delta$.

Claim 10. *D does not contain two disjoint directed cycles of length at least $\delta + 1$.*

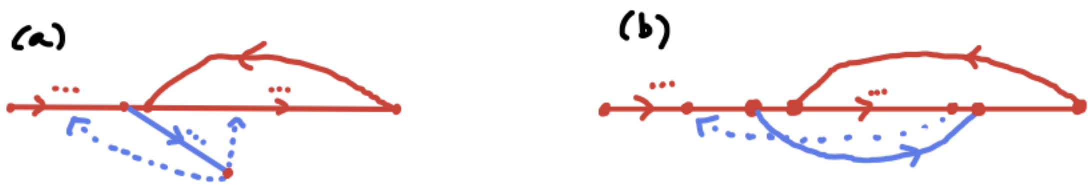

Figure 1: Illustrations for the proofs of Claims 11 and 12.

Proof. Suppose on the contrary that C_1 and C_2 are two such cycles. By strong connectivity, there exists a path P from $u_1 \in C_1$ to $u_2 \in C_2$ with $V(P)$ internally disjoint from $V(C_1) \cup V(C_2)$. Writing u_1u_1' for the out-arc of u_1 in C_1 and $u_2'u_2$ for the in-arc of u_2 in C_2, the path $\{C_1 - u_1u_1'\} + P + \{C_2 - u_2'u_2\}$ has length at least $2\delta + 1$, a contradiction. □

Now let $P = v_0v_1 \cdots v_{\ell(D)}$ be a directed path of maximum length, where $\ell(D) < 2\delta$. By maximality of P, the out-neighbours $N^+(v_{\ell(D)})$ of $v_{\ell(D)}$ must lie on P. Let $v_a \in N^+(v_{\ell(D)})$ such that the index a is minimum among all the out-neighbours of $v_{\ell(D)}$. Thus $C = v_av_{a+1} \cdots v_{\ell(D)}v_a$ is a directed cycle; we call $|C|$ the *cycle bound* of P. For future reference, we record the consequence

$$\ell(D) \geq g(D) \text{ for any digraph } D. \tag{1}$$

Choose P such that the cycle bound of P is also maximum subject to that P is a directed path of length $\ell(D)$. Clearly $a \neq 0$, otherwise using $|V(D)| \geq 2\delta + 1$ and strong connectivity, we can easily add one more vertex to C and get a longer path, contradiction.

Claim 11. *Every vertex in $N^+(v_{a-1})$ must be on P.*

Proof. Suppose on the contrary that there exists an out-neighbour w_1 of v_{a-1} such that $w_1 \in V(D) \setminus V(P)$. Let D_1 be the induced graph of D on $V(D) \setminus V(P)$. We extend the vertex w_1 to a maximal directed path $P_1 = w_1w_2 \cdots w_m$ in D_1. Since P_1 is maximal in D_1, all the out-neighbours of w_m must be on $V(P) \cup V(P_1)$, see Figure 1(a).

We cannot have $w \in N^+(u_m)$ such that $w \in V(C)$. Indeed, writing w^- for the in-neighbour of w in C, the directed path $P' = v_0 \ldots v_{a-1}P_1w + (C - w^-w)$ would be longer than P, a contradiction. Thus we conclude that $N^+(w_m) \subseteq V(P_1) \cup \{v_0, ..., v_{a-1}\}$. Choose a vertex $z \in N^+(w_m)$ that has the largest distance to u_m on the path $P_2 = v_0 \ldots v_{a-1}u_1 \ldots u_m$. Then $P_2 \cup w_mz$ contains a cycle C_1 of length at least $\delta + 2$. Now C_1 and C are two disjoint directed cycles of length at least $\delta + 2$, which contradicts Claim 10. □

Let $A = N^+(v_{a-1}) \cap \{v_0, \ldots, v_{a-1}\}$ and $B = N^+(v_{a-1}) \cap V(C)$. Also, let $B^- = \{u : u \in V(C), uv \in E(C) \text{ for some } v \in B\}$.

Claim 12. *$N^+(B^-) \subseteq V(C)$.*

Proof. Suppose not, then there exists a vertex $w \in V(D) \setminus V(C)$ such that $bw \in A(D)$ for some $b \in B^-$. By definition of B, there exists some vertex $b^+ \in B$ such that $v_{a-1}b^+ \in A(D)$ and $bb^+ \in A(C)$. We cannot have $w \in V(D) \setminus V(P)$, as then the path $v_0v_1 \ldots v_{a-1}b^+ + (C - bb^+) + bw$ has length $\ell(D) + 1$, a contradiction.

It remains to show that we cannot have $w \in V(P) \setminus V(C)$. Suppose that we do, with $w = v_i$ for some $0 \leq i \leq a-1$. Then the cycle $v_i v_{i+1} \dots v_{a-1} b^+ + (C - bb^+) + bv_i$ is longer than C. However, $P_1 = v_0 \dots v_{a-1} b^+ + (C - bb^+)$ has length $\ell(D)$ and cycle bound larger than P, which contradicts our choice of P, see Figure 1(b). $\square$

Now let S be the induced digraph of D on B^-. Fix $x \in B^-$ with $N_S^+(x) = \delta^+(S)$. Then $N^+(x) \subseteq V(C)$ by Claim 12. As $|N^+(x)| = \delta$ we deduce $|C| \geq |V(S)| - \delta^+(S) + \delta$.

Note that $|P| \geq |A| + 1 + |C| \geq |A| + 1 + |B| - \delta^+(S) + \delta$, as $|V(S)| = |B^-| = |B|$ and $A \subseteq \{v_0, \dots, v_{a-1}\}$. But $|A| + |B| \geq |N^+(v_{a-1})| \geq \delta$, so $\ell(D) = |P| \geq 2\delta + 1 - \delta^+(S)$ and $\delta^+(S) \geq 2\delta + 1 - \ell(D)$.

This completes the proof of Lemma 7.

5 Long directed paths in almost-regular digraphs

In this section, we prove Theorem 5. We start by stating some standard probabilistic tools (see [7]). We use the following version of Chernoff's inequality.

Lemma 13. *Let $X_1, \dots, X_n$ be independent Bernoulli random variables with $\mathbb{P}[X_i = 1] = p_i$ and $\mathbb{P}[X_i = 0] = 1 - p_i$ for all $i \in [n]$. Let $X = \sum_{i=1}^n X_i$ and $E[X] = \mu$. Then for every $0 < a < 1$, we have*

$$\mathbb{P}[|X - \mu| \geq a\mu] \leq 2e^{-a^2\mu/3}.$$

We will also use the following version of Lovász Local Lemma.

Lemma 14. *Let $A_1, \dots, A_n$ be a collection of events in some probability space. Suppose that each $\mathbb{P}[A_i] \leq p$ and each A_i is mutually independent of a set of all the other events A_j but at most d, where $ep(d+1) < 1$. Then $\mathbb{P}[\cap_{i=1}^n \overline{A_i}] > 0$.*

Next we deduce the following useful partitioning lemma.

Lemma 15. *For every $C > 0$ there exists $c > 0$ such that for any positive integer d with $t := \lfloor cd/\log d \rfloor \geq 1$, for any (C,d)-regular digraph D there exists a partition of $V(D)$ into $V_1 \cup \dots \cup V_t$ such that $||V_i| - |V_j|| \leq 1$ and $d^+(v, V_j) \geq \frac{\log d}{2c}$ for each $i, j \in [n]$ and $v \in V_i$.*

Proof. We start with an arbitrary partition $U_1 \cup \dots \cup U_s$ of $V(D)$ where $|U_1| = \dots = |U_{s-1}| = t$ and $1 \leq |U_s| \leq t$, so that $n/t \leq s < n/t + 1$. We add $t - |U_s|$ isolated 'fake' vertices into U_s to make it a set of size t. We consider independent uniformly random permutations $\sigma_i = (\sigma_{i,1}, \dots, \sigma_{i,t})$ of each U_i. Now let $V_j = \{\sigma_{1,j}, \dots, \sigma_{s,j}\}$ for each $1 \leq j \leq t$. We will show that $V_1 \cup \dots \cup V_t$ (with fake vertices deleted) gives the required partition with positive probability.

We consider the random variables $X(v,j) := d^+(v, V_j)$ for each $v \in V$ and $j \in [t]$. Note that each is a sum of independent Bernoulli random variables with $\mathbb{E}[X(v,j)] = d^+(v)/t$. We let $E_{v,j}$ be the event that $\left| X(v,j) - \frac{d^+(v)}{t} \right| \geq \frac{d^+(v)}{2t}$. Then $\mathbb{P}[E_{v,j}] \leq 2e^{-\frac{d^+(v)}{12t}} \leq 2e^{-\frac{d}{12t}}$ by Chernoff's inequality.

Now $E_{v,j}$ is determined by those σ_i with $U_i \cap N^+(v) \neq \emptyset$, so is mutually independent of all but at most $C(dt)^2$ other events $E_{v',j'}$, using $\Delta^-(D) \leq Cd$. For c sufficiently small, for example $c \leq \frac{1}{100 \log C}$, we get $2e^{-\frac{d}{12t}+1}(C(dt)^2 + 1) < 1$. By Lemma 14 we conclude that with positive probability no $E_{v,j}$ occurs, and so $V_1 \cup \dots \cup V_t$ (with fake vertices deleted) gives the required partition. $\square$

Proof of Theorem 5. Suppose that D is a (C, d)-regular digraph with girth g. We will show $\ell(D) \geq \frac{cdg}{2\log d}$ with c as in Lemma 15. As $\ell(D) \geq g(D)$ by (1), we can assume $cd/\log d \geq 1$, so $t = \lfloor cd/\log d \rfloor \geq 1$. By Lemma 15 we can partition $V(D)$ into $V_1 \cup \cdots \cup V_t$ such that $||V_i| - |V_j|| \leq 1$ and each $d^+(v, V_j) \geq \frac{\log d}{2c}$. We note that $\frac{\log d}{2c} > 1$ for $c < 0.1$, say.

Let P_1 be a maximal directed path in $D[V_1]$ starting from any vertex x_1, ending at some y_1. Then $|P_1| \geq g$ by (1). By choice of partition, y_1 has an out-neighbour x_2 inside $D[V_2]$. Similarly, we can find a maximal directed path of length at least g inside $D[V_2]$ starting from x_2. We repeat the process until we find t directed paths $P_1, \ldots, P_t$ of length at least g, that can be connected into a directed path of length at least $tg \geq \frac{cdg}{2\log d}$. This completes the proof. □

6 Concluding remarks

We propose the following weaker version of Thomassé's conjecture.

Conjecture 16. *There is some $c > 0$ such that $\ell(D) \geq cg(D)\delta^+(D)$ for any digraph D.*

By Proposition 2, the best possible c in this conjecture satisfies $c \leq 1/2$. We do not even know whether it holds for regular digraphs, or whether $\ell(D)/\delta^+(G) \to \infty$ as $g \to \infty$.

Acknowledgments. We are grateful to António Girão for helpful discussions.

References

[1] Louis Caccetta and Roland Haggkvist. *On minimal digraphs with given girth.* Department of Combinatorics and Optimization, University of Waterloo, 1978.

[2] B.D. Sullivan. A summary of problems and results related to the Caccetta-Häggkvist conjecture. *arXiv:math/0605646*, 2006.

[3] J. Bang-Jensen and G. Gutin. *Digraphs: Theory, Algorithms and Applications.* Springer-Verlag London, 2008.

[4] J.A. Bondy and U.S.R. Murty. *Graph Theory.* Springer, Berlin, 2008.

[5] V. Chvátal and E. Szemerédi. Short cycles in directed graphs. *Journal of Combinatorial Theory, Series B*, 35(3):323–327, 1983.

[6] J. Hladký, D. Král, and S. Norin. Counting flags in triangle-free digraphs. *Combinatorica*, 3(1):49–76, 2017.

[7] N. Alon and J.H. Spencer. *The Probabilistic Method.* John Wiley&Sons, Inc.,, 1992.