

Finding 2-edge and 2-vertex strongly connected components in quadratic time

Monika Henzinger Sebastian Krinninger Veronika Loitzenbauer

University of Vienna

December 19, 2014

Abstract

We present faster algorithms for computing the 2-edge and 2-vertex strongly connected components of a directed graph, which are straightforward generalizations of strongly connected components. While in *undirected* graphs the 2-edge and 2-vertex connected components can be found in linear time, in *directed* graphs only rather simple $O(mn)$ -time algorithms were known. We use a hierarchical sparsification technique to obtain algorithms that run in time $O(n^2)$. For 2-edge strongly connected components our algorithm gives the first running time improvement in 20 years. Additionally we present an $O(m^2/\log n)$ -time algorithm for 2-edge strongly connected components, and thus improve over the $O(mn)$ running time also when $m = O(n)$. Our approach extends to k -edge and k -vertex strongly connected components for any constant k with a running time of $O(n^2 \log^2 n)$ for edges and $O(n^3)$ for vertices.

1 Introduction

Problem Description In a directed graph G two vertices u and v are *2-edge strongly connected* if from u to v and from v to u , respectively, there are two paths that have no common edge. A *2-edge strongly connected component* (2eSCC) of G is a maximal subgraph of G such that in the subgraph every pair of distinct vertices is 2-edge strongly connected. The *2-vertex strongly connected components* (2vSCCs) are defined in a similar but slightly more complicated way (see Section 2 for details). 2-edge- and 2-vertex-connectivity are central properties of graphs and have many applications [1, 23], for example in the construction of reliable communication networks [2] and in the analysis of the structure of networks [25].

Our Results In this work we present an algorithm that computes the 2eSCCs and the 2vSCCs of a directed graph in $O(n^2)$ time. For 2eSCCs we additionally provide an algorithm that runs in time $O(m^2/\log n)$, which is faster than $O(n^2)$ if $m = O(n)$. Thus we significantly improve upon the previous $O(mn)$ -time algorithms for both 2eSCCs [24, 14] and 2vSCCs [21]. For 2eSCCs the previous upper bound stood for 20 years. Our approach immediately generalizes to computing the *k-edge strongly connected components* (keSCCs) and the *k-vertex*

strongly connected components (kvSCCs). We give algorithms that, for any constant k , compute (1) the keSCCs in time $O(n^2 \log^2 n)$ (improving upon the previous upper bound of $O(mn)$ [24]) and (2) the keSCCs in time $O(n^3)$ (improving upon the previous upper bound of $O(mn^2)$ [22]).

Related Work The 2-edge and 2-vertex connected components of an *undirected* graph can be determined in linear time [26, 19]. In *directed* graphs several related problems can be solved in linear time: Testing whether a graph is 2-edge or 2-vertex strongly connected [27, 11, 13], finding all *bridges* and *articulation points* [20], and determining the *2-edge* and *2-vertex strongly connected blocks* [14, 15]. An edge is a bridge and a vertex is an articulation point, respectively, if its removal from the graph increases the number of strongly connected components of the graph. In a 2-edge strongly connected block every pair of distinct vertices is 2-edge strongly connected. However, as opposed to a 2eSCC, the paths to connect the vertices in a block might use vertices that are not in the same block. Each 2eSCC is completely contained in one 2-edge connected block, i.e., the 2eSCCs refine the 2-edge connected blocks.

Georgiadis et al. [14] and Jaberri [21] described simple algorithms to compute the 2eSCCs and 2vSCCs in $O(mn)$ -time, respectively, and posed as an open problem whether this can be improved to linear time as well. An $O(mn)$ running time for computing the 2eSCCs was already achieved by Nagamochi and Watanabe in 1993 [24], which in fact solved the more general problem of computing the keSCCs (also for multigraphs). To the best of our knowledge, the fastest known algorithm for computing the kvSCCs is by Makino [22] and has a running time of $O(mn^2)$. In *undirected* graphs there are linear-time algorithms for computing both the 3-vertex connected components [19] and the 3-edge connected components [12]. The k -edge connected components of an undirected graph can be computed in time $O(m + k^2 n^2)$ [24] (also in multigraphs).

Techniques As the $O(n^2)$ -algorithms for 2eSCCs and 2vSCCs are very similar, we describe here the main ideas for the slightly simpler case of edge-connectivity. At the heart of our algorithm we use a *hierarchical graph sparsification* that was used by Henzinger et al. [17] for undirected graphs and extended to directed graphs and game graphs in [4, 5]. Roughly speaking, this sparsification technique allows us to replace the ‘ m ’ in the $O(mn)$ running time by an ‘ n ’, yielding $O(n^2)$. Our main technical contribution is to find structural properties of edge connectivity that allow us to apply this technique. Note that while various ways of sparsification are used in algorithms for *undirected* graphs, such approaches are rarely found for *directed* graphs. We therefore strongly believe that understanding how to apply sparsification for directed graphs is interesting in its own right.

The straightforward approach for computing 2eSCCs, which gives an $O(mn)$ running time, is the following: Assuming that the graph is strongly connected, first find the bridges of the graph, i.e., the edges whose removal increases the number of strongly connected components. Then remove the bridges from the graph and compute the strongly connected components. Repeat the algorithm for each strongly connected component until the component does not contain a bridge anymore. The components remaining in the end are the 2eSCCs. We now explain in which way our algorithm deviates from this scheme.

Note that a 2eSCC is (a) strongly connected and (b) for any possible split

of the vertex set of a 2eSCC into two sets, there are at least two edges in each direction in the 2eSCC that connect the two parts. Thus if we detect in a graph G a set of vertices S that (a) cannot be reached by the vertices in $V \setminus S$ or (b) that can be reached from $V \setminus S$ only through one edge, then we know that each 2eSCC of G contains either only vertices in S or only vertices in $V \setminus S$. Our algorithms will repeatedly identify such sets S and recurse on the subgraphs induced by S and $V \setminus S$, respectively.

The difference to the straightforward approach is thus the following: Instead of repeatedly removing edges that cannot be in a 2eSCC (in particular bridges), we focus on partitioning the set of vertices. To see why this is useful, note that the incoming edges of the vertices S in G consist of the incoming vertices from other vertices in S and possibly one edge from $V \setminus S$ to S . Thus the number of incoming edges of each vertex in S is bounded by the number of vertices in S . We will use this insight as follows: When searching for such a set S , we start the search in a subgraph of G that includes all vertices but only the first incoming edge of each vertex. If no such set S is found, we repeatedly double the number of incoming edges per vertex in the subgraph until the search is successful. In this way the search will take time $O(n)$ per vertex in S . This will allow us to bound the total runtime by $O(n^2)$.

To correctly identify such sets S by a search in a strict subgraph of G , the algorithm will find *edge-dominators* in slightly modified *flow graphs*. A flow graph is a directed graph with a designated root where all vertices are reachable from the root. An edge is an edge-dominator in a flow graph if some vertex can be reached from the root only through this edge. Our algorithms will use the linear-time algorithms for dominators [27, 11, 13] and strongly connected components [26] as subroutines.

In the $O(m^2/\log n)$ -algorithm for 2eSCCs we search small sets S in subgraphs that are obtained by local breadth-first searches from vertices that lost edges in the previous iteration of the algorithm. Such local breadth-first searches were first used for Büchi games by Chatterjee et al. [6].

Outline The remainder of this paper is organized as follows. In Section 2 the main definitions and the notation is introduced. We also explain the straightforward algorithms for 2eSCCs and 2vSCCs. In Section 3 we show when and how we can identify a set S in a strict subgraph of G . In Section 4 we present the $O(n^2)$ -algorithms for 2eSCCs and 2vSCCs, in Section 5 the $O(m^2/\log n)$ -algorithm for 2eSCCs. In Section 6 we outline how the results from the Sections 3 and 4 extend to keSCCs and kvSCCs.

2 Preliminaries

Let $G = (V, E)$ be a simple directed graph with $m = |E|$ edges and $n = |V|$ vertices. The reverse graph $Rev(G)$ of G is equal to (V, E^R) where E^R is the set containing for each edge $(u, v) \in E$ its reverse (v, u) . For any set $S \subseteq V$ we denote by $G[S]$ the subgraph of G induced by the vertices in S , i.e., the graph $(S, E \cap (S \times S))$. The incoming edges of a vertex v in G will be denoted by $\text{In}_G(v)$, the number of incoming edges by $\text{Indeg}_G(v)$; analogously we use $\text{Out}_G(v)$ and $\text{Outdeg}_G(v)$ for outgoing edges. We denote by $G \setminus V'$ the graph $G[V \setminus V']$ and by $G \setminus E'$ the graph $(V, E \setminus E')$ for any set of vertices $V' \subseteq V$ and

any set of edges $E' \subseteq E$.

Strong connectivity A subgraph $G[S]$ induced by some set of vertices S is *strongly connected* if for every pair of distinct vertices u and v in S there exists a path from u to v and a path from v to u in $G[S]$. A subgraph induced by a single vertex is considered strongly connected. The *strongly connected components* (SCCs) of G are its maximal strongly connected subgraphs and form a partition of V . A strongly connected subgraph with no out-going edges is a bottom SCC (bSCC), a strongly connected subgraph with no in-coming edges a top SCC (tSCC). bSCCs and tSCCs are by definition maximal. Every graph G contains at least one bSCC and one tSCC. If G is not strongly connected, then there exist a bSCC and a tSCC that are disjoint and thus one of them contains at most half of the vertices of G . Note that a bSCC in G is a tSCC in $Rev(G)$ and vice versa. We will further use that when a set of vertices S cannot be reached by the vertices in $V \setminus S$ in G , then $G[S]$ contains a tSCC of G . The SCCs of a graph can be computed in $O(m)$ time [26].

Strong 2-edge and 2-vertex connectivity An edge $e \in E$ is a *bridge* if the removal of e from G increases the number of SCCs in G . Similarly, a vertex $v \in V$ is an *articulation point* if the removal of v from G increases the number of SCCs in G . All bridges and all articulation points of a graph can be found in time $O(m)$ [20]. Two (simple) paths are *edge-disjoint* if they do not share an edge; they are *vertex-disjoint* if they do not share a vertex except possibly their endpoints. Two distinct vertices u and v are *2-edge-connected* (*2-vertex-connected*) in G if they remain strongly connected after the removal of any edge (any vertex except u and v) from G . There exist two edge-disjoint paths from u to v and two edge-disjoint paths from v to u if and only if u and v are 2-edge-connected. If there is no edge between u and v , then it also holds that u and v are 2-vertex-connected if and only if there exists two vertex-disjoint paths from u to v and two vertex-disjoint paths from v to u . A subgraph $G[S]$ induced by some set of vertices S is *2-edge-connected* (*2-vertex-connected*) if every pair of distinct vertices u and v in S is 2-edge-connected (2-vertex-connected) in $G[S]$. The *2-edge strongly connected components* (2eSCCs) of a graph are its maximal 2-edge-connected subgraphs. Equivalently, the 2eSCCs are the maximal strongly connected subgraphs such that none of the subgraphs contains a bridge. Similarly, the *2-vertex strongly connected components*¹ (2vSCCs) of a graph are its maximal 2-vertex-connected subgraphs. Equivalently, the 2vSCCs are the maximal strongly connected subgraphs such that none of the subgraphs contains an articulation point. Note that the definition of 2vSCCs allows for *degenerate* 2vSCCs that consist of a strongly connected subgraph $G[\{u, v\}]$ with two vertices. Observe that the 2eSCCs form a partition of the vertices of the graph while this does not hold for 2vSCCs.

A simple algorithm to compute the 2eSCCs of a graph G [14] works as follows.

1. For each SCC C of G :
 - (a) Find all bridges X in C .

¹Our definitions follow [14], while [21, 15] use slightly different definitions for 2-vertex-connectivity. The 2vSCCs of [21, 15] can be determined in $O(n)$ time from the 2vSCCs defined here.

- (b) If none found, output C .
- (c) Otherwise recurse on $C \setminus X$.

It can easily be seen that the runtime is $O(mn)$: In each iteration SCCs and bridges are found in time $O(m)$. Whenever the number of SCCs does not increase from one iteration to the next, the algorithm terminates. Thus the number of iterations is at most n .

Jaberi [21] gave a similar $O(mn)$ -time algorithm to compute the 2vSCCs of a graph G .

1. For each SCC C of G :
 - (a) Find an articulation point v in C .
 - (b) If none found, output C .
 - (c) Otherwise compute the SCCs of $C \setminus \{v\}$
 - i. For each SCC C' recurse on $C' \cup \{v\}$

To the best of our knowledge, no faster algorithms are known for neither 2eSCCs nor 2vSCCs.

Flow graphs We define the *flow graph* $G(r)$ to be the graph G with a vertex $r \in V$ designated as the root and with all vertices not reachable from r removed. An edge $e \in E$ is an *edge-dominator* in $G(r)$ if there exists a vertex $u \in V \setminus \{r\}$ such that u is reachable from r and every path from r to u contains e . We say that e *dominates* u in $G(r)$. A *vertex-dominator* in $G(r)$ is a vertex $v \in V \setminus \{r\}$ for which there exists a vertex $u \in V \setminus \{r, v\}$ such that u is reachable from r and every path from r to u contains v . We say that v *dominates* u in $G(r)$. The edge- and vertex-dominators of a flow graph can be computed in linear time [16, 27, 11, 3]. Italiano et al. [20] showed the following relations between edge-dominators and vertex-dominators in flow graphs and bridges and articulation points in strongly connected graphs.

Lemma 1 (Italiano et al. [20]). *Let $G = (V, E)$ be a strongly connected graph.*

1. *If $e \in E$ is an edge-dominator in the flow graph $G(r)$ for some $r \in V$, then e is a bridge in G . If $e \in E$ is a bridge in G , then there exists a vertex $r \in V$ such that e is an edge-dominator in $G(r)$.*
2. *If $v \in V$ is a vertex-dominator in the flow graph $G(r)$ for some $r \in V$, then v is an articulation point in G . If $v \in V$ is an articulation point in G , then there exists $r \neq v$ in V such that v is a vertex-dominator in $G(r)$.*

Notation for unified analysis In many places the analysis for 2eSCCs and 2vSCCs will coincide. Thus we introduce the following notation to be able to unify the analysis for both cases. Let an *element* f of a graph G denote either an edge or a vertex of G . A *dominator* f can be either an edge- or a vertex-dominator. A *separator* f can be either a bridge or an articulation point. We say two paths are *disjoint* if we mean edge- or vertex-disjoint, respectively. Similarly, we use *2SCC* to denote 2eSCC or 2vSCC and *2-connected* to denote 2-edge- or 2-vertex-connected. Note that the text is meant to be read consistently for either edge- or vertex-connectivity, e.g., when a statement about 2-connectivity

is interpreted as a statement about 2-vertex-connectivity, then any occurrence of “element” has to be interpreted as a vertex and not as an edge. Additional notation will be defined in Section 3.

3 New top SCCs and dominators in subgraphs

Our algorithms are based on the following question: Can we identify a set of vertices that is not 2-connected to the remaining vertices by searching in a strict subgraph of the graph? Note that whenever a separator f is removed from a strongly connected graph G , then there exist both a tSCC and a bSCC in $G \setminus \{f\}$ that were adjacent to f in G and are disjoint. Observe that if, say, the tSCC in $G \setminus \{f\}$ contains only a few vertices, say the set $T \subsetneq V$, then each vertex in T must have a low in-degree in G because all incoming edges to vertices in T in G must have come from f (or be the edge f if f is an edge) or other vertices in T . In our $O(n^2)$ -time algorithm we will search for such “new tSCCs” in the subgraph of G induced by vertices with low in-degree. We will do the same on $Rev(G)$ to detect small new bSCCs in $G \setminus \{f\}$.

Definition 2. *Let G be a directed graph. A subgraph $G[T]$ induced by some vertices T is a new tSCC in $G \setminus \{f\}$ with respect to element f if it is a tSCC in $G \setminus \{f\}$ but has incoming edges from f in G .*

Given f , the top SCC $G[T]$ in a subgraph of $G \setminus \{f\}$ can be identified in time linear in the number of edges in the subgraph. But how can we identify the element f without looking at the whole graph? Assume there exists a vertex r that is not in T but can reach f (and if f is a vertex also $r \neq f$). Since $G[T]$ is a tSCC in $G \setminus \{f\}$, it follows that f dominates every vertex of T in the flow graph $G(r)$. This still holds in any subgraph of G as long as r can reach T in the subgraph. If additionally all incoming edges of the vertices in T are present in the subgraph, we can identify f and T in time linear in the number of edges in the subgraph by finding the dominator f in the flow graph with root r and the tSCC $G[T]$ in the subgraph with f removed.

As edges are missing in the subgraph, it is not a-priori clear how to choose r . Moreover, in our algorithm we cannot afford to identify all SCCs in the current graph G as we only want to spend time proportional to the edges in a strict subgraph of G ; thus we cannot assume that the graph we are considering is strongly connected. This means that, in contrast to Lemma 1, when we identify a dominator f in $G(r)$, the element f might not necessarily be a separator in G . However, for a new tSCC w.r.t. f we still know for the set of vertices T in the tSCC that all vertices in $V \setminus (T \cup \{f\})$ that can reach T in G can reach T only through f (because T has no incoming edges in $G \setminus \{f\}$). Thus there cannot be two disjoint paths from any vertex in $G \setminus (T \cup \{f\})$ to any vertex in T . We summarize this intuition about new tSCCs in the following lemma.

Lemma 3. *Let f be an element such that there exists a new tSCC with respect to f in $G \setminus \{f\}$. Let T be the set of vertices in this new tSCC. Let $W = V \setminus T$ if f is an edge and let $W = V \setminus (T \cup \{f\})$ if f is a vertex. If $W \neq \emptyset$, then there do not exist two disjoint paths from any vertex in W to any vertex in T in G , i.e., no vertex in W is 2-connected to any vertex in T . Additionally, the element f is a dominator in $G(r)$ for every $r \in W$ that can reach f in G .*

Proof. By the definition of a tSCC, the vertices in T are strongly connected in $G \setminus \{f\}$ but have no incoming edge from vertices in W in $G \setminus \{f\}$. Hence in G every path from a vertex in W to a vertex in T contains f . If f is a vertex, we have that the vertices in W are not adjacent to the vertices in T in G ; hence no vertex in W is 2-connected to a vertex in T . By the definition of a new tSCC, all vertices in T are reachable from f in G . Thus f dominates the vertices in T in $G(r)$ for every $r \in W$ that can reach f . \square

We will consider two different kinds of subgraphs, and will first make some general observations about the relation of new tSCCs and dominators in subgraphs to those in the full graph. Let f be an element for which a new tSCC exists in the full graph G . To find f by searching in a subgraph H of G , we have to search in a flow graph with an appropriate root. To this end we will use graphs created from H , which we define next. We will then show that as long as the vertices in the new tSCC are not missing incoming edges in H a new tSCC w.r.t. an element f in the subgraph is a new tSCC w.r.t. element f in the full graph and vice versa.

We will use the following way of *contracting* a set of vertices in a graph $G = (V, E)$: Let B be a set of vertices and let $A = V \setminus B$ be its complement. Contracting the vertices in B to a new vertex b defines a new multi-graph $G' = (V', E')$ where $V' = A \cup \{b\}$ and E' consists of all edges in $G[A]$ and one edge (u, b) for each edge in $(u, v) \in E \cap (A \times B)$ and one edge (b, v) for each $(u, v) \in E \cap (B \times A)$.

Definition 4. Let $G_h = (V_h, E_h)$ be a subgraph of a directed graph $G = (V, E)$, i.e., $V_h \subseteq V$ and $E_h \subseteq G[V_h]$. Let all vertices in V_h that are missing incoming edges in G_h compared to G be contained in the vertex set $B_{G,h} \subseteq V_h$ and let $A_{G,h}$ denote $V_h \setminus B_{G,h}$. Assume $|B_{G,h}| \geq 1$. Let the graph G'_h be the graph G_h with all vertices in $B_{G,h}$ contracted to a single vertex $b_{G,h}$. If $|B_{G,h}| \geq 2$, let \hat{G}_h be the graph G_h with an additional vertex $s_{G,h}$ and one additional edge from $s_{G,h}$ to each vertex in $B_{G,h}$. For $|B_{G,h}| = 1$ let $s_{G,h}$ be the vertex in $B_{G,h}$ and let \hat{G}_h be equal to G_h .

We will search edge-dominators in flow graphs $G'_h(b_{G,h})$ and vertex-dominators in $\hat{G}_h(s_{G,h})$. For our purpose it would not be sufficient to identify vertex-dominators in $G'_h(b_{G,h})$, where all vertices in $B_{G,h}$ are contracted, because we also want to identify dominators that are in $B_{G,h}$ as long as the dominated vertices are in $A_{G,h}$. Note that if we would define the flow graph $\hat{G}_h(s_{G,h})$ in the case $|B_{G,h}| = 1$ in the same way as for $|B_{G,h}| > 1$, then in $\hat{G}_h(s_{G,h})$ the root $s_{G,h}$ could reach the vertices in $A_{G,h}$ only through the vertex in $B_{G,h}$, i.e., the vertex in $B_{G,h}$ would be a vertex-dominator in $\hat{G}_h(s_{G,h})$ independent of the underlying graph G .

In the following consider a subgraph G_h and a set of vertices V_h partitioned into $B_{G,h}$ and $A_{G,h}$ as in the previous definition.

Lemma 5. A set of vertices $T \subseteq A_{G,h}$ induces a tSCC in G_h if and only if it induces a tSCC in G . If $B_{G,h} \neq \emptyset$, then this also holds with G_h replaced by G'_h or \hat{G}_h .

Proof. Assume $B_{G,h} \neq \emptyset$. By the definition of $A_{G,h}$ we have $G_h[T] = G'_h[T] = \hat{G}_h[T] = G[T]$ and that T has no incoming edges in each of G_h , G'_h , and \hat{G}_h if

and only if it has no incoming edges in G . If $B_{G,h} = \emptyset$, the graphs G'_h and \hat{G}_h are undefined but the claim still holds for G_h by the same arguments. \square

If $T \subseteq A_{G,h}$, then for a new tSCC $G[T]$ with respect to f all incoming edges are present in G_h . Thus, also f has to be in G_h because otherwise at least one vertex in T would miss an incoming edge in G_h . This implies the following corollary.

Corollary 6. *For an element f and a set of vertices $T \subseteq A_{G,h}$ we have that T induces a new tSCC with respect to f in $G_h \setminus \{f\}$ if and only if it induces a new tSCC with respect to f in $G \setminus \{f\}$. If $B_{G,h} \neq \emptyset$, then this also holds with G_h replaced by G'_h or \hat{G}_h .*

We will identify elements f for which a new tSCC exists in $G \setminus \{f\}$ by searching for dominators in the flow graphs $G'_h(b_{G,h})$ and $\hat{G}_h(s_{G,h})$. For this we want to compare the reachability of vertices in G'_h and \hat{G}_h with their reachability in G . In particular, we will consider the reachability of vertices in $A_{G,h}$ from vertices in $V \setminus A_{G,h}$.

Lemma 7. *Let P be any path in G from a vertex $w \in V \setminus A_{G,h}$ to a vertex $v \in A_{G,h}$. Let P' be the subpath of P from the last occurrence of a vertex in $B_{G,h}$ to v . For every path P the path P' exists and is contained in G_h and \hat{G}_h . Further, G'_h contains the corresponding path after the contraction of the vertices in $B_{G,h}$ to vertex $b_{G,h}$, i.e., the path P' with the first vertex of P' replaced by $b_{G,h}$.*

Proof. Each vertex in V is either in $V \setminus V_h$, in $A_{G,h}$, or in $B_{G,h}$. Whenever a vertex u on P is in $V \setminus V_h$, there has to be a vertex $u' \in B_{G,h}$ on path P after the occurrence of u because u is not in V_h and thus one of the subsequent vertices on P is missing an incoming edge in G_h , i.e., is contained in $B_{G,h}$. It follows that P' exists, does not contain a vertex in $V \setminus V_h$, and all vertices of P' except the first vertex are in $A_{G,h}$. Thus all edges in P' are incoming edges for some vertex in $A_{G,h}$. All incoming edges of vertices in $A_{G,h}$ are contained in G_h and \hat{G}_h . Furthermore, there is a one-to-one correspondence between incoming edges of vertices in $A_{G,h}$ in G'_h and in G where in G'_h each vertex in $B_{G,h}$ is replaced with $b_{G,h}$. \square

Before we continue, we introduce additional notation to simultaneously analyze edge- and vertex-dominators.

Definition 8. *Let the graph $F_{G,h}$ be equal to \hat{G}_h and the root $r_{G,h}$ be equal to $s_{G,h}$ when vertex-dominators are searched and let $F_{G,h}$ be equal to G'_h and $r_{G,h}$ equal to $b_{G,h}$ when edge-dominators are searched.*

We say that the reachability condition is satisfied in G for an element f and the set of vertices $B_{G,h}$ if the following holds: If f is an edge, then the condition states that f can be reached from some vertex in $B_{G,h}$. If f is a vertex, then the condition states that either f is not in $B_{G,h}$ and can be reached from some vertex in $B_{G,h}$ in G or f is a vertex in $B_{G,h}$ and $|B_{G,h}| \geq 2$.

With the reachability condition and Lemma 7 we can finally specify which elements f with a new tSCC in $G \setminus \{f\}$ we can find in $F_{G,h}(r_{G,h})$.

Corollary 9 (of Lemmata 3 and 7 and Corollary 6). *Assume $B_{G,h} \neq \emptyset$ and let $T \subseteq A_{G,h}$. An element $f \in G$ satisfies the reachability condition for $B_{G,h}$ in G and there is a new tSCC $G[T]$ with respect to f in $G \setminus \{f\}$ if and only if f is a dominator in $F_{G,h}(r_{G,h})$ and $G[T] = G_h[T]$ is a new tSCC with respect to f in $G_h \setminus \{f\}$.*

Proof. By Corollary 6, $G[T]$ is a new tSCC with respect to f in $G \setminus \{f\}$ if and only if $G[T] = G_h[T] = F_{G,h}[T]$ is a new tSCC with respect to f in $G_h \setminus \{f\}$ and $F_{G,h} \setminus \{f\}$, respectively. Note that $r_{G,h} \notin A_{G,h}$ and thus $r_{G,h} \notin T$. Additionally we have in both directions $r_{G,h} \neq f$ by the reachability condition and the definition of dominators. If T is a new tSCC with respect to f in $F_{G,h} \setminus \{f\}$, we further have by Lemma 3 that the element f is a dominator in $F_{G,h}(r_{G,h})$ if and only if $f \neq r_{G,h}$ and f is reachable from $r_{G,h}$. By Lemma 7 this is the case if and only if f satisfies the reachability condition for $B_{G,h}$ in G . \square

In the following two sections we will define specific subgraphs G_h that will allow us to identify every new tSCCs in G that has at most a certain size by searching for dominators f in $F_{G,h}(r_{G,h})$ and tSCCs in $G_h \setminus \{f\}$.

For vertex-connectivity we additionally have to consider one special case, namely if $|B_{G,h}| = 1$, i.e., $B_{G,h} = \{f\}$ for some vertex f . In this case f does not satisfy the reachability condition and we have $r_{G,h} = f$ by Definition 4. We will explicitly identify new tSCCs with respect to such a vertex.

4 2eSCCs and 2vSCCs in $O(n^2)$ time

In the algorithm presented in this section we consider for $i \in \mathbb{N}$ subgraphs $G_i = (V, E_i)$ of a graph $G = (V, E)$ where E_i contains for each vertex in V its first 2^i incoming edges in E (for some arbitrary but fixed ordering of the incoming edges of each vertex). Note that when $i \geq \log(\max_{v \in V} \text{Indeg}_G(v))$, then $G_i = G$.

We will first provide some intuition for the algorithm and its analysis and then formally prove its correctness and running time. Let S be a set of at most 2^i vertices that induces a strongly connected subgraph $G[S]$ of G such that $G[S]$ is either a top or bottom SCC or a new top or bottom SCC with respect to some element. By applying the results from the previous section, we will show that we can detect each such set S by searching for SCCs and dominators in graphs constructed from G_i as defined in Definition 4 (as long as the reachability condition is satisfied).

We will search for top and bottom SCCs by searching for top SCCs in G and $\text{Rev}(G)$. The search for both top and bottom SCCs will ensure that whenever a (new) tSCC and a disjoint (new) bSCC exist in G , we only spend time proportional to the smaller one.

We will start the search for such a set S at $i = 1$. Whenever the search is not successful, we increase i by one, until we have $G_i = G$ or $\text{Rev}(G)_i = \text{Rev}(G)$. In this way we can show that whenever we had to go up to i^* to identify a (new) tSCC in G_i or $\text{Rev}(G)_i$, this subgraph contains $\Omega(2^{i^*})$ vertices. The search in G_i and $\text{Rev}(G)_i$ for i up to i^* will take time proportional to $n \cdot 2^{i^*}$, i.e., $O(n)$ per vertex in the identified subgraph. This observation will allow us to bound the running time by $O(n^2)$.

Definition 10. For a set of vertices S and a set of elements Z , let $G_{S,Z}$ be equal to $G[S \cup Z]$ if $Z \subseteq V$ and equal to $G[S]$ otherwise. Let $G_{V \setminus S}$ be equal to $G[V \setminus S]$.

ALGORITHM 2SCC: 2-edge and 2-vertex strongly connected components
in $O(n^2)$ time

```

1 2SCC( $G$ ) :
2    $\delta \leftarrow \min(\max_{v \in V} \text{Indeg}_G(v), \max_{v \in V} \text{Outdeg}_G(v))$ 
3   for  $i \leftarrow 1$  to  $\lceil \log \delta \rceil - 1$  do
4      $(S, Z) \leftarrow \text{LevelSearch}(G, i)$ 
5     if  $S \neq \emptyset$  then
6       return  $2\text{SCC}(G_{S,Z}) \cup 2\text{SCC}(G_{V \setminus S})$ 
7    $(S, Z) \leftarrow \text{Search}(G)$ 
8   if  $S \neq \emptyset$  then
9     return  $2\text{SCC}(G_{S,Z}) \cup 2\text{SCC}(G_{V \setminus S})$ 
10  else
11  return  $\{G\}$ 

```

Whenever the algorithm identifies a set of vertices S such that $G[S]$ is a tSCC in G or $Rev(G)$ or a new tSCC with respect to some element f in $G \setminus \{f\}$ or $Rev(G) \setminus \{f\}$, it recursively calls itself on $G_{S,Z}$ and $G_{V \setminus S}$ for $Z = \emptyset$ or $Z = \{f\}$, respectively. We will use Lemma 3 to show that in this case every 2SCC of G is completely contained in one either $G_{S,Z}$ or $G_{V \setminus S}$.

Procedure LevelSearch(G, i)

```

1 foreach  $\mathcal{G} \in \{G, Rev(G)\}$  do
2   /*  $2^i < \max_{v \in V} \text{Indeg}_G(v) \implies B_{\mathcal{G},i} \neq \emptyset$  */
3   construct  $\mathcal{G}_i = (V, E_i)$  with  $E_i = \cup_{v \in V} \{\text{first } 2^i \text{ edges in } \text{In}_G(v)\}$ 
4    $B_{\mathcal{G},i} = \{v \mid \text{Indeg}_G(v) > 2^i\}$ 
5    $S \leftarrow \text{TopSCCWithout}(\mathcal{G}_i, B_{\mathcal{G},i})$ 
6   if  $S \neq \emptyset$  then
7     return  $(S, \emptyset)$ 
8   construct flow graph  $F_{\mathcal{G},i}(r_{\mathcal{G},i})$ 
9   if exists dominator  $f$  in  $F_{\mathcal{G},i}(r_{\mathcal{G},i})$  then
10     $S \leftarrow \text{TopSCCWithout}(\mathcal{G}_i \setminus \{f\}, B_{\mathcal{G},i})$ 
11    return  $(S, \{f\})$ 
12    /* 2vSCCs only: */
13  else if  $|B_{\mathcal{G},i}| = 1$  and  $\exists \text{ tSCC} \subsetneq V \setminus \{r_{\mathcal{G},i}\}$  in  $\mathcal{G}_i \setminus \{r_{\mathcal{G},i}\}$  then
14     $S \leftarrow \text{TopSCC}(\mathcal{G}_i \setminus \{r_{\mathcal{G},i}\})$ 
15    return  $(S, \{r_{\mathcal{G},i}\})$ 
16  /*
17  return  $(\emptyset, \emptyset)$ 

```

For the search for such sets S , the Procedure LevelSearch is used when $2^i < \min(\max_{v \in V} \text{Indeg}_G(v), \max_{v \in V} \text{Outdeg}_G(v))$, i.e., both $B_{G,i}$ and $B_{Rev(G),i}$ are

non-empty, and the Procedure **Search** otherwise. The Procedure **LevelSearch** first searches for a tSCC in $\mathcal{G}_i \in \{G_i, Rev(G)_i\}$ that does not contain a vertex in $B_{\mathcal{G},i}$. Following Definition 4, the set $B_{\mathcal{G},i}$ contains all vertices with an in-degree higher than 2^i in \mathcal{G} . If no such tSCC is found, the flow graph $F_{\mathcal{G},i}(r_{\mathcal{G},i})$ is constructed and searched for dominators. If a dominator f is found, a tSCC in $\mathcal{G}_i \setminus \{f\}$ that does not contain a vertex in $B_{\mathcal{G},i}$ is found. We will show that such a tSCC always exists in this case.

For vertex-connectivity we additionally have to consider the special case when $|B_{\mathcal{G},i}| = 1$. In this case we have $B_{\mathcal{G},i} = \{r_{\mathcal{G},i}\}$ and we want to detect when there exists a new tSCC $\mathcal{G}[T]$ with respect to $r_{\mathcal{G},i}$ in $\mathcal{G}_i \setminus \{r_{\mathcal{G},i}\} = \mathcal{G}_i \setminus B_{\mathcal{G},i}$ such that $V \setminus (T \cup \{r_{\mathcal{G},i}\})$ is not empty.

We use **TopSCCWithout**(H, B) to denote the search for a tSCC in a graph H that does not contain a vertex in B . We use Tarjan's linear-time algorithm for this search. We let all procedures that search for an SCC return the set of vertices S in the SCC instead of the subgraph $G[S]$.

Procedure Search(G)

```

1  $S \leftarrow \text{TopOrBottomSCC}(G)$ 
2 if  $\emptyset \subsetneq S \subsetneq V$  then
3   return  $(S, \emptyset)$ 
4 if exists separator  $f$  in  $G$  then
5    $S \leftarrow \text{TopOrBottomSCC}(G \setminus \{f\})$ 
6   return  $(S, \{f\})$ 
7 return  $(\emptyset, \emptyset)$ 
```

To search for sets S in the full graph G , we cannot use the flow graph $F_{\mathcal{G},i}(r_{\mathcal{G},i})$, as the set $B_{\mathcal{G},i}$ would be empty. However, in this case we can afford to check whether the graph is strongly connected and either make progress by separating the strongly connected components from each other or by finding separators in the strongly connected graph. If a separator f is found, disjoint top and bottom SCCs exist after the removal of the separator f . Procedure **Search** returns a top or bottom SCC in $G \setminus \{f\}$ in this case. If the graph is strongly connected and does not contain a separator, then the currently considered graph is a 2SCC. In this case the Procedure **Search** returns the empty set, the recursion stops, and **2SCC**(G) returns the current graph. To find dominators, separators, and SCCs the known linear time algorithms are used (see Section 2).

4.1 Correctness

To show the correctness of the algorithm for 2SCCs, the following three parts are needed: (1) Every step in the algorithm can be executed as described; in particular, (1a) whenever Procedure **LevelSearch** is called the algorithm ensures $2^i < \max_{v \in V} \text{Indeg}_G(v)$ and $2^i < \max_{v \in V} \text{Outdeg}_G(v)$ and thus $B_{\mathcal{G},i} \neq \emptyset$ and (1b) whenever there is no tSCC in \mathcal{G}_i that does not contain a vertex in $B_{\mathcal{G},i}$ but a dominator f is found in the flow graph $F_{\mathcal{G},i}(r_{\mathcal{G},i})$, then there exists a new tSCC in $\mathcal{G}_i \setminus \{f\}$ w.r.t. f that does not contain a vertex in $B_{\mathcal{G},i}$. (2) Whenever a set S and a set Z are identified, every 2SCC of G is completely contained in either $G_{S,Z}$ or $G_{V \setminus S}$ and both $G_{S,Z}$ and $G_{V \setminus S}$ are non-empty and thus strict

subgraphs of G . (3) Whenever no set S is identified in G , then G is a 2SCC. In the following we will first show Part (1b) and then introduce Lemma 12 and Theorem 13 to show Parts (2) and (3), respectively.

The application of the following lemma to the graph $F_{\mathcal{G},i}$ shows Part (1b) as follows. Recall that by Corollary 6 we have that a set of vertices $T \subseteq V \setminus B_{\mathcal{G},i}$ induces a new tSCC in $F_{\mathcal{G},i}$ if and only if it induces a new tSCC in \mathcal{G}_i . Let f be a dominator in $F_{\mathcal{G},i}(r_{\mathcal{G},i})$. Let $T \subseteq V \setminus (B_{\mathcal{G},i} \cup \{r_{\mathcal{G},i}\})$ be a set of vertices that induces a new tSCC w.r.t. f in $F_{\mathcal{G},i} \setminus \{f\}$. The first part of the following lemma shows that if such a set T exists, then all vertices in T are dominated by f in $F_{\mathcal{G},i}(r_{\mathcal{G},i})$. The second part shows that if no such tSCC exists in $F_{\mathcal{G},i} \setminus \{f\}$ for the dominator f , then there exists a set of vertices $S \subseteq V \setminus B_{\mathcal{G},i}$ that induces a tSCC $\mathcal{G}[S]$ in \mathcal{G}_i that is not reachable from $r_{\mathcal{G},i}$. Note that Procedure `LevelSearch` would detect S before it searches for dominators in $F_{\mathcal{G},i}(r_{\mathcal{G},i})$.

Lemma 11. *Let $G(r)$ be a flow graph for a directed graph G and some vertex r . Let f be a dominator in $G(r)$ and let D be the set of vertices dominated by f .*

1. *If there exists a set of vertices T that induces a new tSCC $G[T]$ in $G \setminus \{f\}$ with respect to f , either (a) T is a subset of D and contains a vertex with an edge from f in G if f is a vertex or the vertex w if $f = (u, w)$ is an edge or (b) f is a vertex and T contains r .*
2. *If for no set $T \subseteq D$ a new tSCC $G[T]$ exists in $G \setminus \{f\}$, then there exists a tSCC in G that is contained in $G \setminus G(r)$.*

Proof. 1. If $f = (u, w)$ is an edge, then the only edge from vertices in $V \setminus T$ to vertices in T in G is the edge f , thus clearly T contains w . Since $G[T]$ is strongly connected in $G \setminus \{f\}$, all vertices in T are reachable from r in G and have a path to each other that avoids f , i.e., we have $T \subseteq D$ (Case (a)).

If f is a vertex and the vertices in T induce a new tSCC $G[T]$ in $G \setminus \{f\}$ with respect to f , some vertex in T has an incoming edge from f in G . As f is reachable from r , also T is in $G(r)$. We have either $T \subseteq D$ (Case (a)) or there is a path from r to T in $G(r)$ that avoids f . Thus if r is not contained in T , then T has an incoming edge from a vertex in $G(r) \setminus \{f\}$, a contradiction to $G[T]$ being a tSCC in $G \setminus \{f\}$. Hence it must hold that $r \in T$ (Case (b)).

2. Since D is dominated by f in $G(r)$, there are no edges from vertices in $G(r) \setminus (D \cup \{f\})$ to vertices in D in $G \setminus \{f\}$. Thus either D contains a tSCC in $G \setminus \{f\}$ or there are vertices in $G \setminus G(r)$ that have edges to vertices in D . These vertices are not reachable from vertices in $G(r)$, i.e., they have to contain a tSCC in G . \square

Lemma 12. *Assume $B_{\mathcal{G},i}$ for $\mathcal{G} \in \{G, Rev(G)\}$ is non-empty whenever Procedure `LevelSearch` is called. If (1) Procedure `LevelSearch` or (2) Procedure `Search` return a non-empty set S and a set Z for graph G , then every 2SCC of G is completely contained in one of $G_{S,Z}$ and $G_{V \setminus S}$. Additionally, the set $V \setminus (S \cup Z)$ is non-empty.*

Proof. 1. Let $\mathcal{G} \in \{G, Rev(G)\}$ be the graph in which S is identified. First note that \mathcal{G}_i is a subgraph of \mathcal{G} such that all vertices in \mathcal{G}_i that are missing

incoming edges in \mathcal{G}_i are contained in $B_{\mathcal{G},i}$. Remember that the flow graph $F_{\mathcal{G},i}(r_{\mathcal{G},i})$ is specified Definition 8.

In Procedure `LevelSearch` a set of vertices S can be identified in three ways: (a) $\mathcal{G}[S]$ is a tSCC in \mathcal{G}_i that does not include vertices in $B_{\mathcal{G},i}$ and $Z = \emptyset$; (b) the algorithm finds a dominator f in $F_{\mathcal{G},i}(r_{\mathcal{G},i})$, $Z = \{f\}$, and $\mathcal{G}[S]$ is a new tSCC in $\mathcal{G}_i \setminus \{f\}$ with respect to f that does not include a vertex in $B_{\mathcal{G},i}$; or (c) vertex-connectivity is considered and we have for $f = r_{\mathcal{G},i}$ the case that $B_{\mathcal{G},i} = \{f\} = Z$ and $\mathcal{G}[S]$ is a new tSCC in $\mathcal{G}_i \setminus \{f\}$ with respect to f .

In Case (a) $\mathcal{G}[S]$ is a tSCC in \mathcal{G} by Lemma 5. Note that every 2SCC is strongly connected, thus it clearly holds that in this case each 2SCC of G is completely contained in one of $G_{S,Z}$ and $G_{V \setminus S}$.

In the Cases (b) and (c) $\mathcal{G}[S]$ is a new tSCC in $\mathcal{G} \setminus \{f\}$ with respect to f by Corollary 6. Thus by Lemma 3 no vertex in $V \setminus (S \cup \{f\})$ is 2-connected to any vertex in S in \mathcal{G} . Thus also in this case each 2SCC of G is completely contained in one of $G_{S,Z}$ and $G_{V \setminus S}$.

Note that $V \setminus (S \cup Z) = \emptyset$ is explicitly avoided in Case (c). Thus $V \setminus (S \cup Z) = \emptyset$ can only happen in Case (b) and if $Z = B_{\mathcal{G},i}$, as otherwise we have $B_{\mathcal{G},i} \cap V \setminus (S \cup Z) \neq \emptyset$. If $Z = B_{\mathcal{G},i}$, we have $|B_{\mathcal{G},i}| = 1$ and $Z = \{f\}$, where f is a vertex. Thus $F_{\mathcal{G},i} = \hat{\mathcal{G}}_i$ by Definition 8 and $B_{\mathcal{G},i} = \{r_{\mathcal{G},i}\}$ by Definition 4. However, the root $r_{\mathcal{G},i}$ cannot be a dominator in $F_{\mathcal{G},i}(r_{\mathcal{G},i})$, a contradiction.

2. In Procedure `Search` a set of vertices S can be identified in two ways: (a) $G[S]$ is a top or bottom SCC in G with $S \neq V$ and $Z = \emptyset$; or (b) G is strongly connected, the algorithm finds a separator f in G , $Z = \{f\}$, and $G[S]$ is a new top or bottom SCC in $G \setminus \{f\}$. In Case (a) each 2SCC is completely contained in either $G_{S,Z}$ or $G_{V \setminus S}$ by the fact that every 2SCC is strongly connected; in Case (b) each 2SCC is completely contained in either $G_{S,Z}$ or $G_{V \setminus S}$ by the definition of a 2SCC and Lemma 3. In both cases it clearly holds that $V \setminus (S \cup Z) \neq \emptyset$ because $G \setminus Z$ contains a top and a bottom SCC that are disjoint and only one of them is contained in $G[S]$. \square

Theorem 13 (Correctness). *Let G^* be a simple directed graph. $2SCC(G^*)$ computes the 2SCCs of G^* .*

Proof. We denote by G the graph for which $2SCC(G)$ was called, which is equal to G^* for the initial call to $2SCC$ and equal to a subgraph of G^* for all recursive calls to $2SCC$. Whenever $2SCC(G)$ identifies a set S and recursively calls itself on $G_{S,Z}$ and $G_{V \setminus S}$ for $Z = \emptyset$ or $Z = \{f\}$, then by Lemma 12 each 2SCC of G is completely contained in one of $G_{S,Z}$ and $G_{V \setminus S}$ and thus is a 2SCC in $G_{S,Z}$ or $G_{V \setminus S}$. Note that each 2SCC of a subgraph of G is by definition a subset of a 2SCC of G . Hence the union of the 2SCCs of $G_{S,Z}$ and $G_{V \setminus S}$ yields the 2SCCs of G . Furthermore, by Lemma 12 both S and $V \setminus (S \cup Z)$ are non-empty, i.e., the recursion is on strict subgraphs of G and thus the algorithm will terminate.

To show that the algorithm will not terminate until it has correctly identified all 2SCCs of G^* , it remains to show that whenever $2SCC(G)$ is called, it either identifies a set S and recursively calls itself on $G_{S,Z}$ and $G_{V \setminus S}$ or the graph G is

strongly connected and does not contain a separator, i.e., the graph G is a 2SCC. For this it is sufficient to show that if the algorithm has not identified a set S in the for-loop, then it either identifies a set by the call to Procedure **Search** or G is strongly connected and does not contain a separator. If G is not strongly connected, then Procedure **Search** will return a tSCC or a bSCC in G . If G is strongly connected, then either G does not contain separators or Procedure **Search** will find a separator f and identify a tSCC or a bSCC in $G \setminus \{f\}$, which both exist by the definition of a separator. \square

4.2 Runtime

Towards the runtime analysis of the algorithm for 2SCCs, we first use the results of Section 3 to show a relation between (new) tSCCs we can find in the subgraph G_i and the number of vertices in the (new) tSCC.

Lemma 14. *Let G be a simple directed graph.*

1. *If a set of vertices $S \subsetneq V$ with $|S| \leq 2^i$ induces a tSCC $G[S]$ in G , then $G_i[S] = G[S]$ is a tSCC in G_i and $S \subseteq V \setminus B_{G,i}$.*
2. *If (a) $B_{G,i} \neq \emptyset$, (b) there is an element f that satisfies the reachability condition, and (c) for some set of vertices S with $|S| \leq 2^i$ there exists a new tSCC $G[S]$ in $G \setminus \{f\}$ with respect to f , then f is a dominator in $F_{G,i}(r_{G,i})$, $S \subseteq V \setminus B_{G,i}$, and $G_i[S] = G[S]$ is a new tSCC in $G_i \setminus \{f\}$ with respect to f .*
3. *If for vertex-connectivity we have $|B_{G,i}| = 1$ and for the element f contained in $B_{G,i}$ there exists a new tSCC $G[S]$ induced by some set of vertices S in $G \setminus \{f\}$ with respect to f , then $S \subseteq V \setminus B_{G,i}$, and $G_i[S] = G[S]$ is a new tSCC in $G_i \setminus \{f\}$ with respect to f .*

Proof. Consider any graph \tilde{G} in which S induces a tSCC. Since a tSCC has no incoming edges, all incoming edges of vertices in S have to come from other vertices in S . As there are at most 2^i vertices in S , each vertex in S can have an in-degree of at most $2^i - 1$ in \tilde{G} .

1. Using $\tilde{G} = G$, we have $S \subseteq V \setminus B_{G,i}$ and the claim follows directly from Lemma 5.
2. Using $\tilde{G} = G \setminus \{f\}$, in G each vertex in S can have at most one additional incoming edge compared to $G \setminus \{f\}$, namely an edge from f if f is a vertex and the edge f if f is an edge. Thus we can bound the in-degree in G of each vertex in S by 2^i . We have $S \subseteq V \setminus B_{G,i}$ and the claim follows from Corollary 9.
3. Since f is the only vertex in $B_{G,i}$, we clearly have $S \subseteq V \setminus B_{G,i}$. The claim follows from Corollary 6. \square

Corollary 15. *If there is no set of vertices T with $T \cap B_{G,i} = \emptyset$ that induces a tSCC $G_i[T]$ in G_i for each of $G \in \{G, Rev(G)\}$ but G is not strongly connected, then each top or bottom SCC in G has more than 2^i vertices.*

- Lemma 16.** 1. If for $\mathcal{G} \in \{G, \text{Rev}(G)\}$ there is a dominator f in $F_{\mathcal{G},i+1}(r_{\mathcal{G},i+1})$ or $B_{\mathcal{G},i+1} = \{f\} = \{r_{\mathcal{G},i+1}\}$ and there is a set of vertices S with $S \cap B_{\mathcal{G},i+1} = \emptyset$ that induces a new tSCC $\mathcal{G}_{i+1}[S]$ in $\mathcal{G}_{i+1} \setminus \{f\}$ with respect to f but f is not a dominator in $F_{\mathcal{G},i}(r_{\mathcal{G},i})$ nor $B_{\mathcal{G},i} = \{f\} = \{r_{\mathcal{G},i}\}$, then $|S| > 2^i$.
2. If G is strongly connected and there is a separator f in G but for each $\mathcal{G} \in \{G, \text{Rev}(G)\}$ there is no dominator in $F_{\mathcal{G},i}(r_{\mathcal{G},i})$ nor $B_{\mathcal{G},i} = \{f\} = \{r_{\mathcal{G},i}\}$, then each top or bottom SCC in $G \setminus \{f\}$ has more than 2^i vertices.

Proof. 1. By Corollary 6 we have that $\mathcal{G}_{i+1}[S]$ is a new tSCC in $\mathcal{G} \setminus \{f\}$ with respect to f . Further, either f satisfies the reachability condition for $B_{\mathcal{G},i+1}$, and thus also satisfies the reachability condition for $B_{\mathcal{G},i} \supseteq B_{\mathcal{G},i+1}$, or we have $B_{\mathcal{G},i+1} = \{f\} = \{r_{\mathcal{G},i+1}\}$. In the second case we have by $B_{\mathcal{G},i} \supseteq B_{\mathcal{G},i+1}$ that either $B_{\mathcal{G},i} = \{f\} = \{r_{\mathcal{G},i}\}$ or f satisfies the reachability condition for $B_{\mathcal{G},i}$. Thus if $|S| \leq 2^i$, we obtain a contradiction by Lemma 14.

2. Let f^R be equal to the reverse edge of f if f is an edge and equal to f if f is a vertex. Since f is a separator in a strongly connected graph, we have that there exists a new tSCC with respect to f in $G \setminus \{f\}$ and a new tSCC with respect to f^R in $\text{Rev}(G) \setminus \{f^R\}$. Let \mathcal{G} be the graph in $\{G, \text{Rev}(G)\}$ that contains the smaller of these two new tSCCs and let S be the set of vertices in this new tSCC. Since \mathcal{G} is strongly connected and we exclude the case $B_{\mathcal{G},i} = \{f\} = \{r_{\mathcal{G},i}\}$, we have that f and f^R satisfy the reachability condition, respectively. Thus if $|S| \leq 2^i$, we obtain a contradiction by Lemma 14.2. \square

In the runtime analysis we will argue that whenever the search for a set S stops at a certain level i^* , the algorithm has spent time at most proportional to the number of vertices in the smaller set of S and $V \setminus S$. Additional to Corollary 15 and Lemma 16, we will use the fact that if a graph is not strongly connected, it contains a top and a bottom SCC that are disjoint, which implies that one of them contains at most half of the vertices of the graph.

Theorem 17 (Runtime 2eSCCs). *Algorithm 2SCC for 2eSCCs can be implemented in time $O(n^2)$.*

Proof. We denote by n and m the number of vertices and edges in the input graph G^* and by n' the number of vertices in the graph G in the current level of recursion in Algorithm 2SCC. Let δ denote the minimum of $\max_{v \in V} \text{Indeg}_G(v)$ and $\max_{v \in V} \text{Outdeg}_G(v)$. Let S be a non-empty set of vertices returned by `LevelSearch` or `Search`.

To efficiently construct the graphs \mathcal{G}_i for $1 \leq i \leq \lceil \log \delta \rceil - 1$ and $\mathcal{G} \in \{G, \text{Rev}(G)\}$ we will maintain for all vertices w a list of $\text{In}(w)$ and a list of $\text{Out}(w)$. We do not update this data structure immediately when G is split into $G[S]$ and $G[V \setminus S]$ but remove obsolete entries in $\text{In}(w)$ and $\text{Out}(w)$ whenever we encounter them while constructing \mathcal{G}_i . This can happen at most once for each entry and thus takes total time $O(m) \in O(n^2)$.

We first analyze the time per iteration i of the for-loop in `2SCC(G)`. In iteration i the algorithm calls `LevelSearch(G, i)`. Searching on both G and $\text{Rev}(G)$ only increases the running time by a factor of two; we analyze the search on G , the search on $\text{Rev}(G)$ is analogous. Given the above data structure,

constructing G_i , $F_{G,i}$, and determining $B_{G,i}$ takes time $O(n' \cdot 2^i)$. Finding SCCs and dominators in G_i and $F_{G,i}$ can all be done in time linear in the number of edges in G_i , i.e., in time $O(n' \cdot 2^i)$. Thus `LevelSearch`(G, i) and hence iteration i of the outer for-loop takes time $O(n' \cdot 2^i)$.

The search for SCCs and separators in `Search` can be done in linear time in the number of edges, which can be bounded by $O(n' \cdot \delta)$.

Let $i^* < \lceil \log \delta \rceil$ be the last iteration of `2SCC`(G), i.e., the iteration before `2SCC`(G) returns G (Case (1)) or recursively calls itself on the subgraphs $G[S]$ and $G[V \setminus S]$ (Case (2)). The time spent in the iterations $i = 1$ up to i^* forms a geometric series that can be bounded by $O(n' \cdot 2^{i^*})$. If $i^* = \lceil \log \delta \rceil - 1$, we have $\delta \leq 2^{i^*+1}$ and thus also the time spent in `Search` can be bounded with $O(n' \cdot 2^{i^*})$.

Case (1): If `2SCC`(G) returns G , the recursion stops. In this case we bound the running time with $c_1 \cdot n' \cdot \delta \leq c_1 \cdot n'^2$ for some constant c_1 .

Case (2): In iteration $i^* - 1$ no top or bottom SCC or new top or bottom SCC was detected in G_{i^*-1} . Note that the case $B_{\mathcal{G}, i^*-1} = \{r_{\mathcal{G}, i^*-1}\}$ can only occur for vertex-connectivity. Thus we have by Corollary 15 and Lemma 16 that $|S| > 2^{i^*-1}$. Let $\mathcal{G} \in \{G, \text{Rev}(G)\}$ be the graph in which S was detected at level i^* . If (a) G is strongly connected, then $\mathcal{G}[S]$ is a new tSCC in $\mathcal{G} \setminus \{f\}$ with respect to some element f and we have that there exists a new bSCC in $\mathcal{G} \setminus \{f\}$ with respect to f . This new bSCC is contained in $V \setminus S$ and by Lemma 16.2 contains more than 2^{i^*-1} vertices, i.e., we have $|V \setminus S| > 2^{i^*-1}$. If (b) G is not strongly connected, then there exist a top and a bottom SCC in G which are disjoint and both have more than 2^{i^*-1} vertices by Corollary 15. Since $G[S]$ is a strongly connected subgraph of G , it has to be completely contained in either the top or the bottom SCC. Thus also in this case we have $|V \setminus S| > 2^{i^*-1}$. Let $\min(|S|, |V \setminus S|)$ be denoted by n_e . By the definition of n_e we have $n_e \leq n'/2$. By the analysis above, without the recursion, `2SCC`(G) spends time at most $c \cdot n' \cdot n_e$ for some constant c in this case. In the following we choose $c \geq c_1$.

We show next that the total time spent in all calls to `2SCC`(G), and thus the total running time of Algorithm `2SCC` for 2eSCCs, can be bounded by $f(n) = 2cn^2$. In Case (1), i.e., if the recursion stops, we have $f(n') = c_1 n'^2 < 2cn'^2$. In Case (2) we have by induction, and in particular for $n' = n$,

$$\begin{aligned} f(n') &\leq f(n_e) + f(n' - n_e) + cn'n_e, \\ &= 2cn_e^2 + 2c(n' - n_e)^2 + cn'n_e, \\ &= 2cn_e^2 + 2cn'^2 - 4cn'n_e + 2cn_e^2 + cn'n_e, \\ &= 2cn'^2 + 4cn_e^2 - 3cn'n_e, \\ &\leq 2cn'^2, \end{aligned}$$

where the last inequality follows from $n_e \leq n'/2$. □

Theorem 18 (Runtime 2vSCCs). *Algorithm 2SCC for 2vSCCs can be implemented in time $O(n^2)$.*

Proof. The proof is analogous to the proof of Theorem 17 except that the recursion is on $G[S \cup \{f\}]$ and $G[V \setminus S]$ whenever a vertex-dominator or articulation point f was found, and we have to exclude the case $B_{\mathcal{G}, i^*-1} = \{f\} = \{r_{\mathcal{G}, i^*-1}\}$ to apply Lemma 16. If we had $B_{\mathcal{G}, i^*-1} = \{f\} = \{r_{\mathcal{G}, i^*-1}\}$ for $\mathcal{G} \in \{G, \text{Rev}(G)\}$ and some vertex f for which a new tSCC $\mathcal{G}[T]$ induced by some set of vertices T

with $V \setminus (T \cup \{f\}) \neq \emptyset$ exists in $\mathcal{G} \setminus \{f\}$, then by Lemma 14.3 the set T would have been identified in iteration $i^* - 1$, a contradiction to i^* being the last iteration of $2\text{SCC}(G)$.

To efficiently update the adjacency lists, we update the lists of f for each of $G[S \cup \{f\}]$ and $G[V \setminus S]$ immediately when the recursive call occurs, while for all other vertices the same argument as before applies.

Let $n_v = \min(|S|, |V \setminus S| - 1)$. By the definition of n_v it holds that $n'/2 \geq n_v$. Let $c \cdot n' \cdot n_v$ be the runtime of $2\text{SCC}(G)$ without recursion whenever the recursion does not stop. We will stop the recursion whenever the number of remaining vertices is at most 8, i.e., a constant. In this case we use the $O(mn)$ -algorithm to determine the 2vSCCs of G in constant time. This can happen at most n times. With $3n'/4 \geq 6n_v/5$ and $n'/4 \geq 2 > 9/5$ we have $5n' \geq 6n_v + 9$, which we use in the following to obtain an upper bound of $f(n) = 3cn^2$ on the total running time of Algorithm 2SCC for 2vSCCs . We have

$$\begin{aligned}
f(n') &\leq f(n_v + 1) + f(n' - n_v) + cn'n_v, \\
&= 3c(n_v + 1)^2 + 3c(n' - n_v)^2 + cn'n_v, \\
&= 3cn_v^2 + 6cn_v + 3c + 3cn'^2 - 6cn'n_v + 3cn_v^2 + cn'n_v, \\
&= 3cn'^2 + 6cn_v^2 - 5cn'n_v + 6cn_v + 3c, \\
&\leq 3cn'^2. \quad \square
\end{aligned}$$

5 An $O(m^2/\log n)$ -time algorithm for 2eSCCs

In the algorithm presented in this section we will assume that each vertex in the input graph has constant in- and out-degree. We will show how every graph with m edges can be transformed in $O(m)$ time into a graph with in- and out-degree at most three with equivalent 2eSCCs , $\Theta(m)$ vertices and $\Theta(m)$ edges. We are not aware of such a transformation for 2vSCCs .

We will combine our results for new tSCCs and dominators in subgraphs in Section 3 with a local-search technique used for Büchi games by Chatterjee et al. [6]. The algorithm will repeatedly remove edges from the graph. We will use that all sets of vertices that have only one or no incoming edge after the removal of some edges from the graph but had more incoming edges before must have lost an incoming edge. In each iteration we want to find all such sets of vertices that are additionally strongly connected, i.e., all tSCCs and new tSCCs that had more incoming edges at the beginning of the previous iteration. Following Chatterjee et al. [6], we start local searches around vertices that lost edges whenever not too many edges have been removed from the graph. Similar to the previous section, we search for tSCCs and new tSCCs . For each of the local searches started at a vertex r we consider the subgraph G_r reached by a breadth-first search on $\text{Rev}(G)$ of depth roughly $\log n$. In this way we will find all such tSCCs and new tSCCs that contain at most $O(\log n)$ vertices. With the results from Section 3, this will imply that when we afterwards search for SCCs and bridges in the full graph, we will separate sets with more than $\Omega(\log n)$ vertices from each other; this can happen at most $O(m/\log n)$ times. The constant-degree assumption allows us to bound the work done in the local searches. Algorithm 2ESCC-SPARSE shows the pseudocode of our approach.

The repeat-until loop without the local searches is equivalent to the simple

$O(mn)$ -algorithm for 2eSCCs. The local searches work analogous to the level searches in the $O(n^2)$ -time algorithm for 2eSCCs. The only differences are that different subgraphs are used and that in Algorithm 2ESCC-SPARSE we do not partition the graph recursively but repeatedly remove edges from the graph. We show next how to transform a graph to a constant degree graph without changing the 2eSCCs.

Lemma 19. *Let $G = (V, E)$ be a simple directed graph with some fixed ordering in $\text{In}(v)$ and $\text{Out}(v)$ for each $v \in V$. Let $t_v = \max(\text{Indeg}(v), \text{Outdeg}(v))$. Let \tilde{G} be G with every vertex $v \in V$ with $t_v > 3$ replaced by t_v vertices V_v as follows. Let v_0, \dots, v_{t_v-1} be the vertices in V_v . We add for each v_i an edge $(v_i, v_{(i+1) \bmod t_v})$ and an edge $(v_i, v_{(i-1) \bmod t_v})$ to \tilde{G} . Additionally we add for each edge $(u, v) \in E$ that is the i -th edge in $\text{Out}(u)$ and the j -th edge in $\text{In}(v)$ an edge (u_{i-1}, v_{j-1}) to \tilde{G} . For vertices $v \in V$ with $t_v \leq 3$, let $V_v = \{v\}$. Then $G[S]$ is a 2eSCC in G if and only if $\tilde{G}[\cup_{u \in S} V_u]$ is a 2eSCC in \tilde{G} .*

Proof. Since every subgraph $\tilde{G}[V_v]$ is 2-edge strongly connected, it is sufficient to show that there are two edge-disjoint paths from a vertex u to a vertex $v \neq u$ in G if and only if there are two edge-disjoint paths from some vertex $\tilde{u} \in V_u$ to some vertex $\tilde{v} \in V_v$ with $v \neq u$ in \tilde{G} . We can assume w.l.o.g. that the edge-disjoint paths are simple.

\Leftarrow : Let \tilde{P}_1 and \tilde{P}_2 be two edge-disjoint paths in \tilde{G} from a vertex $\tilde{u} \in V_u$ to a vertex $\tilde{v} \in V_v$ with $v \neq u$. We can construct two edge-disjoint paths P_1 and P_2 from u to v in G by simply removing all edges that are contained in a subgraph $\tilde{G}[V_w]$ for some $w \in V$ and replacing all vertices $\tilde{w} \in V_w$ with w in the remaining edges. As there is a one-to-one relation between edges between subgraphs $\tilde{G}[V_w]$ and edges in E , the paths P_1 and P_2 must be edge-disjoint.

\Rightarrow : Let P_1 and P_2 be two edge-disjoint paths from a vertex u to a vertex $v \neq u$ in G . We construct two edge-disjoint paths \tilde{P}_1 and \tilde{P}_2 from a vertex $\tilde{u} \in V_u$ to a vertex $\tilde{v} \in V_v$ in \tilde{G} . First add for all edges in the path P_ℓ , $\ell \in \{1, 2\}$, the corresponding edges between two different subgraphs $\tilde{G}[V_w]$ in \tilde{G} to \tilde{P}_ℓ . It remains to connect the edges in \tilde{P}_ℓ within subgraphs $\tilde{G}[V_w]$ with $t_w > 3$. For any w that is only contained in one of P_1 and P_2 , we can select some arbitrary path within $\tilde{G}[V_w]$ that connects the edges in \tilde{P}_ℓ . Let w be a vertex in both P_1 and P_2 . Let w_0, \dots, w_{t_w-1} be the vertices in V_w . We connect the path \tilde{P}_1 in the subgraph $\tilde{G}[V_w]$ by using the edges from w_i to $w_{(i+1) \bmod t_w}$ for $0 \leq i < t_w$ and we connect the path \tilde{P}_2 in the subgraph $\tilde{G}[V_w]$ by using the edges from w_i to $w_{(i-1) \bmod t_w}$. In this way the paths \tilde{P}_1 and \tilde{P}_2 are edge-disjoint. \square

5.1 Correctness

The following lemma shows the correctness of Procedure LocalSearch.

Lemma 20. *Let $\mathcal{G} \in \{G, \text{Rev}(G)\}$.*

1. *If there is no tSCC without a vertex in $B_{\mathcal{G},r}$ in \mathcal{G}_r , then $B_{\mathcal{G},r} \neq \emptyset$.*
2. *Let $B_{\mathcal{G},r} \neq \emptyset$. Whenever there is an edge-dominator e in $\mathcal{G}'_r(b_{\mathcal{G},r})$, then there exists a tSCC without vertices in $B_{\mathcal{G},r}$ in $\mathcal{G}_r \setminus \{e\}$.*

ALGORITHM 2ESCC-SPARSE: 2-edge strongly connected components in $O(m^2/\log n)$ time

Input: graph G with $\text{Indeg}(v) \leq 3$ and $\text{Outdeg}(v) \leq 3$ for all $v \in V$

```

1  $q \leftarrow \lceil \log n \rceil$ 
2  $d \leftarrow \lceil \varepsilon \log n \rceil$  for some  $\varepsilon \in (0, 1)$ 
3 repeat
4   find SCCs  $C_1, \dots, C_g$  of  $G$ 
5   remove edges between SCCs of  $G$ 
6    $J \leftarrow \emptyset$ 
7   for  $i \leftarrow 1$  to  $g$  do
8      $X \leftarrow \text{Bridges}(C_i)$ 
9      $G \leftarrow G \setminus X$ 
10    add vertices adjacent to  $X$  to  $J$ 
11  if  $0 < |J| < q$  then
12    repeat
13       $S \leftarrow \text{LocalSearch}(G, J)$ 
14      remove edges between  $S$  and  $V \setminus S$  from  $G$ 
15      add vertices that lost edges to  $J$ 
16    until  $S = \emptyset$  or  $|J| \geq q$ 
17 until  $J = \emptyset$ 
18 return  $C_1, \dots, C_g$ 
19 LocalSearch( $G, J$ ):
20   foreach  $r \in J$  do
21     foreach  $\mathcal{G} \in \{G, \text{Rev}(G)\}$  do
22       find set of vertices  $V_{\mathcal{G},r}$  with distance at most  $d$  to  $r$  in  $\mathcal{G}$  with
23       BFS from  $r$  in  $\text{Rev}(\mathcal{G})$ 
24        $\mathcal{G}_r \leftarrow \mathcal{G}[V_{\mathcal{G},r}]$ 
25        $B_{\mathcal{G},r} \leftarrow \{v \in V_{\mathcal{G},r} \mid \text{In}_{\mathcal{G}}(v) \cap ((V \setminus V_{\mathcal{G},r}) \times \{v\}) \neq \emptyset\}$ 
26        $T \leftarrow \text{TopSCCWithout}(\mathcal{G}_r, B_{\mathcal{G},r})$ 
27       /*  $B_{\mathcal{G},r} = \emptyset \Rightarrow$  no edge from  $V \setminus V_{\mathcal{G},r}$  to  $V_{\mathcal{G},r} \Rightarrow T \neq \emptyset$  */
28       if  $T \neq \emptyset$  then
29         if exists edge from  $T$  to  $V \setminus T$  in  $\mathcal{G}$  then
30           return  $T$ 
31         else /*  $T = V_{\mathcal{G},r}$  and  $\mathcal{G}_r$  is tSCC and bSCC in  $\mathcal{G}$  */
32           if exists bridge  $e$  in  $\mathcal{G}_r$  then
33              $U \leftarrow \text{TopOrBottomSCC}(\mathcal{G}_r \setminus \{e\})$ 
34             return  $U$ 
35         else
36           let  $\mathcal{G}'_r$  be  $\mathcal{G}_r$  with  $B_{\mathcal{G},r}$  contracted to  $b_{\mathcal{G},r}$ 
37           if exists edge-dominator  $e$  in  $\mathcal{G}'_r(b_{\mathcal{G},r})$  then
38              $U \leftarrow \text{TopSCCWithout}(\mathcal{G}_r \setminus \{e\}, B_{\mathcal{G},r})$ 
39             return  $U$ 
40   return  $\emptyset$ 

```

3. If Procedure `LocalSearch` returns a non-empty set of vertices S , then
 - (a) there exists an edge between vertices in S and vertices in $V \setminus S$ in G
 - and (b) each 2eSCC of G is completely contained in either $G[S]$ or $G[V \setminus S]$.

Proof. Note that for each vertex r and each graph $\mathcal{G} \in \{G, Rev(G)\}$ the graphs \mathcal{G}_r and \mathcal{G}'_r and the set $B_{\mathcal{G},r}$ satisfy Definition 4, respectively.

1. If $B_{\mathcal{G},r} = \emptyset$, we have that there are no edges from vertices in $V \setminus V_{\mathcal{G},r}$ to vertices in $V_{\mathcal{G},r}$ in \mathcal{G} . Thus in this case the subgraph \mathcal{G}_r has to contain a tSCC. This tSCC is also a tSCC in \mathcal{G} by Lemma 5.
2. By the construction of \mathcal{G}'_r , all vertices dominated by e in $\mathcal{G}'_r(b_{\mathcal{G},r})$ are in $V_{\mathcal{G},r} \setminus B_{\mathcal{G},r}$. By Lemma 11 there either exists a tSCC in $\mathcal{G}'_r \setminus \{e\}$ that consists of vertices dominated by e or there exists a tSCC in $\mathcal{G}'_r \setminus \mathcal{G}'_r(b_{\mathcal{G},r})$. In the first case the tSCC is also a tSCC in $\mathcal{G}_r \setminus \{e\}$ by Lemma 5. For the second case note that $\mathcal{G}'_r \setminus \mathcal{G}'_r(b_{\mathcal{G},r}) \subseteq \mathcal{G}_r \setminus B_{\mathcal{G},r}$. Thus by Lemma 5 a tSCC in $\mathcal{G}'_r \setminus \mathcal{G}'_r(b_{\mathcal{G},r})$ is a tSCC in \mathcal{G}_r (and $\mathcal{G}_r \setminus \{e\}$) that does not contain vertices in $B_{\mathcal{G},r}$.
3. Procedure `LocalSearch` returns a non-empty set for some $r \in J$ in three cases. It first searches for a set T with $T \cap B_{\mathcal{G},r} = \emptyset$ that induces a tSCC $\mathcal{G}[T]$ in \mathcal{G}_r . If the search is successful, it determines whether there exists an edge between vertices in T and in $V \setminus T$ in G . If this is satisfied, it returns the set T (Case (1)). If this is not satisfied, i.e., $G[T]$ is a top and a bottom SCC in G , then the procedure returns set of vertices U that induces a new top or bottom SCC $G[U]$ with respect to a bridge e in $G[T] \setminus \{e\}$ if one exists (Case (2)). If \mathcal{G}_r does not contain a tSCC without a vertex in $B_{\mathcal{G},r}$, then the procedure searches for an edge-dominator e in $\mathcal{G}'_r(b_{\mathcal{G},r})$ and a set of vertices U with $U \cap B_{\mathcal{G},r} = \emptyset$ that induces new tSCC $\mathcal{G}[U]$ with respect to e in $\mathcal{G}_r \setminus \{e\}$ and returns U if it exists (Case (3)).

In Case (1) the existence of edges between vertices in T and vertices in $V \setminus T$ is explicitly checked. By Lemma 5 $\mathcal{G}[T]$ is a tSCC in \mathcal{G} . Since every 2eSCC is strongly connected, each 2eSCC of G is completely contained in either $G[T]$ or $G[V \setminus T]$.

In the Cases (2) and (3) we have by Corollary 6 that $\mathcal{G}[U]$ is new tSCC in $\mathcal{G} \setminus \{e\}$ with respect to e . Thus by Lemma 3 no vertex in $V \setminus U$ has two edge-disjoint paths to any vertex in U in \mathcal{G} . Hence each 2eSCC of G has to be completely contained in either $G[U]$ or $G[V \setminus U]$. The identified edge e has one endpoint in U and one in $V \setminus U$, i.e., there exists at least one edge between vertices in U and vertices in $V \setminus U$ in G . \square

Theorem 21 (Correctness). *Algorithm 2ESCC-SPARSE computes the 2eSCCs of the input graph.*

Proof. Algorithm 2ESCC-SPARSE repeatedly removes edges from the input graph until the SCCs in the remaining graph correspond to the 2eSCCs of the input graph. We will show the correctness of the algorithm by showing that (a) the removed edges cannot be in a 2eSCC, (b) when the algorithm terminates, each SCC is a 2eSCC, and (c) the algorithm terminates.

By definition, 2eSCCs are strongly connected subgraphs that do not contain a bridge. Thus (a) clearly holds when edges between SCCs or bridges are removed

in the repeat-until loop without the calls to `LocalSearch`. By Lemma 20.3, (a) also holds for the edges removed after `LocalSearch` returns a non-empty set of vertices.

To show (b), first note that whenever `LocalSearch` is called, there will be another iteration of the repeat-until loop. This is because the algorithm terminates only in the case that J is empty but `LocalSearch` is only called when there are vertices in J and `LocalSearch` does not remove vertices from J . Consider the last iteration of the repeat-until loop. In this iteration no bridges were identified as otherwise J cannot be empty. Thus no SCC in G contains a bridge, i.e., each SCC in G is a 2eSCC.

For (c) we will show that in each iteration of the inner and the outer-repeat until loop either edges are removed or the algorithm terminates. The inner repeat-until loop terminates when Procedure `LocalSearch` returns an empty set. Whenever Procedure `LocalSearch` returns a non-empty set S , by Lemma 20.3 there exist edges between S and $V \setminus S$ in G , which are then removed from G . For the outer repeat-until loop we have that the algorithm terminates if no bridges are identified in the for-loop; otherwise at least the bridges are removed from G . \square

5.2 Runtime

In the runtime analysis we need that the algorithm indeed starts local searches in each strongly connected subgraph with no or one incoming edge that was no such subgraph when the SCCs and bridges of the graph were determined for the last time. For this we will use the following observation.

Observation 22. *Let G be a directed graph. Let X be a set of edges in G and let J be the set of vertices adjacent to an edge in X in G . Let H be a subgraph of G . Every tSCC in $H \setminus X$ that has incoming edges in H contains a vertex of J .*

Proof. Let S be a set of vertices that induces a tSCC in $H \setminus X$ that has incoming edges in H . Then the set of edges X has to contain an incoming edge (u, v) for some vertex $v \in S$. We have that v is in J . \square

The next lemma applies the results of Section 3 to the subgraphs used in Algorithm 2ESCC-SPARSE to show that we can indeed find all desired subgraphs with at most d vertices in `LocalSearch`.

Lemma 23. *Let $V_{\mathcal{G},r}$ be the set of vertices that can reach a vertex r in $\mathcal{G} \in \{G, Rev(G)\}$ using a path containing at most d edges. Assume there exists a set of vertices S with $r \in S$ and $|S| \leq d$ that induces (a) a tSCC $\mathcal{G}[S]$ in \mathcal{G} or (b) a new tSCC $\mathcal{G}[S]$ in $\mathcal{G} \setminus \{e\}$ with respect to some edge e . Let $B_{\mathcal{G},r}$ be the vertices in $V_{\mathcal{G},r}$ that are missing incoming edges in \mathcal{G}_r compared to \mathcal{G} . Then we have $S \cap B_{\mathcal{G},r} = \emptyset$ and in Case (a) $\mathcal{G}[S]$ is a tSCC in \mathcal{G}_r and in Case (b) either the edge e is an edge-dominator in $\mathcal{G}'_r(b_{\mathcal{G},r})$ and $\mathcal{G}[S]$ is a new tSCC in $\mathcal{G}_r \setminus \{e\}$ with respect to e or there exists a tSCC $\mathcal{G}[T]$ in \mathcal{G}_r such that $T \cap B_{\mathcal{G},r} = \emptyset$.*

Proof. Since $|S| \leq d$ and $r \in S$, any simple path in $\mathcal{G}[S]$ from a vertex in $S \setminus \{r\}$ to r can contain at most $d - 1$ edges. Thus all vertices with incoming edges to vertices in S in \mathcal{G} can reach r using a path with at most d edges in \mathcal{G} . Hence S is a subset of $V_{\mathcal{G},r}$ and all incoming edges of S are contained in \mathcal{G}_r , i.e., $B_{\mathcal{G},r} \cap S = \emptyset$. By Lemma 5 this implies Case (a). For Case (b) we have that

either (b1) $B_{\mathcal{G},r} \neq \emptyset$ and the edge e (and thus S) is reachable from a vertex in $B_{\mathcal{G},r}$ in \mathcal{G} and thus by Corollary 9 the edge e is a dominator in $\mathcal{G}'_r(b_{\mathcal{G},r})$ and $\mathcal{G}[S]$ is a new tSCC in $\mathcal{G}_r \setminus \{e\}$ with respect to e , (b2) $B_{\mathcal{G},r} \neq \emptyset$ and the edge e (and thus S) is not reachable from $B_{\mathcal{G},r}$ in \mathcal{G} and thus there exists a set of vertices in $V_{\mathcal{G},r} \setminus B_{\mathcal{G},r}$ that is not reachable from $B_{\mathcal{G},r}$ in \mathcal{G} , or (b3) $B_{\mathcal{G},r} = \emptyset$ and thus no vertex in $V \setminus V_{\mathcal{G},r}$ can reach a vertex in $V_{\mathcal{G},r}$. In the Cases (b2) and (b3) there has to exist a tSCC $\mathcal{G}[T]$ in \mathcal{G}_r with $T \cap B_{\mathcal{G},r} = \emptyset$. \square

Theorem 24 (Runtime). *Algorithm 2ESCC-SPARSE can be implemented in time $O(m^2/\log n)$.*

Proof. Let the input graph have n vertices and m edges. Let G denote the graph maintained by the algorithm.

Using the $O(m)$ -time algorithms to compute SCCs and find all bridges in a graph G , an iteration of the outer repeat-until loop without the calls to `LocalSearch` takes time $O(m)$. We will show that there can be only $O(m/q + n/d) \in O(m/\log n)$ iterations of the outer repeat-until loop. We will bound the time spent in the inner repeat-until loop separately.

A new iteration of the outer repeat-until loop is started in two cases.

Case 1: $|J| \geq q$. A vertex is in J only when one of its adjacent edges was deleted from G since the last time J was initialized with the empty set. Thus Case 1 can happen at most $2m/q$ times.

Case 2: `LocalSearch` returned an empty set. Let G and J be as maintained by the algorithm at the beginning of the subsequent iteration of the outer repeat-until loop. We distinguish two subcases. Let a subgraph $G[W]$ induced by some set of vertices W be *connected* if for every partition of W into two subsets there are edges between the subsets. In Case (2a) there exists a subgraph in G that is connected but not strongly connected, in Case (2b) no such subgraph exists in G .

Case (2a): Let W be a set of vertices that induces a maximal connected but not strongly connected subgraph of G . Since $G[W]$ is connected, the vertices in W were strongly connected when the edges between SCCs were removed at the beginning of the previous iteration of the outer repeat-until loop. As $G[W]$ is not strongly connected, it contains a top and a bottom SCC that are disjoint. By the maximality of $G[W]$ this top and this bottom SCC in $G[W]$ are also a top and a bottom SCC in G , respectively. By Observation 22 each of them contains a vertex of J . Further, by Lemma 23 each of them has more than d vertices as otherwise at least one of them would have been identified by Procedure `LocalSearch` and thus they would not be connected in G . Both of them are identified when the SCCs of G are determined and at least the outgoing edges of the tSCC and the incoming edges of the bSCC are removed from G . Thus in this case two vertex sets that each contain more than d vertices are separated from each other by deleting all edges between them. This can happen at most n/d times.

Case (2b): After determining the SCCs of G , the algorithm searches for bridges in every SCC of G . If no bridge is found, the algorithm terminates. Assume that a bridge e was found in an SCC $G[W]$ induced by some set of vertices W . The edge e cannot have been a bridge when bridges were identified in the previous iteration of the outer repeat-until loop, as otherwise it would not be in G . Since e is a bridge, there exist a new top and a new bottom SCC with

respect to e in $G[W] \setminus \{e\}$. As every connected subgraph of G is also strongly connected, there are no edges between W and $V \setminus W$ in G . Thus the new top and the new bottom SCC in $G[W]$ are also a new top and a new bottom SCC in G . By Observation 22 each of them contains a vertex of J . By Lemma 23 each of them has more than d vertices as otherwise the edge e would have been identified by Procedure `LocalSearch` and removed from the graph. Thus in Case (2b) either the algorithm terminates or two vertex sets that each contain more than d vertices and were strongly connected to each other are separated from each other by deleting the edge e (such that they are no longer strongly connected but they might still be connected). This can happen at most n/d times.

It remains to bound the time spent in `LocalSearch`. To this end note that each time *before* `LocalSearch` is called either (a) `LocalSearch` was called and a set of vertices S that induces a top or bottom SCC $G[S]$ or a new top or bottom SCC $G[S]$ in $G \setminus \{e\}$ with respect to some edge e was identified and separated from the remaining graph by deleting the edges between S and $V \setminus S$ or (b) `Bridges` identified a bridge e and increased the number of SCCs in G by removing e . Both (a) and (b) can happen at most n times.

We now consider the time for one call to `LocalSearch`. For each $r \in J$ `LocalSearch` runs a breadth-first search of depth d on each of $\mathcal{G} \in \{G, Rev(G)\}$ to identify the subgraphs \mathcal{G}_r . Considering G and $Rev(G)$ only increases the running time by a factor of two. The number of edges explored by a breadth-first-search of depth d on a graph with out-degree at most three is $O(3^d)$. Thus with $d = \lceil \varepsilon \log n \rceil$ for some $0 < \varepsilon < 1$ we have that the number of edges in \mathcal{G}_r is $O(n^\varepsilon)$. `LocalSearch` computes SCCs and bridges or edge-dominators in \mathcal{G}_r . This can be done in time linear in the number of edges in G_r , i.e., in time $O(n^\varepsilon)$. Thus with $|J| < q$ we obtain a time bound of $O(q \cdot n^\varepsilon)$ for one call to `LocalSearch`. Hence the total time spent in `LocalSearch` can be bounded with $O(q \cdot n^\varepsilon \cdot n) = O(n^{1+\varepsilon} \log n)$. We have $O(n^{1+\varepsilon} \log n) \in O(m^2 / \log n)$ for any $\varepsilon \in (0, 1)$. \square

6 Extension to kSCCs

For a constant $k > 2$ the algorithm based on the hierarchical graph decomposition technique, presented in Section 4, extends to computing the k -edge and the k -vertex strongly connected components. In this section we outline the necessary changes, we will provide the details in a later version of this manuscript.

The main change is the extension of separators and dominators from a single element to sets of elements with size less than k . A k -separator is a minimal set of vertices such that the set contains less than k elements and its removal from the graph increases the number of SCCs in the graph. Two distinct vertices u and v are k -connected if they remain strongly connected after the removal of any less than k elements different from u and v from G . The k -strongly connected components (kSCCs) of a graph G are its maximal subgraphs $G[S]$ such that every pair of distinct vertices u and v in S is k -connected. This definition allows for degenerate kvSCCs with less than k vertices. Given the kvSCCs, the degenerate kvSCCs can be identified in linear time. A k -dominator Z in a flow graph $G(r)$ is a minimal set of less than k elements in $G(r) \setminus \{r\}$ such that there exists a vertex $u \in G(r) \setminus (\{r\} \cup Z)$ such that u is reachable from r and

every path from r to u contains an element of Z . We say that Z k -dominates u in $G(r)$. A subgraph $G[T]$ induced by some set of vertices T is a *new tSCC* in $G \setminus Z$ with respect to a set of elements Z if it is a tSCC in $G \setminus Z$ but has, for vertex-connectivity, incoming edges from each of the vertices in Z , or, for edge-connectivity, the edges in Z as incoming edges in G . We will only consider new tSCC with respect to sets Z with $|Z| < k$. With these definitions it is rather straightforward to extend Lemma 3 to k -connectivity and k -dominators:

Lemma 25 (Extension of Lemma 3 to k -connectivity). *Let Z be a set of elements with $|Z| < k$ such that there exists a new tSCC with respect to Z in $G \setminus Z$. Let T be the set of vertices in this new tSCC. Let $W = V \setminus T$ if the elements of Z are edges and let $W = V \setminus (T \cup Z)$ if the elements of Z are vertices. If $W \neq \emptyset$, then no vertex in W is k -connected to any vertex in T . Additionally, the set $Z \cap G(r)$ is a k -dominator in $G(r)$ for every $r \in W$ that can reach T in G .*

6.1 k -edge strongly connected components

For edge-connectivity the remaining parts extend in the same, straightforward way. Let $\mathcal{G} \in \{G, Rev(G)\}$. The same flow graphs $\mathcal{G}'_i(b_{\mathcal{G},i})$ as in Definition 8 can be used and the reachability condition is satisfied for a set of edges Z and the set of vertices $B_{\mathcal{G},i}$ if *every* edge in Z can be reached from some vertex in $B_{\mathcal{G},i}$. Corollary 9 still holds for new tSCCs with respect to sets Z with $|Z| < k$, k -dominators, and the adapted reachability condition.

To find k -dominators in a flow graph $\mathcal{G}'_i(b_{\mathcal{G},i})$ and k -separators in G , the $O(m \log n)$ -time algorithm by Gabow [9] can be used. Recall that in \mathcal{G}_i the number of edges is $O(n \cdot 2^i)$, i.e., the search in \mathcal{G}_i takes time $O(n \log n \cdot 2^i)$.

The correctness proof mainly depends on Lemmata 3 and 5 (i.e. the properties of tSCCs and new tSCCs in G and \mathcal{G}'_i) and thus can be extended in a rather straightforward way by using the generalized definition of new tSCCs. For the runtime note that Lemma 14 still holds for $|S| \leq 2^i - k + 2$. The lower bounds on the sizes of top and bottom SCCs and new top and bottom SCCs in Corollary 15, Lemma 16, and the runtime proof have to be changed to $2^i - k + 2$ accordingly. To take the additional $\log n$ -factor in the runtime into account, the following argument can be used: Whenever a set of vertices S that induces a top or bottom SCC or a new top or bottom SCC $G[S]$ is identified and the algorithm recurses on $G[S]$ and $G[V \setminus S]$, the algorithm spends time proportional to $n \log n \cdot \min(|S|, |V \setminus S|)$. We charge $O(n \log n)$ time to each vertex in the smaller set of S and $V \setminus S$. A vertex can be in the smaller set at most $O(\log n)$ times. Thus the total runtime is bounded by $O(n^2 \log^2 n)$.

6.2 k -vertex strongly connected components

The extension to $k > 2$ is more complicated for vertex-connectivity. We outline here the changes needed *in addition* to the respective changes for edge-connectivity. In particular, we have to deal with the case $0 < |B_{\mathcal{G},i}| < k$. Note that in this case we cannot use an additional vertex that we connect to the vertices in $B_{\mathcal{G},i}$ as root in the flow graph because the vertices in $B_{\mathcal{G},i}$ would be a k -dominator in this flow graph independent of the underlying graph G . Assume we are in this case and there exists a k -separator Z in G such that the new tSCC $\mathcal{G}[T]$ with respect to Z is induced by some set of vertices T that is contained in $V \setminus B_{\mathcal{G},i}$. Assume for simplicity that G is strongly connected.

Then we want to be able to identify Z and T in \mathcal{G}_i (similar to Corollary 9 and the special case for $B_{\mathcal{G},i} = 1$ for 2vSCCs). We distinguish two cases: Either there exists a vertex in $B_{\mathcal{G},i}$ that is not in Z or $B_{\mathcal{G},i}$ is contained in Z . For the first case we will use $|B_{\mathcal{G},i}| < k$ different flow graphs (which we will define below) and extend the reachability condition such that it characterizes which new tSCCs we can identify by searching for k -dominators in these flow graphs. If the search in the $|B_{\mathcal{G},i}|$ flow graphs is not successful, we consider the second case by searching for a $(k - |B_{\mathcal{G},i}|)$ -separator in $\mathcal{G}_i \setminus B_{\mathcal{G},i}$. If such a separator Z' exists, then $Z = Z' \cup B_{\mathcal{G},i}$ contains less than k vertices and there is a new tSCC induced by some set of vertices $T \subsetneq V \setminus Z$ in $\mathcal{G}_i \setminus Z$. The correctness of this approach will follow from a generalization of Lemma 5 and Corollary 6 described below, and Lemma 25; for the runtime analysis this second case corresponds to a generalization of Lemma 14.3.

We now define the flow graphs that we use in the algorithm for kvSCCs. If $|B_{\mathcal{G},i}| \geq k$, the same flow graph $\hat{\mathcal{G}}_i(s_{\mathcal{G},i})$ as for 2vSCCs (Definition 8) and $|B_{\mathcal{G},i}| \geq 2$ can be used. If $|B_{\mathcal{G},i}| < k$, let $\hat{\mathcal{G}}_{i,w}(s_{\mathcal{G},i,w})$ denote the flow graph for $w \in B_{\mathcal{G},i}$. Then the root $s_{\mathcal{G},i,w}$ is equal to w and the graph $\hat{\mathcal{G}}_{i,w}$ is the graph \mathcal{G}_i with additional edges from w to each vertex in $B_{\mathcal{G},i} \setminus \{w\}$. Lemmata 5 and 7 and Corollary 6 still hold for these graphs as the incoming edges of vertices in $A_{\mathcal{G},i}$ are the same in \mathcal{G} and $\hat{\mathcal{G}}_{i,w}$.

The reachability condition can be extended as follows. Let the reachability condition be satisfied for a set of vertices Z and the set of vertices $B_{\mathcal{G},i}$ if *every* vertex in Z is a vertex in $B_{\mathcal{G},i}$ or can be reached from a vertex in $B_{\mathcal{G},i}$ and either $|B_{\mathcal{G},i}| \geq k$ or there exists a vertex in $B_{\mathcal{G},i} \setminus Z$. With this generalization of the reachability condition, Corollary 9 can be extended to k -vertex connectivity, where in the case of $|B_{\mathcal{G},i}| < k$ a set Z has to be a k -dominator in *one of* the flow graphs $\hat{\mathcal{G}}_{i,w}(s_{\mathcal{G},i,w})$.

To find k -dominators in the flow graphs and k -separators in G and \mathcal{G}_i , the $O(mn)$ -time algorithms of Henzinger et al. [18] can be used (see Even [7] and Even and Tarjan [8] for the first and Gabow [10] for the newest related algorithms). For the search in \mathcal{G}_i this gives a runtime of $O(n^2 \cdot 2^i)$.

The correctness proof can be extended to kvSCCs in a similar way as for edge-connectivity. For the runtime proof we additionally have to take into account the increased runtime in each recursive step of the algorithm and that in the recursion on $G_{S,Z}$ and $G_{V \setminus S}$ the set Z might contain up to $k - 1$ vertices. For this we define $n_v = \min(|S|, |V \setminus S| - k)$. Note that $n_v \leq n'/2$. We stop the recursion when the number of vertices n' is less than $14k^3$, i.e., a constant; in this case we can use the known $O(mn^2)$ -time algorithm for kvSCCs to compute the kvSCCs of G in constant time. Recall that it takes time $O(n^2 \cdot n_v)$ to identify S and thus we can bound the runtime without recursion with $c \cdot n'^2 \cdot n_v$ for some constant c . With this we obtain an upper bound of $f(n) = cn^3$ on the total runtime of the algorithm for kvSCCs as follows:

$$\begin{aligned}
f(n') &\leq f(n_v + k) + f(n' - n_v) + cn'^2 n_v, \\
&= c(n_v + k)^3 + c(n' - n_v)^3 + cn'^2 n_v, \\
&= cn'^3 - 2cn'^2 n_v + 3cn' n_v^2 + 3cn_v^2 k + 3cn_v k^2 + ck^3, \\
&\leq cn'^3 - 2cn'^2 n_v + 3cn' n_v^2 + 7ck^3 n_v^2, \\
&\leq cn'^3.
\end{aligned}$$

Acknowledgements

We would like to thank Giuseppe Italiano for suggesting the problem and Slobodan Mitrović for helpful discussions. V. L. would like to thank Christian Tschabuschnig for his help in improving the readability of the algorithms.

This work was funded by the Austrian Science Fund (FWF): P23499-N23. Additionally, the research leading to these results has received funding from the European Research Council under the European Union’s Seventh Framework Programme (FP/2007-2013) / ERC Grant Agreement no. 340506.

References

- [1] Jørgen Bang-Jensen and Gregory Gutin. *Digraphs: Theory, algorithms and applications*. Springer Monographs in Mathematics. Springer, London, 2nd edition, 2009.
- [2] John A. Bondy and U. S. R. Murty. *Graph Theory with Applications*. Macmillan, London, 1976.
- [3] Adam L. Buchsbaum, Loukas Georgiadis, Haim Kaplan, Anne Rogers, Robert E. Tarjan, and Jeffery R. Westbrook. Linear-time algorithms for dominators and other path-evaluation problems. *SIAM Journal on Computing*, 38(4):1533–1573, 2008.
- [4] Krishnendu Chatterjee and Monika Henzinger. Efficient and Dynamic Algorithms for Alternating Büchi Games and Maximal End-component Decomposition. *Journal of the ACM*, 61(3):15, 2014.
- [5] Krishnendu Chatterjee, Monika Henzinger, and Veronika Loitzenbauer. Improved Algorithms for One-Pair and k -Pair Streett Objectives. <http://arxiv.org/abs/1410.0833v2>, October 2014.
- [6] Krishnendu Chatterjee, Marcin Jurdziński, and Thomas A. Henzinger. Simple stochastic parity games. In *Computer Science Logic*, pages 100–113. Springer, 2003.
- [7] Shimon Even. An Algorithm for Determining Whether the Connectivity of a Graph is at Least k . *SIAM Journal on Computing*, 4(3):393–396, 1975.
- [8] Shimon Even and Robert E. Tarjan. Network Flow and Testing Graph Connectivity. *SIAM Journal on Computing*, 4(4):507–518, 1975.
- [9] Harold N. Gabow. A matroid approach to finding edge connectivity and packing arborescences. *Journal of Computer and System Sciences*, 50(2):259–273, 1995.
- [10] Harold N. Gabow. Using expander graphs to find vertex connectivity. *Journal of the ACM*, 53(5):800–844, 2006.
- [11] Harold N. Gabow and Robert E. Tarjan. A linear-time algorithm for a special case of disjoint set union. In *STOC*, pages 246–251, 1983.
- [12] Zvi Galil and Giuseppe F. Italiano. Reducing edge connectivity to vertex connectivity. *ACM SIGACT News*, 22(1):57–61, 1991.

- [13] Loukas Georgiadis. Testing 2-vertex connectivity and computing pairs of vertex-disjoint s-t paths in digraphs. In *ICALP*, pages 433–442, 2010.
- [14] Loukas Georgiadis, Giuseppe F. Italiano, Luigi Laura, and Nikos Parotsidis. 2-Edge Connectivity in Directed Graphs. <http://arxiv.org/abs/1407.3041v2>, to appear in SODA 2015, July 2014.
- [15] Loukas Georgiadis, Giuseppe F. Italiano, Luigi Laura, and Nikos Parotsidis. 2-Vertex Connectivity in Directed Graphs. <http://arxiv.org/abs/1409.6277>, September 2014.
- [16] Loukas Georgiadis and Robert E. Tarjan. Finding dominators revisited. In *SODA*, pages 862–871, 2004.
- [17] Monika Henzinger, Valerie King, and Tandy Warnow. Constructing a Tree from Homeomorphic Subtrees, with Applications to Computational Evolutionary Biology. *Algorithmica*, 24(1):1–13, 1999.
- [18] Monika R. Henzinger, Satish Rao, and Harold N. Gabow. Computing Vertex Connectivity: New Bounds from Old Techniques. *Journal of Algorithms*, 34(2):222–250, 2000.
- [19] John E. Hopcroft and Robert Endre Tarjan. Dividing a graph into triconnected components. *SIAM Journal on Computing*, 2(3):135–158, 1973.
- [20] Giuseppe F. Italiano, Luigi Laura, and Federico Santaroni. Finding strong bridges and strong articulation points in linear time. *Theoretical Computer Science*, 447:74–84, 2012.
- [21] Raed Jaber. On computing the 2-vertex-connected components of directed graphs. <http://arxiv.org/abs/1401.6000v1>, January 2014.
- [22] Seishi Makino. An algorithm for finding all the k -components of a digraph. *International Journal of Computer Mathematics*, 24(3–4):213–221, 1988.
- [23] Hiroshi Nagamochi and Toshihide Ibaraki. *Algorithmic aspects of graph connectivity*. Cambridge University Press, New York, 2008.
- [24] Hiroshi Nagamochi and Toshimasa Watanabe. Computing k -edge-connected components of a multigraph. *IEICE TRANSACTIONS on Fundamentals of Electronics, Communications and Computer Sciences*, E76-A(4):513–517, 1993.
- [25] Mark E. J. Newman. *Networks: An Introduction*. Oxford University Press, Oxford, 2010.
- [26] Robert E. Tarjan. Depth-first search and linear graph algorithms. *SIAM Journal of Computing*, 1(2):146–160, 1972.
- [27] Robert E. Tarjan. Edge-disjoint spanning trees and depth-first search. *Acta Informatica*, 6(2):171–185, 1976.