From Pathwidth to Connected Pathwidth
Dariusz Dereniowski

To cite this version:
Dariusz Dereniowski. From Pathwidth to Connected Pathwidth. Symposium on Theoretical Aspects of Computer Science (STACS2011), Mar 2011, Dortmund, Germany. 9, pp.416-427, 0000. <hal-00573616>

HAL Id: hal-00573616
https://hal.archives-ouvertes.fr/hal-00573616
Submitted on 5 Mar 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
From Pathwidth to Connected Pathwidth

Dariusz Dereniowski

1 Department of Algorithms and System Modeling, Gdańsk University of Technology, Narutowicza 11/12, 80-233 Gdańsk, Poland
deren@eti.pg.gda.pl

Abstract

It is proven that the connected pathwidth of any graph \( G \) is at most \( 2 \cdot \text{pw}(G) + 1 \), where \( \text{pw}(G) \) is the pathwidth of \( G \). The method is constructive, i.e. it yields an efficient algorithm that for a given path decomposition of width \( k \) computes a connected path decomposition of width at most \( 2k + 1 \). The running time of the algorithm is \( O(dk^2) \), where \( d \) is the number of ‘bags’ in the input path decomposition.

The motivation for studying connected path decompositions comes from the connection between the pathwidth and some graph searching games. One of the advantages of the above bound for connected pathwidth is an inequality \( \text{cs}(G) \leq 2\sigma(G) + 3 \), where \( \text{cs}(G) \) is the connected search number of a graph \( G \) and \( \sigma(G) \) is its search number, which holds for any graph \( G \). Moreover, the algorithm presented in this work can be used to convert efficiently a given search strategy using \( k \) searchers into a connected one using \( 2k + 3 \) searchers and starting at arbitrary homebase.

1998 ACM Subject Classification G.2.2 Graph Theory, F.2.2 Nonnumerical Algorithms and Problems

Keywords and phrases connected pathwidth, connected searching, fugitive search games, graph searching, pathwidth

Digital Object Identifier 10.4230/LIPIcs.STACS.2011.416

1 Introduction

The notions of pathwidth and treewidth are receiving increasing interest since the series of Graph Minor articles by Robertson and Seymour. The importance of those parameters is due to their numerous practical applications, connections with several graph parameters and usefulness in designing graph algorithms. Informally speaking, the pathwidth of a graph \( G \), denoted by \( \text{pw}(G) \), says how closely \( G \) is related to a path. Moreover, a path decomposition captures the linear path-like structure of \( G \). (For a definition see Section 2.)

Here we briefly describe a graph searching game that is the main motivation for the results presented in this paper. A team of \( k \) searchers is given and the goal is to find an invisible and fast fugitive located in a given graph \( G \). The fugitive has also the complete knowledge about the graph and about the strategy used by the searchers, and therefore he will avoid being captured as long as possible. The fugitive is captured when a searcher reaches his location. In this setting the game is equivalent to the problem of clearing all edges of a graph that is initially entirely contaminated. There are two main types of this graph searching problem. In the node searching two moves are allowed: placing a searcher on a vertex and removing a searcher from a vertex. An edge becomes clear whenever both of its
endpoints are simultaneously occupied by searchers. In the edge searching we have, besides to the two mentioned moves, a move of sliding a searcher along an edge. In this model an edge \( \{u, v\} \) becomes clear if a searcher slides from \( u \) to \( v \) and either all the other edges incident to \( u \) have been previously cleared or another searcher occupies \( u \). In both cases the goal is to find a search strategy (a sequence of moves of the searchers) that clears all the edges of \( G \). The node (edge) search number of \( G \), denoted by \( ns(G) \) (\( s(G) \), respectively), equals the minimum number of searchers sufficient to construct a node (edge, respectively) search strategy. An important property is that \( pw(G) = ns(G) - 1 \) for any graph \( G \) \([13, 14, 15, 18]\). The edge searching problem is closely related to node searching, i.e. \( |s(G) - ns(G)| \leq 1 \) \([4]\), and consequently to pathwidth, \( pw(G) \leq s(G) \leq pw(G) + 2 \).

In this work we are interested in special types of path decompositions called connected path decompositions. The motivation comes from the need of creating connected search strategies. An edge search strategy is connected if the subgraph of \( G \) that is clear is always connected. The minimum number of searchers that guarantees the capture of the fugitive by a connected (edge) search strategy, denoted by \( cs(G) \), is the connected search number of \( G \). This model of graph searching receives recently a growing interest, because in many applications the connectedness is an requirement.

Related work. There are several results that give a relation between the connected and the ‘classical’ search numbers of a graph. Fomin et al. proved in \([7]\) that the connected search number of an \( n \)-node graph of branchwidth \( b \) is bounded by \( O(b \log n) \) and this bound is tight. One of the implications of this result is that \( cs(G) = O(\log n)pw(G) \). Nisse proved in \([19]\) that \( cs(G) \leq (\text{tw}(G) + 2)(2s(G) - 1) \) for any chordal graph \( G \). Barrière et al. obtained in \([2]\) a constant upper bound for trees, namely for each tree \( T \), \( cs(T)/s(T) \leq 2 \). On the other hand, there exists an infinite family of graphs \( G_k \) such that \( cs(G_k)/s(G_k) \) approaches \( 2 \) when \( k \) goes to infinity \([3]\).

Fraigniaud and Nisse presented in \([9]\) a \( O(nk^2) \)-time algorithm that takes a width \( k \) tree decomposition of a graph and returns a connected tree decomposition of the same width. (For definition of treewidth see e.g. \([5, 21]\).) Therefore, \(tw(G) = ctw(G)\) for any graph \( G \). This method cannot be applied for proving the same result for connected path decompositions, because the decomposition that the algorithm in \([9]\) constructs is not, in general, a path decomposition even when a path decomposition is given as an input. That result also yields an upper bound of \( cs(G) \leq (\log n + 1)s(G) \) for any graph \( G \). The problems of computing the pathwidth (the search number) and the connected pathwidth (the connected search number) are NP-hard, also for several special classes of graphs, see e.g. \([6, 11, 12, 16, 17, 20]\).

This work. This paper presents an efficient algorithm that takes a (connected) graph \( G \) and its path decomposition \( \mathcal{P} = (X_1, \ldots, X_d) \) of width \( k \) as an input and finds in time \( O(dk^2) \) a connected path decomposition \( \mathcal{C} = (Z_1, \ldots, Z_m) \) of width at most \( 2k + 1 \), where \( m \leq kd \). This solves an open problem stated in several papers, e.g. in \([1, 2, 3, 7, 8, 9, 10, 22]\), since it implies that for any graph \( G \), \( \text{cpw}(G) \leq 2\text{pw}(G) + 1 \), and improves previously known estimations \([7, 19]\). The path decomposition \( \mathcal{C} \) can be turned into a monotone connected search strategy using at most \( 2k + 3 \) searchers. Thus, in terms of the graph searching terminology, the bound immediately implies that \( mcs(G) \leq \text{cpw}(G) + 2 \leq 2\text{pw}(G) + 3 \leq 2s(G) + 3 \), where \( mcs(G) \) is the monotone connected search number of \( G \). (A search strategy is monotone if the fugitive cannot reach a previously cleared edge.) Since \( cs(G) \leq mcs(G) \), the bound can be restated for the connected search number of a graph, \( cs(G) \leq 2s(G) + 3 \). Moreover, the factor 2 in the bound is tight \([3]\). The bound finds also applications in designing approximation algorithms, for it implies that the pathwidth and the connected pathwidth (the search
number, the connected search number, and some other search numbers not mentioned here, e.g. the internal search number) are within a constant factor of each other.

2 Preliminaries and basic definitions

Given a simple graph $G = (V(G), E(G))$ and its subset of vertices $X \subseteq V(G)$, the subgraph of $G$ induced by $X$ is $G[X] = (X, \{u, v \in E(G) : u, v \in X\})$. For a simple (not necessarily connected) graph $G$, $H$ is its connected component if $H$ is connected, that is, there exists a path in $H$ between each pair of vertices, and each proper supergraph of $H$ is not a subgraph of $G$. For $X \subseteq V(G)$ let $N_G(X) = \{x \in V(G) \setminus X : \{x, x' \in E(G)\mbox{ for some } x \in X\}$.

Definition 1. A path decomposition of a simple graph $G = (V(G), E(G))$ is a sequence $\mathcal{P} = (X_1, \ldots, X_d)$, where $X_i \subseteq V(G)$ for each $i = 1, \ldots, d$, and

- $\bigcup_{i=1}^{d} X_i = V(G)$,
- for each $\{u, v\} \in E(G)$ there exists $i \in \{1, \ldots, d\}$ such that $u, v \in X_i$,
- for each $i, j, k, 1 \leq i \leq j \leq k \leq d$ it holds $X_i \cap X_k \subseteq X_j$.

The width of the path decomposition $\mathcal{P}$ is $\text{width}(\mathcal{P}) = \max_{i=1}^{d} |X_i| - 1$. The pathwidth of $G$, $\text{pw}(G)$, is the minimum width over all path decompositions of $G$.

A path decomposition $\mathcal{P}$ is connected if $G[X_1 \cup \cdots \cup X_i]$ is connected for each $i = 1, \ldots, d$. Then, $\text{cpw}(G)$ denotes the minimum width over all connected path decompositions of $G$.

Definition 2. Given a graph $G$ and its path decomposition $\mathcal{P} = (X_1, \ldots, X_d)$, a node-weighted graph $\mathcal{G} = (V(\mathcal{G}), E(\mathcal{G}), \omega)$ derived from $G$ and $\mathcal{P}$ is the graph with vertex set $V(\mathcal{G}) = V_1 \cup \cdots \cup V_d$, where $V_i = \{v_i(H) : H$ is a connected component of $G[X_i]\}$, $i = 1, \ldots, d$, and edge set $E(\mathcal{G}) = \{(v_i(H), v_{i+1}(H')) : v_i(H) \in V_i, v_{i+1}(H') \in V_{i+1}, i \in \{1, \ldots, d-1\}$, and $V(H) \cap V(H') \neq \emptyset\}$. The weight of a vertex $v_i(H) \in V(\mathcal{G})$, $i \in \{1, \ldots, d\}$, is $\omega(v_i(H)) = |V(H)|$. The width of $\mathcal{G}$, denoted by $\text{width}(\mathcal{G})$, equals $\text{width}(\mathcal{P}) + 1$.

In the following we omit a subgraph $H$ of $G$ and the index $i \in \{1, \ldots, d\}$ whenever they are not important when referring to a vertex of $\mathcal{G}$ and we write $v$ instead of $v_i(H)$. For brevity, $\omega(X) = \sum_{x \in X} \omega(x)$ for any subset $X \subseteq V(\mathcal{G})$.

Figures 1(a) and 1(b) present a graph $G$ and its path decomposition $\mathcal{P}$, respectively, where the subgraph structure in each bag $X_i$ is also given. Figure 1(c) depicts the derived graph $\mathcal{G}$. Note that $\mathcal{P}$ is not connected: the subgraphs $G[X_1 \cup \cdots \cup X_i]$ are not connected for $i = 2, 3, 4$. Let $C \subseteq V(\mathcal{G})$. The border $\delta(C)$ of the set $C$ is its subset consisting of all the vertices $v \in C$ such that there exists $u \in V(\mathcal{G}) \setminus C$ adjacent to $v$ in $\mathcal{G}$, i.e. $\delta(C) = N_{\mathcal{G}}(V(\mathcal{G}) \setminus C)$.

Given a set $X \subseteq V(\mathcal{G})$, $X \neq \emptyset$, we define the left (right) extremity of $X$ as $l(X) = \min \{i : V_i \cap X \neq \emptyset\}$ (resp. $r(X) = \max \{i : V_i \cap X \neq \emptyset\}$), respectively.

A path $P$ in $\mathcal{G}$ is progressive if $|V(P) \cap V_i| \leq 1$ for each $i = 1, \ldots, d$.

Definition 3. Given $\mathcal{G}$, $C \subseteq V(\mathcal{G})$ and $X \subseteq \delta(C)$, a left (right) branch $B_l(C, X, i)$, where $1 \leq i \leq r(X)$ (respectively $B_r(C, X, i)$, where $l(X) \leq i \leq d$) is the subgraph of $\mathcal{G}$ induced by the vertices in $X$ and by the vertices of all progressive paths contained in $(V(\mathcal{G}) \setminus C) \cup X$ and connecting $x \in X \cap V_j$ and $v \in V_k \setminus C$, where $i \leq k \leq j$ ($j \leq k \leq i$, respectively).

We sometimes write $B$ to refer to a branch whenever its ‘direction’ or $C, X, i$ are clear from the context. A branch $B = B_l(C, X, i)$, $i \leq r(X)$, (resp. $B = B_r(C, X, i)$, $i \geq l(X)$) is continuous if $V_j \cap V(B) \neq \emptyset$ for each $j = i, \ldots, r(X)$ (resp. $j = l(X), \ldots, i$), respectively. A vertex $v$ of $B$ is external if $N_{\mathcal{G}}(v) \not\subseteq C \cup V(B)$. The branch $B$ is proper if it has no external vertices in
This flexibility is provided due to the potential application of this algorithm: in graph searching games the bags of path decompositions correspond to the vertices occupied by the searchers while the search proceeds; in this way the selected vertex \( v \) can be the first one that becomes guarded in a connected search strategy and it is called the *homebase*. The first step performed by \( CP \) is the construction of the derived graph \( \mathcal{G} \) and in the subsequent

**Figure 1** (a) a graph \( G \); (b) a path decomposition \( \mathcal{P} \) of \( G \); (c) the graph \( \mathcal{G} \) derived from \( G \) and \( \mathcal{P} \)

\[ V_{i+1} \cup \cdots \cup V_{i+1} \cup \cdots \cup V_{i-1}, \text{respectively}, \] while \( \mathcal{B} \) is *maximal* if it is continuous, proper and \( B_{i+1}(C, X, i-1) \) is not a proper branch. An integer \( j \) is a *cut* of the branch \( \mathcal{B} \) if \( i \leq j \leq r(X) \) (\( l(X) \leq j \leq i \), respectively). The *weight* of the cut \( j \) of \( \mathcal{B} \) is \( \omega((V(\mathcal{B}) \cap V_j) \cup (X \cap Y)) \), where \( Y = V_{i(X)} \cup \cdots \cup V_{j-1} \) for a left branch, and \( Y = V_{i+1} \cup \cdots \cup V_{i(X)} \) for a right branch. A cut of minimum weight is a *bottleneck* of \( \mathcal{B} \).

Figure 2 illustrates the above definitions. (In all cases the branch is distinguished by the dark area.) Let \( X = \{x_1, x_2, x_3\} \) be a subset of \( \delta(C) \). Figure 2(a) gives \( B_{i+1}(C, X, i_2) \) and this branch is continuous and proper, but not maximal for each \( i_2, i_3 < i_2 \leq i_1 \). The branch \( B_{i+1},(C, X, i_3) \) (see Figure 2(b)) is maximal (thus continuous and proper), which follows from the fact that any branch \( B_{i+(C, X, i_4)} \) where \( i_4 < i_3 \) is not proper, because it contains an external vertex \( u_2 \), as shown in Figure 2(c). Note that the vertices of \( \mathcal{G}_1 \) and \( \mathcal{G}_2 \) (except for \( u_1 \) and \( u_2 \)) do not belong to any branch \( B_{i+1}(C, X, i) \), because they are not connected by progressive paths to \( x_2 \) or \( x_3 \). Figure 2(d) depicts a branch \( B_{i+(C, X, i_\tau)} \) that is not continuous for any \( \tau < i_5 \), because \( V_{i_4} \cap V(\mathcal{B}_{i+(C, X, i_\tau)}) = \emptyset \). In our algorithm we ensure that each branch we use is continuous and proper.

Given \( C, X \) and \( i \), a left (respectively right) branch \( \mathcal{B} \) can be calculated efficiently as follows. Initially \( \mathcal{B} \) satisfies \( V(\mathcal{B}) = X \). We start with \( j = r(X) \) (\( j = l(X) \), resp.) and we add to the vertex set of the branch all vertices in \( V_{j-1} \setminus C \) (\( V_{j+1} \setminus C \), resp.) that have a neighbor in \( V_j \cap V(\mathcal{B}) \). Then, we decrement (increment, resp.) \( j \) and repeat this step. The computation stops when \( j < i \) (\( j > i \), respectively).

### 3 The algorithm

The algorithm \( CP \) (*Connected Pathwidth*) for finding a connected path decomposition of a graph \( G \) takes \( G \), a vertex \( v \) of \( G \), and a path decomposition \( \mathcal{P} \) of \( G \) as an input. The vertex \( v \) is guaranteed to belong to the first bag of the resulting connected path decomposition. This flexibility is provided due to the potential application of this algorithm: in graph searching games the bags of path decompositions correspond to the vertices occupied by the searchers while the search proceeds; in this way the selected vertex \( v \) can be the first one that becomes guarded in a connected search strategy and it is called the *homebase*. The first step performed by \( CP \) is the construction of the derived graph \( \mathcal{G} \) and in the subsequent
steps the algorithm works on $G$. (Also, most parts of our analysis use $G$ rather than $G$.) The algorithm computes a sequence of sets $C_j \subseteq V(G)$, $j = 1, \ldots, m$, called expansions. The expansion $C_1$ consists of $v$ and one of its neighbors, and $C_m = V(G)$ at the end of the execution of CP. Moreover, $C_j \subseteq C_{j+1}$ for each $j = 1, \ldots, m - 1$. Informally speaking, $C_{j+1}$ is obtained from $C_j$ by adding to $C_j$ some vertices from $N_G(C_j)$. This guarantees that the final path decomposition obtained from $\delta(C_1), \ldots, \delta(C_m)$ is valid and is connected, as proved in Lemma 7. On the other hand, the particular vertices in $N_G(C_j)$, used to obtain $C_{j+1}$, are selected in a way to guarantee that $\omega(\delta(C_j))$ is bounded by $2 \cdot \text{width}(G)$ for each $j = 1, \ldots, m$. By construction, $\omega(\delta(C_j))$ is the size of the corresponding $j$th bag in the resulting connected path decomposition.

In this section we give the statement of the algorithm and we prove that it computes a connected path decomposition $C$. Then, in Section 4 we analyze the width of $C$. Due to the space limitations, the proofs of several results (marked with $\triangle$) are omitted.

The algorithm computes for each expansion $C_j$ two sets called the left and right borders of $C_j$, denoted by $\delta_L(C_j)$ and $\delta_R(C_j)$, respectively. It is guaranteed that $\delta_L(C_j) \cup \delta_R(C_j) = \delta(C_j)$ for each $j = 1, \ldots, m$ (see Lemma 6). As it is proven later, the left and right borders are special types of partitions of $\delta(C_j)$. In particular, there exists an integer $i \in \{1, \ldots, d\}$ such that the left border $\delta_L(C_j)$ is contained in $V_1 \cup \cdots V_i$ and the right border $\delta_R(C_j)$ is a subset of $V_{i+1} \cup \cdots \cup V_d$. For brevity let in the following $l(\delta_L(C_j)) = r(\delta_L(C_j)) = 0$ if $\delta_L(C_j) = \emptyset$ and $l(\delta_R(C_j)) = r(\delta_R(C_j)) = d$ if $\delta_R(C_j) = \emptyset$, where $C_j$ is any expansion.

We start by describing a subroutine EE (Extend Expansion) that is used by the main procedure CP given below. The input to EE consists of two integers $i$ and $k$, $i, k \in \{1, \ldots, d\}$. Informally speaking, the procedure adds, in its subsequent iterations, to the current expansion $C_m$ each vertex in $V_j$ that is connected by a progressive path to a vertex in $V_j \cap \delta(C_m)$ for each $j = i$ to $k$ and for some $i', i \leq i' \leq j$ if $i < k$, and for each $j = i$ down to $k$, $j \leq i' \leq i$ if $i > k$, which we formally prove in Lemma 4 below. All the ‘intermediate’ expansions are recorded as they will give us the corresponding bags in the final path decomposition. The procedures EE$(i, k)$ and CP$(G, P)$ are as follows.
**Procedure** EE (Extend Expansion)

**Input:** integers $i$ and $k$. ($G$, $m$, $C_m$ are used as global variables)

while $k \neq i$ do

if $k < i$ then

**EL:** Increment $m$, decrement $i$ and set:

$C_m = C_{m-1} \cup (V_i \cap N_G(C_{m-1}))$,

$\delta_L(C_m) = (\delta_L(C_{m-1}) \cup V_i) \cap \delta(C_m)$,

$\delta_R(C_m) = \delta_R(C_{m-1}) \cap \delta(C_m)$.

else ($k > i$)

**ER:** Increment $m$, increment $i$ and set:

$C_m = C_{m-1} \cup (V_i \cap N_G(C_{m-1}))$,

$\delta_R(C_m) = (\delta_R(C_{m-1}) \cup V_i) \cap \delta(C_m)$,

$\delta_L(C_m) = \delta_L(C_{m-1}) \cap \delta(C_m)$.

end if

end while.

end procedure EE.

**Algorithm** CP (Connected Pathwidth)

**Input:** a simple graph $G$, a path decomposition $P$ of $G$, and a vertex $v \in V(G)$.

**Output:** a connected path decomposition $C$ of $G$.

(Initialization.)

I.1: Use $G$ and $P$ to calculate the derived graph $G$. Let $v$ be any vertex of $G$. Let $C_1 = \{x,y\}$, where $v \in C_1$, $x,y$ are adjacent in $G$, and $x \in V_i$, $y \in V_{i+1}$ for some $i \in \{1, \ldots, d-1\}$. Let $m = 1$.

I.2: If $x \in \delta(C_1)$, then set $\delta_L(C_1) = \{x\}$, compute the maximal left branch $B_L(C_1, \delta_L(C_1), a_0)$ with a bottleneck $a_0' (a_0' \geq a_0)$ and with no external vertices in $V_i$ and call EE$(i, a_0')$; otherwise $\delta_L(C_1) = \emptyset$.

I.3: If $y \in \delta(C_1)$, then set $\delta_R(C_1) = \{y\}$, compute the maximal right branch $B_R(C_1, \delta_R(C_1), b_0)$ with a bottleneck $b_0' (b_0' \leq b_0)$ and with no external vertices in $V_i$ and call EE$(i+1, b_0')$; otherwise $\delta_R(C_1) = \emptyset$.

(Main loop.)

while $C_m \neq V(G)$ do

if $\omega(\delta_L(C_m)) > \omega(\delta_R(C_m))$ then

L.1: Compute the maximal left branch $B_1 = B_L(C_m, \delta_L(C_m), k_1)$. If $B_1$ has no external vertex in $V_i$, $i = r(\delta_L(C_m))$, then call EE$(r(\delta_L(C_m)), k_1)$, otherwise let $k_1 = r(\delta_L(C_m))$.

L.2: Compute the maximal right branch $B_2 = B_R(C_m, \delta_R(C_m) \cup (V_{i} \cap \delta_L(C_m)), k_2)$. Let $k_2'$ be its minimum weight cut such that $k_2' > k_1$. Call EE$(k_1, k_2')$.

L.3: If $r(\delta_L(C_m)) = k_1$, then compute the maximal left branch $B_3 = B_L(C_m, \delta_L(C_m), k_3)$ with bottleneck $k_3'$ and call EE$(k_1, k_3')$.

else ($\omega(\delta_L(C_m)) \leq \omega(\delta_R(C_m))$)

R.1: Compute the maximal right branch $B_1 = B_R(C_m, \delta_R(C_m), k_1)$. If $B_1$ has no external vertex in $V_i$, $i = l(\delta_R(C_m))$, then call EE$(l(\delta_R(C_m)), k_1)$, otherwise let $k_1 = l(\delta_R(C_m))$.

end if

end while.
Lemma 4.

Lemma 5.

4.2 From Pathwidth to Connected Pathwidth

R.2: Compute the maximal left branch $B_2 = B_L(C_m, \delta_L(C_m)) \cup (V_{k_1} \cap \delta_R(C_m))$. Let $k_2'$ be its minimum weight cut such that $k_2' < k_1$. Call $EE(k_1, k_2')$.

R.3: If $l(\delta_R(C_m)) = k_1$, then compute the maximal right branch $B_3 = B_R(C_m, \delta_R(C_m), k_3)$ with bottleneck $k_3'$ and call $EE(k_1, k_3')$.

end if

end while.

Let $Z_j = \bigcup_{v \in H \in \delta(C_j)} V(H)$ for each $j = 1, \ldots, m$. Return $C = (Z_1, \ldots, Z_m)$.

end procedure $CP$.

First we briefly discuss the initialization stage of $CP$. In Step 1, an expansion $C_1$ is constructed in such a way that it contains any two adjacent vertices (the adjacency guarantees the connectedness of the final path decomposition) such that one of them is the input vertex $v$. (W.l.o.g. $v$ has a neighbor in $G$, because otherwise $G$ contains a single vertex and therefore $P$ is connected.) Steps 1.2 and 1.3 are symmetric. In Step 1.2 (1.3) the algorithm finds the maximal left (right) branch ‘emanating’ from $x$ (resp. $y$) provided that the vertex belongs to the border of $C_1$, otherwise the left (right, respectively) border is empty.

In the following, one iteration of $CP$ or $EE$ means one iteration of the ‘while’ loop in the corresponding procedure. Thus, in the case of $CP$, one iteration reduces to executing Steps L.1-L.3 or R.1-R.3 within the ‘if’ statement, while in the procedure $EE$ one iteration results in executing the instructions in Step EL or in Step ER. We use the symbols $B_1, B_2, B_3, a_0, a_1, b_0, b_1, k_1, k_2'$ to refer to the variables used in $CP$, where in the case of $k_1$ we refer to its value at the end of Step L.1 or Step R.1. In what follows we denote for brevity $B'_2 = B_R(C_m, \delta_R(C_m)) \cup (V_{k_1} \cap \delta_L(C_m)), k_2'$ and $B'_3 = B_L(C_m, \delta_L(C_m), k_3')$ if Steps L.1-L.3 have been executed in this particular iteration of $CP$, and $B'_2 = B_L(C_m, \delta_L(C_m) \cup (V_{k_1} \cap \delta_R(C_m)), k_2')$, $B'_3 = B_R(C_m, \delta_R(C_m), k_3')$ otherwise (i.e. in Steps R.1-R.3). Informally speaking, $B'_2$ and $B'_3$ are the branches $B_2$ and $B_3$, respectively, restricted to the vertices up to the corresponding cut $k_2'$ or $k_3'$. The vertex $v$ selected to be in $C_1$ is called the starting vertex.

The branches are used in the subsequent iterations of the algorithm in the way presented in Figure 3, where $C_{s_1}$ refers to the expansion obtained at the end of Step L.1 of an iteration of $CP$, $i = 1, 2, 3$ (the execution of Steps R.1-R.3 is symmetric), and $C_{s_1}$ is the expansion from the beginning of the iteration. First, a branch $B_1$ is used to obtain $C_{s_1}$ from $C_{s_1}$ during the execution of Step L.1 of $CP$. It holds in particular $C_{s_1} = C_{s_1} \cup V(B_1)$, as stated in Lemma 5 below. It is guaranteed that $r(\delta_L(C_{s_1})) \leq k_1$. The (external) vertices in $V(B_1) \cap V_{k_1}$ have some neighbors in $V_{k_1-1} \setminus C_{s_1}$ and the algorithm calculates the right branch $B'_2$ (‘emanating’ from $k_1$) in Step L.2. Its right extremity, $k_2'$, may be strictly less than the left extremity of the new right border $\delta_R(C_{s_2})$ if $B'_2$ has no external vertices in $V_{k_2}$. Finally, a branch $B'_3$ is calculated in a symmetric way (this step is omitted if the vertices in $C_{s_1} \cap V_{k_1}$ have no neighbors in $V_{k_1-1} \setminus C_{s_1}$, and in such case $\delta_L(C_{s_1}) \subseteq \delta_L(C_{s_1})$).

The following lemmas are used to prove that the computation stops and they also demonstrate how the expansions change between the subsequent calls of $EE$.

► Lemma 4. Given an expansion $C_j$ and $X \subseteq \delta(C_j)$, after the execution of the $i$th iteration of the procedure $EE(r(X), k)$, where $k \leq r(X)$ (respectively $EE(l(X), k)$, where $k \geq l(X)$) it holds $C_{j+1} = C_j \cup V(B_L(C_j, X, r(X)-i))$ $(C_{j+1} = C_j \cup V(B_R(C_j, X, l(X)+i))$, respectively), $i \geq 1$. △

► Lemma 5. Let $C_{s_1}$ be an expansion from the beginning of an iteration of $CP$, and let $C_{s_1}$, $i = 1, 2, 3$, be the expansions obtained at the end of Steps L.1, L.2 and L.3 or R.1, R.2
Figure 3 The expansion $C_{s3}$ obtained from $C_{s0}$ by including three branches $B_1, B'_2$ and $B'_3$ calculated in Steps L.1, L.2 and L.3 of $CP$, respectively.

The expansion $C_{s3}$ obtained from $C_{s0}$ by including three branches $B_1, B'_2$ and $B'_3$ calculated in Steps L.1, L.2 and L.3 of $CP$, respectively. Then, $C_{s1} = C_{s0} \cup V(B_1)$, $C_{s2} = C_{s1} \cup V(B'_2)$ and $C_{s3} = C_{s2} \cup V(B'_3)$. Moreover, $C_{s3} \neq C_{s0}$.

Figure 4 gives an example of the execution of $CP$. In all cases (including the following figures) ◇ and □ are used to denote the vertices of the right and left borders, respectively. In particular Figure 4(a) presents a graph $G$ and $C_2$ (this is the expansion obtained at the end of initialization of $CP$, where the starting vertex and its neighbor in $C_1$ are among the three vertices in $C_2$). Figures 4(b)-(d) depict the state of the algorithm at the end of the first three iterations. (The fourth iteration executes the Steps L.1-L.3, which ends the computation.)

The lemma below follows directly from the instructions in procedure EE.

**Lemma 6.** $\delta(C_j) = \delta_L(C_j) \cup \delta_R(C_j)$ for each $j = 1, \ldots, m$.

The connectedness of $C$ is due to the fact that $G[C_j]$ is connected for each $j = 1, \ldots, m$, while the fact that $C$ is a path decomposition follows from the definition of $G$.

**Lemma 7.** Given a simple graph $G$ and its path decomposition $P = (X_1, \ldots, X_d)$, $CP$ returns a connected path decomposition $C = (Z_1, \ldots, Z_m)$ of $G$.

4 The approximation guarantee of the algorithm

In this section we analyze the width of the path decomposition $C$ calculated by $CP$ for the given $G$ and $P$. First we introduce the concept of a nested expansion, which, informally
speaking, is as follows. The first condition for $C$ to be nested states that the weight of $V_i \cap C$ for any $i$ ‘between’ the right extremity of the left border and the left extremity of the right border (by Lemma 9 the former is less than the latter) is greater than or equal to the weight of the left or the right border of $C$. The remaining conditions refer to the situation ‘inside’ borders and are analogous in both cases. The condition (ii) for the left border requires that the weight of $V_i \cap C$, where $i \leq r(\delta_L(C))$, is not less than the weight of the left border restricted to the vertices in $V_i \cup \cdots \cup V_j$. Finally, condition (iii) gives a ‘local’ minimality, that is, if we take a left branch $B_L(C, \delta_L(C), i)$ (where $i$ by the definition is $\leq r(\delta_L(C))$) and we include several vertices of the branch, as it is done in procedure EE, then we ‘arrive’ at some cut of this branch, and (iii) for $C$ guarantees that the weight of the left border of the new expansion is greater than or equal to the weight of the left border of $C$.

We say that an expansion $C$ is nested if it satisfies the following conditions:

(i) for each $i = r(\delta_L(C)), \ldots, l(\delta_R(C))$, $\min\{\omega(\delta_L(C)), \omega(\delta_R(C))\} \leq \omega(V_i \cap C)$,

(ii) for each $i \leq r(\delta_L(C))$, $\omega(V_i \cap C) \geq \sum_{j \leq i} \omega(V_j \cap \delta_C(C))$, and for each $i \geq l(\delta_R(C))$, $\omega(V_i \cap C) \geq \sum_{j \geq i} \omega(V_j \cap \delta_R(C))$,

(iii) $r(\delta_L(C)) ((l(\delta_R(C))))$ is a bottleneck of each branch $B_L(C, \delta_L(C), i)$ (respectively, $B_R(C, \delta_R(C), i)$).

Figure 5 presents a subgraph of $G$ on the vertices that belong to an expansion $C$. For this expansion to be nested it holds in particular: (ii) implies $\omega(V_{i+1} \cap C) \geq \omega(\{x_1, x_2, x_3\})$, $\omega(V_{i+3} \cap C) \geq \omega(\{x_1, x_2, x_3\})$; (i) implies $\omega(V_{i+6} \cap C) \geq \min\{\omega(\delta_L(C)), \omega(\delta_R(C))\} = \min\{\omega(\{x_1, x_3\}), \omega(\{y_1, y_4\})\}$.

**Figure 5** A nested expansion with left and right borders distinguished

Not all expansions computed by $\text{CP}$ are nested, but we prove that all of them satisfy (ii) (Lemmas 10-12). The fact that the expansions obtained in Steps I.1-I.3 of $\text{CP}$ satisfy (ii) follows from the observation that both the left and right border or each such expansion is a subset of a single set $V_i$. For this reason we focus on analyzing the subsequent iterations of $\text{CP}$, and we proceed with the assumption that an expansion from the beginning of an iteration (i.e. obtained at the end of the previous iteration, or at the end of Step I.3) is nested. We justify this assumption in Lemma 12.

First we introduce the following concept of moving the borders of an expansion. We say that $C_j$ moves the right (left) border of $C_{j-1}$ if $l(\delta_R(C_j)) > l(\delta_R(C_{j-1}))$ ($r(\delta_L(C_j)) < r(\delta_L(C_{j-1}))$, respectively). The following lemma states that in each iteration of $\text{CP}$ at most one expansion may be computed that does not move the left or right border of its predecessor. Moreover, this is the first expansion calculated in Step L.2 or in Step R.2, depending on the condition checked in the ‘if’ statement in the main loop of $\text{CP}$.

**Lemma 8.** If $C_j, j \in \{2, \ldots, m\}$, is any expansion calculated in Step EL (Step ER) of $\text{EE}$ invoked in an iteration of $\text{CP}$, except for an expansion obtained in the first iteration of $\text{EE}$ executed in Step L.2 or in Step R.2 of $\text{CP}$, then $C_j$ moves the left (right, resp.) border of $C_{j-1}$. △
Lemma 6 states that $\delta(C_j) = \delta_L(C_j) \cup \delta_R(C_j)$ for each expansion obtained in $CP$, while the following lemma implies that the left and right borders of any expansion obtained during the execution of $CP$ are disjoint.

**Lemma 9.** $r(\delta_L(C_j)) < l(\delta_R(C_j))$ for each $j = 1, \ldots, m$. \hfill $\triangle$

Provided that an expansion from the beginning of an iteration of $CP$ is nested, the following, together with Lemma 8, implies that if the first expansion computed in Step L.2 or R.2 of $CP$ satisfies (ii), then all expansions from the iteration satisfy (ii).

**Lemma 10.** Let $j \in \{2, \ldots, m\}$. If $C_{j-1}$ satisfies (ii) and $C_j$ moves the right or the left border of $C_{j-1}$, then $C_j$ satisfies (ii). \hfill $\triangle$

Due to the above, we analyze the first expansion obtained in Step L.2 or R.2 of $CP$. Let $C_{s_0}$ be the expansion from the beginning of an iteration of $CP$ and let $C_{s_1}$ be the expansion obtained at the end of Step L.1 or Step R.1 of this iteration. We obtain that if $C_{s_0}$ is nested, then $C_{s_1}$ satisfies (i) (the proof is omitted due to lack of space). This allows us to prove that the first expansion obtained in Step L.2 or R.2 of $CP$ also satisfies (ii). Thus, due to Lemmas 8 and 10 we obtain that each expansion in an iteration of $CP$ satisfies (ii), when the expansion from the beginning of the iteration is nested.

**Lemma 11.** If an expansion $C_{s_0}$ from the beginning of an iteration of $CP$ is nested, then each expansion computed by $CP$ in this iteration satisfies (ii). \hfill $\triangle$

Finally, we finish our argument that each expansion calculated by $CP$ satisfies (ii), by proving the following.

**Lemma 12.** Let $C_{s_0}$ and $C_{s_3}$ be the expansions from the beginning of two consecutive iterations of $CP$. If $C_{s_0}$ is nested, then $C_{s_3}$ is nested. \hfill $\triangle$

Therefore, an induction on the number of iterations of $CP$ allows us to prove the claim that each expansion computed by $CP$ satisfies (ii), as shown in Lemma 14. Lemma 13 below, together with Lemma 6, gives an upper bound for $\omega(\delta(C_j))$ for each $j = 1, \ldots, m$.

**Lemma 13.** If an expansion $C$ satisfies (ii), then $\omega(\delta_L(C)) \leq \text{width}(G)$ and $\omega(\delta_R(C)) \leq \text{width}(G)$.

**Proof.** Suppose w.l.o.g. that $\omega(\delta_R(C)) \geq \omega(\delta_L(C))$. Then, it is enough to argue that $\omega(\delta_L(C)) \leq \text{width}(G)$. To that end observe that by (ii), where $i = l(\delta_R(C))$, $\omega(V_i \cap C) \geq \sum_{k \geq i} \omega(\delta_R(C) \cap V_k) = \omega(\delta_R(C))$. Since $\omega(V_i \cap C) \leq \omega(V_i) \leq \text{width}(G)$, the thesis follows. \hfill $\triangleright$

**Lemma 14.** If $\mathcal{C} = (Z_1, \ldots, Z_m)$ is a path decomposition calculated by $CP$ for the given $G$ and $P$, then $\text{width}(\mathcal{C}) \leq 2 \cdot \text{width}(P) + 1$.

**Proof.** The expansion obtained at the end of Step I.3 of $CP$ is nested. Indeed, (i) and (iii) follow from the fact that $a_0^r$ and $b_0^r$ are the bottlenecks of the corresponding branches used in Steps I.2 and I.3, respectively, while (ii) trivially holds, for both the left and right border is contained in a single set $V_i$. Using an induction (on the number of iterations of $CP$) we obtain by Lemma 12 that any expansion from the beginning of an iteration of $CP$ is nested. Note that for each expansion $C_j$ obtained in Steps I.1-I.3 of $CP$ it holds $\delta_L(C_j) \subseteq V_i$ and $\delta_R(C_j) \subseteq V_{i'}$ for some $i, i' \in \{1, \ldots, d\}$, which implies (ii) for $C_j$. This, together with Lemma 11, implies that $C_j$ satisfies (ii) for each $j = 1, \ldots, m$. By Lemmas 6 and 9, $\omega(\delta(C_j)) = \omega(\delta_L(C_j)) + \omega(\delta_R(C_j))$ for each $j = 1, \ldots, m$. By Lemma 13, $\omega(\delta(C_j)) \leq 2 \cdot \text{width}(G)$. By the definition, $\text{width}(\mathcal{C}) = \max\{\omega(\delta(C_j)) : j = 1, \ldots, m\} - 1$. Thus, by the definition, $\text{width}(\mathcal{C}) \leq 2 \cdot \text{width}(G) - 1 = 2 \cdot \text{width}(P) + 1$. \hfill $\triangleright$
Lemma 15. Let $G$ be a simple connected graph and let $\mathcal{P} = (X_1, \ldots, X_d)$ be its path decomposition of width $k$. The running time of CP executed for $G$ and $\mathcal{P}$ is $O(dk^2)$.

Proof. Since each edge of $G$ is contained in one of the bags of $\mathcal{P}$, $|E(G)| \leq dk$. The number of vertices and edges in $G$ is $O(kd)$ and $O(dk^2)$, respectively. Thus, the complexity of constructing $G$ is $O(dk^2)$.

If a branch is given, then the weights of all its cuts can be calculated in time linear in the number of edges and vertices of the branch. The time of finding any branch $B$ in an iteration of CP is $O(|E(B)|)$. The complexity of calculating the weight of all cuts of $B$, and thus finding its bottleneck, is $O(|E(B)|)$. Whenever two branches overlap, we do not have to repeat the computation. Therefore, the time complexity of determining all branches and their bottlenecks is $O(dk^2)$. This includes the complexity of all executions of the procedure EE, because, by Lemma 5, the procedure ‘follows’ the previously calculated branches by including their vertices into the expansions $C_j$. It holds that $m \leq kd$, because (by Lemmas 4 and 5) $C_j \subseteq C_{j+1}$ and $C_j \neq C_{j+1}$ for each $j = 1, \ldots, m - 1$. By Lemma 14, $\omega(C_j) = O(k)$ for each $j = 1, \ldots, m$. Thus, $\sum_{1 \leq j \leq m} |Z_j| = O(dk^2)$. Thus, the complexity of CP is $O(dk^2)$.

Theorem 16. There exists a $O(dk^2)$-time algorithm that for given connected graph $G$ and its path decomposition $\mathcal{P} = (X_1, \ldots, X_d)$ of width $k$ returns a connected path decomposition $\mathcal{C} = (Z_1, \ldots, Z_m)$ such that $\text{width}(\mathcal{C}) \leq 2 \cdot \text{width}(\mathcal{P}) + 1$ and $m \leq kd$.

Proof. The correctness of CP is due to Lemma 7. The inequality $\text{width}(\mathcal{C}) \leq 2\text{width}(\mathcal{P}) + 1$ follows from Lemma 14, and the complexity of CP is due to Lemma 15. As argued in the proof of Lemma 15, $m \leq kd$.

Theorem 17. For each connected graph $G$, $\text{cpw}(G) \leq 2 \cdot \text{pw}(G) + 1$.

The inequalities $\text{pw}(G) \leq s(G) \leq \text{pw}(G) + 2$ and $\text{cpw}(G) \leq \text{cs}(G) \leq \text{cpw}(G) + 2$ [4] and Theorem 17 give the following

Corollary 18. For each graph $G$ it holds $\text{cs}(G) \leq 2 \cdot s(G) + 3$.

5 Conclusions

The advances in graph theory presented in this paper are three-fold:

- A bound for connected pathwidth is given, $\text{cpw}(G) \leq 2\text{pw}(G) + 1$, where $G$ is any graph, which bounds the connected search number of a graph by its search number, $\text{cs}(G) \leq 2s(G) + 3$. Moreover, the input vertex $v$ that belongs to the first bag in the resulting connected path decomposition is selected arbitrarily, which implies a stronger fact, namely a connected $(2s(G) + 3)$-search strategy can be constructed with any vertex of $G$ playing the role of the homebase. This provides an efficient algorithm for converting a search strategy into a connected one with an arbitrary homebase.

- An efficient method is given for calculating a connected pathwidth of width at most $2k + 1$, provided that a graph $G$ and its path decomposition of width $k$ are given.

- It is a strong assumption that the algorithm requires a path decomposition to be given, because calculating $\text{pw}(G)$ is a hard problem in general. However, this algorithm can be used to approximate the connected pathwidth for the classes of graphs for which the approximate algorithms for pathwidth exist.
References