Excluded Grid Theorem: Improved and Simplified

Julia Chuzhoy[*]
Toyota Technological Institute at Chicago
6045 S. Kenwood Ave
Chicago, IL 60637
cjulia@ttic.edu

ABSTRACT

We study the Excluded Grid Theorem of Robertson and Seymour. This is a fundamental result in graph theory, that states that there is some function $f : \mathbb{Z}^+ \to \mathbb{Z}^+$, such that for any integer $g > 0$, any graph of treewidth at least $f(g)$, contains the $(g \times g)$-grid as a minor. Until recently, the best known upper bounds on f were super-exponential in g. A recent work of Chekuri and Chuzhoy provided the first polynomial bound, by showing that treewidth $f(g) = O(g^{98} \operatorname{poly} \log g)$ is sufficient to ensure the existence of the $(g \times g)$-grid minor in any graph. In this paper we provide a much simpler proof of the Excluded Grid Theorem, achieving a bound of $f(g) = O(g^{36} \operatorname{poly} \log g)$. Our proof is self-contained, except for using prior work to reduce the maximum vertex degree of the input graph to a constant.

1. INTRODUCTION

We study the Excluded Grid Theorem of Robertson and Seymour [29] - a fundamental and widely used result in graph theory. Informally, the Excluded Grid Theorem states that for any undirected graph G, if the treewidth of G is large, then G contains a large grid as a minor. Graph treewidth is an important and extensively used graph parameter, that, intuitively, measures how close a given graph G is to being "tree-like". The treewidth of a graph is usually defined via tree-decompositions. A valid tree-decomposition of a graph G consists of a tree T, and, for each node $a \in V(T)$, a subset $X(a) \subseteq V(G)$ of vertices of G, sometimes called a *bag*. We require that for each edge $e = (u, v) \in E(G)$, there is a node $a \in V(T)$, whose bag $X(a)$ contains both u and v, and for each vertex $v \in V(G)$, the set $\mathcal{X} = \{a \in V(T) \mid v \in X(a)\}$ of nodes of T whose bags contain v forms a non-empty connected sub-tree of T. The *width* of a given tree decomposition (T, X) is $\min_{a \in V(T)} |X(a)| - 1$, and the *treewidth* of a graph G, denoted by $\operatorname{tw}(G)$, is the smallest width of any valid tree-decomposition of G. For example, the treewidth of a tree is 1; the treewidth of the $(g \times g)$-grid is $\Theta(g)$; and the treewidth of an n-vertex constant-degree expander is $\Theta(n)$. Many combinatorial optimization problems that are hard on general graphs, have efficient algorithms on trees, often via the dynamic programming technique. Such algorithms can frequently be extended to bounded-treewidth graphs, usually by applying the dynamic programming-based algorithms to the bounded-width tree-decomposition T of G. However, for large-treewidth graphs, a different toolkit is often needed. The Excluded Grid Theorem provides a useful insight into the structure of large-treewidth graphs, by showing that any such graph must contain a large grid as a minor. Recall that a graph H is a *minor* of a graph G, iff H can be obtained from G by a series of edge-deletion, edge-contraction, and vertex-deletion operations. We are now ready to formally state the Excluded Grid Theorem.

Theorem 1.1 *[29] There is some function $f : \mathbb{Z}^+ \to \mathbb{Z}^+$, such that for any integer $g \geq 1$, any graph of treewidth at least $f(g)$ contains the $(g \times g)$-grid as a minor.*

The Excluded Grid Theorem plays an important role in Robertson and Seymour's seminal graph minor series, and it is one of the key elements in their efficient algorithm for the Node-Disjoint Paths problem (where the number of the demand pairs is bounded by a constant) [30]. It is also widely used in Erdos-Pósa-type results (see, e.g. [32, 17, 29]) and in Fixed Parameter Tractability; in fact the Excluded Grid Theorem is the key tool in the bidimentionality theory [12, 14].

It is therefore important to study the best possible upper bounds on the function f, for which Theorem 1.1 holds. Besides being a fundamental graph-theoretic question in its own right, better upper bounds on f immediately result in faster algorithms and better parameters in may applications, e.g. in Fixed Parameter Tractability and Erdos-Pósa-type results. The original upper bound on f of [29] was substantially improved by Robertson, Seymour and Thomas [28] to $f(g) = 2^{O(g^5)}$. Diestel et al. [16] (see also [15]) provide a simpler proof with a slightly weaker bound. This was in turn improved by Kawarabayashi and Kobayashi [19], and by Leaf and Seymour [22], to $f(g) = 2^{O(g^2/\log g)}$. Finally, a recent work of Chekuri and Chuzhoy [2] provides the first polynomial upper bound on the function $f(g)$, by showing that Theorem 1.1 holds for $f(g) = O(g^{98} \operatorname{poly} \log g)$. On the negative side, Robertson et al. [28] show that $f(g) = \Omega(g^2 \log g)$ must hold, and they conjecture that this value is

[*]Supported in part by NSF grant CCF-1318242.

STOC'15, June 14–17, 2015, Portland, Oregon, USA.

http://dx.doi.org/10.1145/2746539.2746551 .

sufficient. Demaine et al. [13] conjecture that the bound of $f(g) = \Theta(g^3)$ is both necessary and sufficient.

In this paper we provide a proof of Theorem 1.1 with an improved bound of $f(g) = O(g^{36} \operatorname{poly} \log g)$. The main advantage of our proof is that, unlike the proof of [2], it is very simple conceptually. Our proof is almost self-contained, in the following sense: we provide a self-contained proof of Theorem 1.1 for bounded-degree graphs G. In order to handle general graphs, we need to use previously known results to reduce the maximum vertex degree of the input graph to a constant, while approximately preserving its treewidth. This is the only part of our proof that is not self-contained. We discuss this in more detail below. Unlike the proof of [2], that relies on many known technical tools, such as the cut-matching game of Khandekar, Rao and Vazirani [20], graph-reduction step preserving element-connectivity [18, 8], edge-splitting [23], and LP-based approximation algorithms for bounded-degree spanning tree [31] to name a few, our proof is entirely from first principles. The contribution of this paper is therefore two-fold: we provide a conceptually simple framework for proving the Excluded Grid Theorem, and show that it can be used to obtain a polynomial bound of $f(g)$; and we improve the bound of [2] on $f(g)$ from $O(g^{98} \operatorname{poly} \log g)$ to $O(g^{36} \operatorname{poly} \log g)$. We note that we have tried to present a simple proof, while simultaneously optimizing the bound on $f(g)$. Unfortunately, these two goals are sometimes conflicting, and we chose to compromise somewhat on both of them. Namely: the proof can be significantly simplified if we are only interested in obtaining a polynomial bound on $f(g)$; on the other hand, the bound that we obtain can be further improved, at the cost of making the proof more technical. Our hope is that the simple conceptual framework introduced in this paper will lead to significantly better bounds for the Excluded Grid Theorem.

There are two caveats in our proof. The first one, that we have already mentioned, is that it requires that the input graph G has a bounded degree. This can be achieved in several ways, using prior work. Reed and Wood [27] showed that any graph of treewidth k contains a sub-graph of maximum vertex degree 4, and treewidth $\Omega(k^{1/4}/\log^{1/8} k)$. Kreutzer and Tazari [21] gave a constructive proof of a similar result, with slightly weaker bounds. The algorithm of Chekuri and Ene [4] can be used to construct a sub-graph G' of the input treewidth-k graph G, such that the treewidth of G' is $\Omega(k/\operatorname{poly}\log k)$, and maximum vertex degree bounded by some constant. Finally, Chekuri and Chuzhoy [3] have recently shown that any graph G of treewidth k contains a sub-graph of maximum vertex degree 3, and treewidth $\Omega(k/\operatorname{poly}\log k)$. Unfortunately, this latter result builds on parts of the previous proof of the Excluded Grid Theorem of [2]. Therefore, if one is interested in a simple self-contained proof of Theorem 1.1, one should use the result of [27] as a starting point. In this paper we chose instead to use the result of [3] as our starting point, for two reasons. First, it gives the best bounds on both the degree and the treewidth of the resulting graph. Second, working with graphs whose maximum vertex degree is 3 is easier than with general constant-degree graphs, since routing on edge-disjoint and node-disjoint paths in such graphs is very similar. This saves on a number of technical steps and makes the proof easier to follow. The second caveat is that, unlike the proof of [2], that also provides an algorithm, whose running time is polynomial in n and g, to construct the grid minor, our proof is non-constructive. We believe that it can be turned into an algorithm whose running time is $2^{O(k)} \cdot \operatorname{poly}(n)$, where $k = \operatorname{poly}(g)$ is the treewidth of the input graph, using methods similar to those used in [2], but we have decided to keep the proof non-constructive for the sake of simplicity. It is however unlikely that our methods can give an algorithm whose running time is polynomial in both k and n, since we need to solve the sparsest cut problem (with k terminals) exactly. We note that most applications of the Excluded Grid Theorem (e.g. in Fixed-Parameter Tractability and in Erdos-Pósa–type results) only use the non-constructive version of the theorem. In other results, where a constructive version is used, such as the algorithm of Robertson and Seymour for the Node-Disjoint Paths problem [30], a running time $2^{O(k)} \cdot \operatorname{poly}(n)$ for finding the grid minor is acceptable, since the rest of the algorithm inherently incurs this (and in fact much higher) running time.

As in much prior work in this area, we use the notion of well-linkedness. We say that a set T of vertices is α-well-linked in graph H, for $0 < \alpha < 1$, iff for any pair $T', T'' \subseteq T$ of disjoint equal-sized subsets of vertices of T, there is a set $\mathcal{Q}(T', T'')$ of paths in H, connecting every vertex of T' to a distinct vertex of T'', such that every edge of H participates in at most $1/\alpha$ such paths. We will informally say that a set T of vertices is well-linked, if T is α-well-linked for some constant α. A central combinatorial object used in the proof of the Excluded Grid Theorem of [2], and that we also use here, is the Path-of-Sets system. We note that Leaf and Seymour [22] used a very similar, but somewhat weaker object, called a *grill*. A path-of-sets system of width r and height h consists of a sequence $\mathcal{S} = (S_1, \ldots, S_r)$ of r clusters, where for each cluster $S_i \subseteq V(G)$, we are given two disjoint subsets $A_i, B_i \subseteq S_i$ of h vertices each. We require that the vertices of $A_i \cup B_i$ are well-linked in $G[S_i]$. Additionally, for each $1 \leq i < r$, the path-of-sets system contains a set $\mathcal{P}_i$ of h paths, connecting every vertex of B_i to a distinct vertex of A_{i+1}. The paths in $\bigcup_i \mathcal{P}_i$ must be all mutually disjoint, and they cannot contain the vertices of $\bigcup_{i'=1}^{r} S_{i'}$ as inner vertices. Chekuri and Chuzhoy [2], strengthening a similar result of Leaf and Seymour [22], showed that if a graph G contains a path-of-sets system of height $\Theta(g^2)$ and width $\Theta(g^2)$, then G contains the $(g \times g)$-grid as a minor. Therefore, in order to prove the Excluded Grid Theorem, it is now enough to prove that there is a function $f : \mathbb{Z}^+ \to \mathbb{Z}^+$, such that any graph of treewidth at least $f(g)$ contains a path-of-sets system of height $\Omega(g^2)$ and width $\Omega(g^2)$.

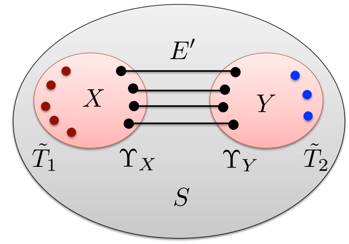
Chekuri and Chuzhoy [2] showed this to be true for $f(g) = O(g^{98} \operatorname{poly} \log g)$, and we prove it here for $f(g) = O(g^{36} \operatorname{poly} \log g)$.

We now briefly summarize the proof of [2], before we describe our proof. It is well-known (see e.g. [26]), that if a graph G has treewidth k, then there is a subset $T \subseteq V(G)$ of $k/4$ vertices, such that T is well-linked in G. Throughout the proof, we will refer to the vertices of T as *terminals*. Given any cluster $C \subseteq V(G)$, we will denote by $\operatorname{out}(C)$ the set of edges of G with exactly one endpoint in C, and by $\Gamma(C)$ the *boundary* of C — the set of vertices of C incident on the edges of $\operatorname{out}(C)$.

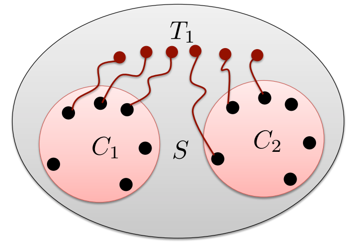
The proof of [2] consists of four steps. In the first step, they show that any graph G of treewidth k contains a large collection $\mathcal{S}$ of disjoint *good routers*. Informally, a good router is a cluster $C \subseteq V(G)$, such that (i) the boundary of C is well-linked in $G[C]$; and (ii) there is a set $\mathcal{P}(C)$ of k^ϵ disjoint paths, for some constant $0 < \epsilon < 1$, connecting the terminals to the vertices of C. The construction of the routers involves several old and new techniques, such as building a contracted graph that "hides" irrelevant information about G by contracting some clusters; random partitions of graphs; and the so-called well-linked decompositions. In the second step, the clusters of $\mathcal{S}$ are "organized" into a tree: that is, we construct an object, called a tree-of-sets system, that is similar to the path-of-sets system, except that the clusters are connected via a tree-like structure instead of a path-like structure. This step involves carefully removing vertices of G that do not belong to the clusters of $\mathcal{S}$, while preserving the connectivities between the clusters of $\mathcal{S}$, by using graph-reduction steps preserving element-connectivity, together with standard edge-splitting. If the resulting tree-of-sets system has a long root-to-leaf path, then we can use this path as the final path-of-sets system. Otherwise, let $\mathcal{S}' \subseteq \mathcal{S}$ be the subset of clusters that serve as the leaves of the tree. In the third step, we repeat Step 2 on the clusters of $\mathcal{S}'$ instead of the clusters of $\mathcal{S}$, and a carefully selected sub-graph G' of G, to ensure that the tree corresponding to the resulting tree-of-sets system has maximum vertex degree at most 3. This step relies on an LP-based approximation algorithm for bounded-degree spanning trees of [31]. Finally, in the fourth step, we turn the resulting tree-of-sets system into a path-of-sets system, by carefully simulating a DFS tour of the corresponding tree.

In contrast, our algorithm consists of only one subroutine, that, intuitively, shows that, given any path-of-sets system of width 1 and height h, we can obtain a path-of-sets system of width 2 and height h/c', for some constant c'. More specifically, suppose we are given some subset S of vertices of G, and two disjoint subsets $T_1, T_2 \subseteq S$ of vertices, such that $|T_1| = h/c$ (where c is some constant), $|T_2| = h$, and $(T_1 \cup T_2)$ is well-linked in $G[S]$. We show that there are two disjoint clusters X, Y in S, a subset $E' \subseteq E(X,Y)$ of h/c^2 edges whose endpoints are all distinct, and two subsets $\tilde{T}_1 \subseteq X \cap T_1$ of at least h/c^2 vertices and $\tilde{T}_2 \subseteq Y \cap T_2$ of at least h/c vertices, such that, if we denote by Υ_X and Υ_Y the endpoints of the edges of E' that belong to X and Y, respectively, then $\Upsilon_X \cup \tilde{T}_1$ is well-linked in $G[X]$, and $\Upsilon_Y \cup \tilde{T}_2$ is well-linked in $G[Y]$ (see Figure 1(a)). We call the corresponding tuple $(X, Y, \tilde{T}_1, \tilde{T}_2, E')$ a 2-*cluster chain*, and we call this procedure a *splitting of a cluster*. Using this procedure, it is now easy to complete the proof of the Excluded Grid Theorem. Let k be the treewidth of the input bounded-degree graph G. Our algorithm performs $2\log_2 g$ phases, where each phase j starts with a path-of-sets system of width 2^{j-1} and height $k/(8c^{2(j-1)})$, and produces a path-of-sets system of width 2^j and height $k/(8c^{2j})$. For our initial path-of-sets system of width 1 and height $k/8$, we use $S_1 = V(G)$, and we let (A_1, B_1) be any partition of the terminals into equal-sized subsets. Clearly, after $2\log_2 g$ phases, we obtain a path-of-sets system of width g^2 and height $k/(8g^{4\log c})$. Each phase is executed by simply splitting each cluster of the current path-of-sets system into two, using the cluster-splitting procedure described above. We omit the technical details, that can be found in Section 3.

We now briefly sketch our algorithm for splitting a cluster S. We note that this is an informal and imprecise overview, that is only intended to provide intuition. Let $k' = |T_1|$. We start by defining a slightly weaker object, called a *weak* 2-*cluster chain*. This object consists of two disjoint clusters $C_1, C_2 \subseteq S \setminus (T_1 \cup T_2)$, such that for $i \in \{1, 2\}$, the interface



(a) Strong 2-cluster chain.



(b) Weak 2-cluster chain.

Figure 1: Splitting a cluster.

vertices of C_i are well-linked in $G[C_i]$, and there is a set $\mathcal{P}_i$ of $\Omega(k')$ node-disjoint paths, connecting the vertices of C_i to the terminals of T_1, such that the paths in $\mathcal{P}_1 \cup \mathcal{P}_2$ are disjoint from each other, and do not contain the vertices of $C_1 \cup C_2$ as inner vertices (see Figure 1(b)). We show that the existence of the weak 2-cluster chain is sufficient to guarantee the existence of the (strong) 2-cluster chain in $G[S]$: the idea is to use the well-linkedness of the set $T_1 \cup T_2$ of vertices, to carefully connect the two clusters C_1, C_2 to each other, and to connect one of them to the set T_2 of vertices, by large enough collections of disjoint paths. The main technical ingredient of the paper is therefore showing that any cluster S, with two disjoint subsets $T_1, T_2 \subseteq S$ of vertices, where $(T_1 \cup T_2)$ is well-linked in $G[S]$, contains a weak 2-cluster chain.

From now on we denote $G[S]$ by H, and we call the vertices in $T = T_1 \cup T_2$ terminals. Let $C \subseteq V(H) \setminus T$ be any cluster of non-terminal vertices of H, and let (A, B) be the minimum balanced cut of $H \setminus C$, with respect to T_1. In other words, (A, B) is a partition of $V(H) \setminus C$, with $|A \cap T_1|, |B \cap T_1| \geq |T_1|/4$, such that $|E(A,B)|$ is minimized among all such partitions. Informally, we say that cluster C is a good cluster if $|E(A,B)| < k'/28$, and we say that it is a perfect cluster if $k'/28 \leq |E(A,B)| \leq 7k'/32$ (where $k' = |T_1|$). The main observation in the proof is that if we find a perfect cluster C, such that $\Gamma(C)$ is α^*-well-linked in $H[C]$, for some constant α^*, then we can find a weak 2-cluster chain in H. In order to find such a cluster C, we first observe that there is a good cluster $C_0 \subseteq V(H) \setminus T$, whose interface vertices are well-linked in $H[C_0]$: the cluster $C_0 = V(H) \setminus T$. Among all such good clusters C_0, we choose one minimizing $|\operatorname{out}(C_0)|$, and among all such clusters, we choose one minimizing $|C_0|$. We then compute a minimum balanced cut (Z, Z') of C_0, with respect to the interface vertices $\Gamma(C_0)$. Assume without loss of generality that $|\Gamma(C_0) \cap Z| \leq |\Gamma(C_0) \cap Z'|$. It is not hard to see that if $|\Gamma(C_0) \cap Z|$ is small enough, then Z' must be a perfect cluster, whose interface vertices are sufficiently well-linked in

$H[Z']$ to guarantee the existence of the weak 2-cluster chain. Otherwise, we use the two clusters Z and Z', together with some useful properties of balanced cuts, in order to construct a weak 2-cluster chain directly.

Organization. We start with Preliminaries in Section 2, and provide an overview of our algorithm in Section 3. We then provide our algorithm for splitting a cluster in Section 4.

2. PRELIMINARIES

Given a graph $G = (V, E)$ and a set $A \subseteq V$ of vertices, we denote by $E_G(A)$ the set of edges with both endpoints in A. For two disjoint sets $A, B \subseteq V$, the set of edges with one endpoint in A and the other in B is denoted by $E_G(A, B)$. The degree of a vertex $v \in V$ is denoted by $d_G(v)$. Given a set $\mathcal{P}$ of paths in G, we denote by $V(\mathcal{P})$ the set of all vertices participating in paths in $\mathcal{P}$. We sometimes refer to sets of vertices as *clusters*. Given a cluster $C \subseteq V$, we denote by $\text{out}_G(C)$ the set of edges with exactly one endpoint in C, and by $\Gamma_G(C)$ the set of vertices of C incident on the edges of $\text{out}_G(C)$. We sometimes call $\Gamma_G(C)$ *the boundary of* C. We may omit the subscript G if it is clear from the context. We say that a path P is *internally disjoint* from a set U of vertices, if no vertex of U serves as an inner vertex of P. We say that two paths P, P' are *internally disjoint*, iff for each vertex $v \in V(P) \cap V(P')$, v is an endpoint of both paths. Let $\mathcal{P}$ be any collection of paths in graph G. We say that the paths in $\mathcal{P}$ cause edge-congestion η, if every edge $e \in E$ is contained in at most η paths in $\mathcal{P}$.

Assume that we are given two subsets $S, T \subseteq V$ of vertices. We denote by $\mathcal{P} : S \rightsquigarrow T$ a collection $\mathcal{P} = \{P_v \mid v \in S\}$ of paths, where path P_v has v as its first vertex and some vertex of T as its last vertex. Notice that each path of $\mathcal{P}$ originates from a distinct vertex of S, and $|\mathcal{P}| = |S|$. If additionally the set $\mathcal{P}$ of paths causes edge-congestion at most η, then we denote this by $\mathcal{P} : S \rightsquigarrow_\eta T$. Assume now that $|S| = |T| = |\mathcal{P}|$, and each path in $\mathcal{P}$ connects a distinct vertex of S to a distinct vertex of T. Then we denote $\mathcal{P} : S \overset{1:1}{\rightsquigarrow} T$, and if the paths in $\mathcal{P}$ cause edge-congestion at most η, then we denote $\mathcal{P} : S \overset{1:1}{\rightsquigarrow}_\eta T$. Notice that the paths of $\mathcal{P}$ are allowed to contain the vertices of $S \cup T$ as inner vertices. Similarly, flow F from the vertices of S to the vertices of T, where every vertex of S sends one flow unit, every vertex of T receives one flow unit, and every edge carries at most η flow units is denoted by $F : S \overset{1:1}{\rightsquigarrow}_\eta T$. We will repeatedly use the following simple observation, whose proofis omitted here.

Observation 2.1 *Let G be any graph, with maximum vertex degree at most 3, and $T_1, T_2 \subseteq V(G)$ any pair of* **disjoint** *equal-sized subset of vertices, such that the degree of every vertex in $T_1 \cup T_2$ is at most 2. Let $\mathcal{P} : T_1 \overset{1:1}{\rightsquigarrow}_1 T_2$ be any set of edge-disjoint paths connecting every vertex of T_1 to a distinct vertex of T_2. Then the paths in $\mathcal{P}$ are node-disjoint.*

2.1 Linkedness, Well-Linkedness, and Bandwidth Property

The notion of well-linkedness has played a central role in algorithms for routing problems (see e.g. [24, 7, 6, 25, 1, 10, 11, 4]), and is also often used in graph theory. Several different variations of this notion were used in the past. The definitions we use here are equivalent to those used in [10, 11, 4, 2], but for convenience we define them slightly differently.

Definition 2.1 *Given a graph G, a subset $T \subseteq V(G)$ of vertices, and a parameter $0 < \alpha < 1$, we say that T is α-well-linked in G, iff for any pair of disjoint equal-sized subsets $T', T'' \subseteq T$, there is a flow $F : T' \overset{1:1}{\rightsquigarrow}_{1/\alpha} T''$ in G.*

Notice that from the integrality of flow, if T is α-well-linked, then for any pair of disjoint equal-sized subsets $T', T'' \subseteq T$, there is also a set $\mathcal{P} : T' \overset{1:1}{\rightsquigarrow}_{\lceil 1/\alpha \rceil} T''$ of paths in G. The next observation relates our definition to the one used in [7, 6, 10, 11, 4, 2], and its proof is omitted here.

Observation 2.2 *Assume that we are given a vertex set $T \subseteq V(G)$ and a parameter $0 < \alpha < 1$, such that T is* **not** *α-well-linked in G. Then there is a partition (A, B) of $V(G)$, with $|E(A, B)| < \alpha \cdot \min\{|A \cap T|, |B \cap T|\}$.*

We call the partition given in Observation 2.2 *an α-violating partition* of G with respect to T. We also need a slightly more general definition of well-linkedness, similar to that introduced in [10].

Definition 2.2 *Given an integer k', and a parameter $0 < \alpha < 1$, we say that a set T of vertices is (k', α)-well-linked in graph G, iff for any pair of disjoint subsets $T', T'' \subseteq T$, with $|T'| = |T''| \leq k'$, there is a flow $F : T' \overset{1:1}{\rightsquigarrow}_{1/\alpha} T''$ in G.*

Notice that if $|T| \leq 2k'$, then T is α-well-linked in G iff it is (k', α)-well-linked in G. Notice also that if a set T of terminals is (k', α)-well-linked in G, then so is any subset $T' \subseteq T$. As before, if set T is (k', α)-well-linked in G, then for any pair of disjoint subsets $T', T'' \subseteq T$, with $|T'| = |T''| \leq k'$, there is a set $\mathcal{P} : T' \overset{1:1}{\rightsquigarrow}_{\lceil 1/\alpha \rceil} T''$ of paths in G. The following observation is an analogue of Observation 2.2, and its proof is omitted here.

Observation 2.3 *Assume that we are given a set T of vertices of G, an integer $k' > 0$, and a parameter $0 < \alpha < 1$. Assume further that T is* **not** *(k', α)-well-linked in G. Then there is a partition (A, B) of $V(G)$, such that $|E(A, B)| < \alpha \cdot \min\{|A \cap T|, |B \cap T|, k'\}$.*

We call the partition given in Observation 2.3 a (k', α)-*violating partition* with respect to T. We next define the notion of bandwidth property, somewhat similar to the one defined in [24].

Definition 2.3 *Given an integer k' and a parameter $\alpha > 0$, we say that a cluster C has the (k', α)-bandwidth property, iff $\Gamma_G(C)$ is (k', α)-well-linked in $G[C]$. We say that it has the α-bandwidth property, iff $\Gamma_G(C)$ is α-well-linked in $G[C]$.*

The following observation is immediate from the definition of the bandwidth property.

Observation 2.4 *Let G be any connected graph, and let C be any cluster of G that has the (k', α)-bandwidth property, for any integer $k' \geq 2$, and any parameter $\alpha > 0$. Then $G[C]$ is connected.*

We now define a stronger notion of well-linkedness, called *node-well-linkedness.*

Definition 2.4 *We say that a set T of vertices is* node-well-linked *in G, iff for any pair (T', T'') of disjoint equal-sized subsets of T, there is a collection $\mathcal{P} : T' \overset{1:1}{\leadsto} T''$ of* **node-disjoint** *paths in G.*

Notice that from Observation 2.1, if G is a graph with maximum vertex degree at most 3, and T is a set of vertices of degree at most 2 each, then T is node-well-linked in G iff T is 1-well-linked in G. Finally, we define the notion of linkedness between a pair of vertex subsets.

Definition 2.5 *We say that two disjoint subsets T_1, T_2 of vertices of G are α-linked for $0 < \alpha < 1$, iff for any pair $T_1' \subseteq T_1$ and $T_2' \subseteq T_2$ of equal-sized vertex subsets, there is a flow $F : T_1' \overset{1:1}{\leadsto}_{1/\alpha} T_2'$ in G.*

Notice that as before, if T_1 and T_2 are α-linked, then for any pair $T_1' \subseteq T_1$ and $T_2' \subseteq T_2$ of equal-sized vertex subsets, there is a set $\mathcal{P} : T_1' \overset{1:1}{\leadsto}_{\lceil 1/\alpha \rceil} T_2'$ of paths G.
The following lemma summarizes an important connection between the graph treewidth, and the size of the largest node-well-linked set of vertices in it.

Lemma 2.1 *[26] Let k be the size of the largest node-well-linked vertex set in G. Then $k \leq \operatorname{tw}(G) \leq 4k$.*

2.2 Balanced Cuts

Definition 2.6 *Let G be any graph, and $T \subseteq V(G)$ any subset of its vertices. Given a parameter $0 < \rho \leq 1/2$, a partition (A, B) of $V(G)$ is called a* ρ-balanced cut *of G with respect to T, iff $|A \cap T|, |B \cap T| \geq \rho |T|$. It is called a* minimum ρ-balanced cut *of G with respect to T, if it minimizes $|E(A, B)|$ among all ρ-balanced cuts, and subject to this, minimizes $\min\{|A \cap T|, |B \cap T|\}$.*

We will use the following lemma, whose proof uses standard techniques (a variation of the so-called well-linked decompositions). The proof is omitted from this extended abstract.

Lemma 2.2 *Let $G = (V, E)$ be any graph, and $C \subsetneq V$ any cluster that has the α-bandwidth property, for some $0 < \alpha \leq 1$. Let (A, B) be the minimum ρ-balanced cut of $G[C]$ with respect to $\Gamma(C)$, for some $0 < \rho \leq 1/4$, and assume that $|\Gamma(C) \cap A| \geq |\Gamma(C) \cap B|$. Then A has the $\alpha/(2+\alpha)$-bandwidth property.*

2.3 Treewidth and Degree Reduction

Our proof of Theorem 1.1 assumes that the maximum vertex degree of the input graph G is bounded by a constant. There are several known results, that, given a graph G of treewidth k, find a sub-graph G' of G, whose maximum vertex degree is bounded by a constant, and whose treewidth is close to $\operatorname{tw}(G)$. For example, Reed and Wood [27] have shown that any graph of treewidth k contains a sub-graph of maximum vertex degree at most 4, and treewidth $\Omega(k^{1/4}/\log^{1/8} k)$. The algorithm of Chekuri and Ene [4] can be used to construct a sub-graph G' of G of treewidth $k/\operatorname{poly}\log k$, and maximum vertex degree bounded by some constant. We use the following stronger result of [3]:

Theorem 2.1 *[3] Let G be any graph of treewidth k. Then there is a sub-graph G' of G, whose maximum vertex degree is 3, and $\operatorname{tw}(G') = \Omega(k/\operatorname{poly}\log k)$. Moreover, there is a set $T \subseteq V(G')$ of $\Omega(k/\operatorname{poly}\log k)$ vertices, such that T is 1-well-linked in G', and each vertex of T has degree 1 in G'.*

The starting point of the above theorem is a path-of-sets system of width $\operatorname{poly}\log k$, and height $k/\operatorname{poly}\log k$, whose existence follows from [2]. We chose to use Theorem 2.1 as our starting point, since it provides the best parameters, and, due to Observation 2.1, degree-3 graphs are somewhat easier to work with. But our proof can work as well using the result of [27] as a starting point instead.

2.4 A Path-of-Sets System

A central combinatorial object that we use is the path-of-sets system, introduced in [2]. A closely related object, called a grill, was previously defined by Leaf and Seymour [22].

Definition 2.7 *A path-of-sets system $(\mathcal{S}, \bigcup_{i=1}^{r-1} \mathcal{P}_i)$ of width r and height h in graph G consists of:*

- *A sequence $\mathcal{S} = (S_1, \ldots, S_r)$ of r disjoint vertex subsets of G, where for each i, $G[S_i]$ is connected;*
- *For each $1 \leq i \leq r$, two disjoint sets $A_i, B_i \subseteq S_i$ of h vertices each; the vertices of $A_1 \cup B_r$ must have degree 1 in G; and*
- *For each $1 \leq i < r$, a set $\mathcal{P}_i : B_i \overset{1:1}{\leadsto} A_{i+1}$ of h paths, such that all paths in $\bigcup_i \mathcal{P}_i$ are mutually node-disjoint, and do not contain the vertices of $\bigcup_{S_j \in \mathcal{S}} S_j$ as inner vertices.*

*We say that it is an α-*weak *path-of-sets system, if for all $1 \leq i \leq r$, $A_i \cup B_i$ is α-well-linked in $G[S_i]$; we say that it is a* good *path-of-sets system, if for all $1 \leq i \leq r$, B_i is 1-well-linked in $G[S_i]$, and (A_i, B_i) are $\frac{1}{2}$-linked in $G[S_i]$. Finally, we say that it is a* perfect *path-of-sets system, if for each $1 \leq i \leq r$, A_i is node-well-linked in $G[S_i]$, B_i is node-well-linked in $G[S_i]$, and (A_i, B_i) are 1-linked in $G[S_i]$.*

The following theorem allows us to turn an α-weak path-of-sets system into a good one, and eventually into a perfect one, with only a small loss in the system's height. The proof is omitted from this extended abstract. A simpler proof, with somewhat weaker parameters, can be found in [2].

Theorem 2.2 *Let G be a graph with maximum vertex degree 3, and suppose we are given an α-weak path-of-sets system of height h and width r in G, where $0 < \alpha \leq 1$, and $1/\alpha$ is an integer. Then G contains a good path-of-sets system of height $\lceil \alpha h/4 \rceil$ and width r, and it contains a perfect path-of-sets system of height at least $\alpha h/c^*$, for some constant c^*, and width r.*

The following theorem, whose proof appears in [2], slightly improves upon a similar result of [22].

Theorem 2.3 *[2] Let G be any graph and let $(\mathcal{S}, \bigcup_{i=1}^{r-1} \mathcal{P}_i)$ be a perfect path-of-sets system of height h and width h in G. Then G contains the $\left(\Omega(\sqrt{h}) \times \Omega(\sqrt{h})\right)$-grid as a minor.*

3. OVERVIEW

In this section we provide an overview of our proof of Theorem 1.1. Let G be any graph of treewidth $\kappa = \Omega(g^{36} \operatorname{poly}\log g)$. We assume that g is a power of 2 - otherwise we round it up to the nearest power of 2, thereby at most doubling it. We use Theorem 2.1 to obtain a sub-graph G' of G, whose maximum vertex degree is 3, together with a set T of $\kappa^* = \Omega(\kappa/\operatorname{poly}\log\kappa)$ terminals, such that the terminals of T are 1-well-linked in G', and the degree of every terminal is 1. From now on we will be working with graph G' only, so to simplify the notation, we denote G' by G, and κ^* by κ. We assume that $\kappa > cg^{36}$ for some large constant $c > 24000$, and it is a power of 2 - otherwise we round it down to the closest power of 2. We discard terminals from T until $|T| = \kappa$ holds.

For $0 \le j \le 2\log g$, let $r_j = 2^j$ and $h_j = \frac{\kappa}{2^{17j+1}}$. We perform $2\log_2 g$ phases. The input to phase j, for $1 \le j \le 2\log_2 g$, is a good path-of-sets system of width r_{j-1} and height h_{j-1}, and the output is a good path-of-sets system of width r_j and height h_j. The input to the first phase is a path-of-sets system of width 1 and height $\kappa/2$, constructed as follows. We let $\mathcal{S} = (S_1)$, where $S_1 = V(G)$, and we let (A_1, B_1) be is any partition of T into two equal-sized subsets. Since the terminals of T are 1-well-linked in G, it is immediate to verify that this is a good path-of-sets system of width $r_0 = 1$ and height $h_0 = \kappa/2$. Clearly, after $2\log_2 g$ iterations, we will obtain a good path-of-sets system of width g^2 and height $\frac{\kappa}{2^{34\log_2 g+1}} = \frac{\kappa}{2g^{34}} = \Omega(g^2)$. From Theorem 2.2, there is a perfect path-of-sets system of width g^2 and height g^2 in G, and from Theorem 2.3, G contains the $(\Omega(g) \times \Omega(g))$-grid as a minor. The execution of each phase is summarized in the following theorem, whose proof finishes the proof of Theorem 1.1.

Theorem 3.1 *Suppose we are given a graph G with maximum vertex degree 3, and a good path-of-sets system $(\mathcal{S}, \bigcup_{i=1}^{r-1}\mathcal{P}_i)$ of width r and height h, where h is a power of 2. Then there is a good path-of-sets system of width $2r$ and height $h/2^{17}$ in G.*

From now on we focus on proving Theorem 3.1. The central combinatorial object that we use is a two-cluster chain (that can intuitively be thought of as a path-of-sets system of width 2, except that the sizes of A_1, B_1, A_2, B_2 are no longer uniform).

Definition 3.1 *Let G be a graph, T_1, T_2 two disjoint sets of vertices, with $|T_1| = k$ and $|T_2| = k' = k/64$, where $k \ge 12000$ is a power of 2. A 2-cluster chain $(X, Y, \tilde{T}_1, \tilde{T}_2, E')$ consists of:*

- *two disjoint clusters $X, Y \subseteq V(G)$;*
- *a subset $\tilde{T}_1 \subseteq T_1 \cap X$, with $|\tilde{T}_1| = k'$, and a subset $\tilde{T}_2 \subseteq T_2 \cap Y$, with $|\tilde{T}_2| = k/512$;*
- *a set $E' \subseteq E(X, Y)$ of $k/512$ edges, whose endpoints are all distinct;*

  Let $\Upsilon_X \subseteq X$ be the subset of vertices of X incident on the edges of E', and let $\Upsilon_Y \subseteq Y$ be the subset of vertices of Y incident on the edges of E'. Then:
- *$\tilde{T}_1 \cup \Upsilon_X$ is $(k/512, \alpha^*)$-well-linked in $G[X]$ and $\tilde{T}_2 \cup \Upsilon_Y$ is $(k/512, \alpha^*)$-well-linked in $G[Y]$, for $\alpha^* = 1/64$. (See Figure 1(a)).*

The main technical contribution of our paper is the following theorem, that is proved in Section 4.

Theorem 3.2 *Suppose we are given a graph G, with maximum vertex degree at most 3, and two disjoint subsets of vertices, T_1 of size k (where $k \ge 12000$ is an integral power of 2), and T_2 of size $k' = k/64$, such that the degree of every vertex in $T_1 \cup T_2$ is 1 in G, the vertices of T_1 are 1-well-linked, and (T_1, T_2) are $\frac{1}{2}$-linked in G. Then there is a 2-cluster chain in G.*

We are now ready to complete the proof of Theorem 3.1. Let $\mathcal{P} = \bigcup_{i=1}^{r-1}\mathcal{P}_i$. For convenience, for each path $P \in \mathcal{P}$, we delete all edges an inner vertices of P from the graph, and instead add a new vertex t_P, that connects to the two endpoints u, v of P. Let $P' = (u, t_P, v)$ be the resulting path. We denote the resulting graph by G'. For each $1 \le i \le r-1$, let $Z_i = \{t_P \mid P \in \mathcal{P}_i\}$. We also let Z'_0 be any subset of $h/64$ vertices of A_1, and $Z_r = B_r$. We perform r iterations, where the ith iteration splits cluster S_i. We assume that for each $1 \le i \le r$, when iteration i starts, we are given a subset $Z'_{i-1} \subseteq Z_{i-1}$ of $h/64$ vertices. In the ith iteration, we apply Theorem 3.2 to graph $G_i = G'[S_i \cup Z'_{i-1} \cup Z_i]$, with $T_1 = Z_i$, and $T_2 = Z'_{i-1}$. Since the path-of-sets system is good, it is easy to see that T_1 is 1-well-linked in G_i, and (T_1, T_2) are $\frac{1}{2}$-linked. Let (X_i, Y_i) be the resulting pair of clusters, $E_i = E'$ the corresponding set of edges, and $Z''_{i-1} = \tilde{T}_2$, $Z'_i = \tilde{T}_1$ the corresponding vertex subsets. We then continue to the next iteration. Consider the final collection $(Y_1, X_1, \ldots, Y_r, X_r)$ of clusters obtained after r iterations. Then for each $0 \le i \le r$, $|Z''_i| = h/512$. We build an α^*-weak path-of sets system $(\mathcal{S}', \bigcup_{i=1}^{2r-1}\mathcal{P}'_i)$, as follows. We let $\mathcal{S}' = (S'_1, \ldots, S'_{2r})$, where for $1 \le i \le r$, $S'_{2i-1} = Y_i$ and $S'_{2i} = X_i$. For $1 \le i \le r$, let $\mathcal{P}'_{2i-1}$ be any subset of $h/512$ edges of E_i. For $1 \le i < r$, we let $\mathcal{P}'_{2i}$ be the set of paths $P \in \mathcal{P}_i$, where $t_P \in Z''_i$. We let $A'_1 = Z''_0$, and B'_{2r} any subset of $h/512$ vertices of Z'_r. For $1 \le i < 2r$, we let B'_i be the subset of the endpoints of the paths in $\mathcal{P}'_i$ that lie in S'_i, and we let A'_{i+1} be the subset of their endpoints that lie in S'_{i+1}. It is now easy to verify that we obtain an α^*-weak path-of-sets system of width $2r$ and height $h/512$, where $\alpha^* = 1/64$. From Theorem 2.2, we can now obtain a good path-of-sets system of width $2r$ and height at least $\frac{h}{512} \cdot \frac{\alpha^*}{4} = \frac{h}{2^{17}}$.

4. SPLITTING A CLUSTER

The goal of this section is to prove Theorem 3.2. We denote $T = T_1 \cup T_2$, and we call the vertices of T *terminals*. Recall that $|T_1| = k$, $|T_2| = k' = k/64$, and we denote $k'' = k/512$. Let G^* be a minimal (with respect to edge- and vertex-deletion) sub-graph of G, in which T_1 is $(k/4, 1)$-well-linked, and (T_1, T_2) are $\frac{1}{2}$-linked. For each terminal $t \in T$, we subdivide the unique edge incident on t by a new vertex v_t. It is easy to see that a 2-cluster chain in G^* immediately defines a 2-cluster chain in G. From now on we will be working with graph G^*, and for simplicity of notation, we denote G^* by G. Our goal is to show that G contains a 2-cluster chain. Notice that from the minimality of G, it is a connected graph.

Given a cluster $C \subseteq V(G) \setminus T$, we denote by $\mathcal{P}(C)$ the maximum-cardinality set of node-disjoint paths connecting the terminals in T to $\Gamma(C)$, and we denote $p(C) = |\mathcal{P}(C)|$. We assume w.l.o.g. that the paths in $\mathcal{P}(C)$ are internally disjoint from $C \cup T$. The following lemma can be seen as a variation of the Deletable Edge Lemma of Chekuri, Khanna and Shepherd [5] (the proof of their original lemma can be found in [9]), though it is somewhat simpler. The proof is omitted from this extended abstract.

Lemma 4.1 *Let C be any cluster of G, with $C \cap T = \emptyset$, such that C has the $(k/4, 1)$-bandwidth property. Then $p(C) = |\Gamma(C)|$.*

4.1 Weak 2-Cluster Chain

In this section we define a weak 2-cluster chain, which is somewhat weaker than the 2-cluster chain defined in Section 3. We then show that if G contains a weak 2-cluster chain, then it must contain a (strong) 2-cluster chain.

Definition 4.1 *A weak 2-cluster chain consists of two disjoint clusters X' and Y', and a set $\mathcal{P} = \mathcal{P}_1 \cup \mathcal{P}_2$ of node-disjoint paths, such that:*

- *$T \cap X', T \cap Y' = \emptyset$;*
- *X' has the (k'', α^*)-bandwidth property, and Y' has the (k'', α^*)-bandwidth property in G; and*
- *$|\mathcal{P}_1| = |\mathcal{P}_2| = 2k'$; paths in $\mathcal{P}_1$ connect vertices in T_1 to vertices in X', and paths in $\mathcal{P}_2$ connect vertices in T_1 to vertices in Y'. Moreover, the paths in $\mathcal{P}$ are internally disjoint from $X' \cup Y'$.*

We will refer to the 2-cluster chain defined in Section 3 as a strong 2-cluster chain from now on. In the next theorem we show how to obtain a strong 2-cluster chain from a weak one. The proof is omitted from this extended abstract.

Theorem 4.1 *If G contains a weak 2-cluster chain $(X', Y', \mathcal{P})$, then it contains a strong 2-cluster chain.*

4.2 Good Clusters and Perfect Clusters

Definition 4.2 *Let $C \subseteq V(G) \setminus T$ be any cluster containing non-terminals vertices only, and let (A, B) be the minimum 1/4-balanced cut in $G \setminus C$ with respect to T_1. We say that C is a* good *cluster, iff $|E(A, B)| < k/28$. We say that it is a* perfect *cluster, iff $k/28 \le |E(A, B)| \le 7k/32$.*

The proof of the following theorem uses standard techniques, namely variations of the so-called well-linked decompositions, and it is omitted from this extended abstract.

Theorem 4.2 *If there is a perfect cluster $C \subsetneq V(G) \setminus T$, such that $|\operatorname{out}(C)| \le k + k' + 1$, and C has the (k'', α^*)-bandwidth property, then G contains a 2-cluster chain.*

We will also use the following simple observation, whose proof is omitted from this extended abstract.

Observation 4.1 *Let C be a good cluster, and let (C', C'') be any partition of C, where $|\operatorname{out}(C) \cap \operatorname{out}(C'')| \le 27k/80$. Then C' is a good or a perfect cluster.*

4.3 Splitting the Cluster

We are now ready to prove that G contains a 2-cluster chain. We are interested in a cluster $C \subseteq V(G) \setminus T$, with the following properties: C is a good or a perfect cluster, and it has the 1/23-bandwidth property. Among all such clusters C, let C^* be the one minimizing $|\operatorname{out}(C^*)|$, and subject to this, minimizing $|C^*|$. We note that $|\operatorname{out}(C^*)| \le k + k'$ must hold, since $V(G) \setminus T$ is a good cluster, and, from the well-linkedness properties of the terminals, it is not hard to see that it has the 1/23-bandwidth property. We need the following two claims.

Claim 4.1 *If C^* is a good cluster, then every vertex $v \in \Gamma(C^*)$ is incident on exactly one edge of $\operatorname{out}(C^*)$.*

Proof. Assume otherwise, and let $v \in \Gamma(C^*)$ be incident on more than one edge of $\operatorname{out}(C^*)$. Since maximum vertex degree in G is 3, and $G[C^*]$ is connected, due to the 1/23-bandwidth property of C^*, v has exactly one neighbor $u \in C^*$. Consider the cluster $C' = C^* \setminus \{v\}$. Then $|\operatorname{out}(C')| < |\operatorname{out}(C^*)|$, and it is easy to see that C' has the 1/23-bandwidth property. Moreover, from Observation 4.1, C' is a good or a perfect cluster, contradicting the choice of C^*. $\square$

Claim 4.2 $p(C^*) = |\Gamma(C^*)|$.

Proof. Assume otherwise. Intuitively, if $p(C^*) < |\Gamma(C^*)|$, then there is a small cut separating $\Gamma(C^*)$ from the terminals in T. We use this cut to define a new cluster C', such that C' is either a good or a perfect cluster, and it has the 1/23-bandwidth property, while $|\operatorname{out}(C')| < |\operatorname{out}(C^*)|$, contradicting the choice of C^*.
Let $p = p(C^*)$. From Menger's theorem, there is a tripartition (X, Y, Z) of $V(G)$, such that $|Y| = p$, Y separates X from Z in G, $C^* \subseteq Y \cup Z$, and $T \subseteq X \cup Y$. Among all such tri-partitions, we choose the one minimizing $|Y| + |Z|$. As each terminal has degree 1 in G, it is easy to see that $Y \cap T = \emptyset$, and so $T \subseteq X$. Recall that $\mathcal{P}(C^*)$ is the largest-cardinality set of node-disjoint paths connecting the terminals of T to C^*, and the paths in $\mathcal{P}(C^*)$ are internally disjoint from C^*. Therefore, Y contains exactly one vertex from each path in $\mathcal{P}(C^*)$, and so $Y \cap C^* \subseteq \Gamma(C^*)$. We let C' be the set of vertices of the connected component of $G[Y \cup Z]$, containing C^*. Notice that $\Gamma(C') \subseteq Y$, and so $|\Gamma(C')| < |\Gamma(C^*)|$. Since $C^* \subseteq C'$, and C^* is a good or a perfect cluster, it is easy to see that C' is also a good or a perfect cluster. Moreover, the set $\mathcal{P}(C^*)$ of paths defines a collection $\mathcal{P}'$ of node-disjoint paths, connecting every vertex of $\Gamma(C')$ to some vertex in $\Gamma(C^*)$, such that the paths in $\mathcal{P}'$ are internally disjoint from C^*. Using the fact that C^* has the 1/23-bandwidth property, it is easy to see that C' also has the 1/23-bandwidth property.
In order to reach a contradiction, it is now enough to show that $|\operatorname{out}(C')| < |\operatorname{out}(C^*)|$. We partition the edges of $\operatorname{out}(C')$ into two subsets: set E_1 contains all edges incident on the vertices of $Y \cap \Gamma(C^*)$, and set E_2 contains all remaining edges. Similarly, we partition the edges of $\operatorname{out}(C^*)$ into two subsets: set E_1' contains all edges incident on the vertices of $Y \cap \Gamma(C^*)$, and set E_2 contains all remaining edges. Observe first that the edges of E_1 and E_1' are incident on the same subset of vertices: $Y \cap \Gamma(C^*)$, and, since $C^* \subseteq C'$, it is easy to see that $|E_1| \le |E_1'|$.

Let $Y' = Y \cap \Gamma(C^*)$. Since $|\Gamma(C')| < |\Gamma(C^*)|$, and $Y' \subseteq \Gamma(C')$, $|\Gamma(C') \setminus Y'| < |\Gamma(C^*) \setminus Y'|$. Every vertex of $\Gamma(C^*) \setminus Y'$ has at least one edge incident to it in E_2', and every edge of E_2 is incident on some vertex of $\Gamma(C') \setminus Y'$. Therefore, it is enough to show that every vertex in $\Gamma(C') \setminus Y'$ is incident on exactly one edge of E_2. Assume otherwise, and let $v \in \Gamma(C') \setminus Y'$ be any vertex incident on at least two edges of E_2. Since $v \notin Y'$, it does not belong to $\Gamma(C^*)$, or to C^*. Moreover, v has at most one neighbor in Z - denote it by u. Therefore, we can obtain a new tri-partition $(X \cup \{v\}, (Y \setminus \{v\}) \cup \{u\}, Z \setminus \{u\})$ separating the terminals from C^*, contradicting the choice of the partition (X, Y, Z). We conclude that $|E_2| \leq |\Gamma(C') \setminus Y'| < |\Gamma(C^*) \setminus Y'| \leq |E_2'|$, and $|\operatorname{out}(C')| < |\operatorname{out}(C^*)|$, contradicting the choice of C^*. □

If C^* is a perfect cluster, then, from Theorem 4.2, we obtain a 2-cluster chain in G. Therefore, we assume from now on that C^* is a good cluster. We use the following theorem to finish our proof. Its statement is slightly stronger than what we need, but this stronger statement will be used in the proof itself.

Theorem 4.3 *Let $C \subseteq V(G) \setminus T$ be any good cluster with $|\operatorname{out}(C)| \leq k + k' + 1$, and $p(C) \geq |\Gamma(C)| - 1$, such that C has the $1/23$-bandwidth property, and every vertex of $\Gamma(C)$ is incident on exactly one edge of $\operatorname{out}(C)$. Then either there is a strong 2-cluster chain in G, or there is a good or a perfect cluster $C' \subsetneq C$, with $|\operatorname{out}(C')| \leq |\operatorname{out}(C)|$, such that C' has the $1/23$-bandwidth property.*

From Theorem 4.3, either G has a strong 2-cluster chain, or there is a good cluster $C' \subsetneq C^*$ with $|\operatorname{out}(C')| \leq |\operatorname{out}(C)|$, such that C' has the $1/23$-bandwidth property. The latter is impossible from the definition of C^*, so G must contain a 2-cluster chain. From now on we focus on proving Theorem 4.3.

Proof of Theorem 4.3

We start with the following two theorems, whose proofs use standard techniques, and are omitted from this extended abstract.

Theorem 4.4 *If there is a good cluster $C' \subseteq C$, with $|\operatorname{out}(C')| \leq k + k' + 1$ and $|\Gamma(C')| \leq 7k/8$, such that C' has the (k'', α^*)-bandwidth property, then there is a weak 2-cluster chain in G.*

Theorem 4.5 *Suppose there is some value ρ, such that $\rho|\Gamma(C)| \leq (k + k' + 1)/4$, and a minimum ρ-balanced cut (A, B) of C with respect to $\Gamma(C)$, such that $|\Gamma(A) \cap \Gamma(C)| \geq |\Gamma(B) \cap \Gamma(C)| \geq 27k/80$. Then there is a cluster $C' \subseteq B$ that has the (k'', α^*)-bandwidth property, and $|\Gamma(C') \cap \Gamma(C)| > 3k/64$.*

If $|\Gamma(C)| \leq 7k/8$, then from Theorem 4.4 there is a 2-cluster chain in G. We assume from now on that $|\Gamma(C)| > 7k/8$. Let α be the largest value for which C has the α-bandwidth property, so $\alpha \geq 1/23$. We distinguish between three cases. The first case is when $\alpha < 1/5$; the second case is when $\alpha \geq 1/5$ but C does not have the $(k/4, 1)$-bandwidth property, and the third case is when C has the $(k/4, 1)$-bandwidth property.

Case 1: $\alpha < 1/5$.
Let (Z, Z') be the minimum $1/4$-balanced cut of C with respect to $\Gamma(C)$, where $|Z \cap \Gamma(C)| \geq |Z' \cap \Gamma(C)|$. We consider three sub-cases.
Subcase 1a. This first sub-case happens if $|E(Z, Z')| > \frac{\alpha}{1-\alpha}|\Gamma(C)|$. Observe that for $0 < \alpha' < 1$, function $\frac{\alpha'}{1-\alpha'}$ monotonously increases in α'. So there is some value $\alpha < \alpha' < 1$, such that $|E(Z, Z')| > \frac{\alpha'}{1-\alpha'}|\Gamma(C)|$. The following lemma uses standard techniques, and its proof is omitted from this extended abstract.

Lemma 4.2 *There is a cluster $C' \subsetneq C$, such that $|\operatorname{out}(C')| < |\operatorname{out}(C)|$, $|\Gamma(C) \cap \Gamma(C')| \geq 3|\Gamma(C)|/4$, and C' has the α'-bandwidth property.*

Let C' be the cluster given by Lemma 4.2, and let $C'' = C \setminus C'$. Then $|\Gamma(C'') \cap \Gamma(C)| \leq |\Gamma(C)|/4 \leq (k + k' + 1)/4$. Since every vertex of $\Gamma(C)$ is incident on exactly one edge of $\operatorname{out}(C)$, we get that $|\operatorname{out}(C'') \cap \operatorname{out}(C)| \leq |\Gamma(C'') \cap \Gamma(C)| \leq (k + k' + 1)/4 \leq 27k/80$. Therefore, from Observation 4.1, C' is a good or a perfect cluster. Since $\alpha' > \alpha \geq 1/23$, C' is a valid output for the theorem. Notice that $|\operatorname{out}(C')| < |\operatorname{out}(C)|$.
Subcase 1b. This case happens if $|E(Z, Z')| \leq \frac{\alpha}{1-\alpha}|\Gamma(C)|$, and Z is a good or a perfect cluster. In this case, from Lemma 2.2, cluster Z has the α'-bandwidth property, for $\alpha' = \frac{\alpha}{2+\alpha}$. Moreover, since $\alpha < 1/5$, $|E(Z, Z')| < |\Gamma(C)|/4 \leq |\Gamma(C) \cap Z'|$, and so $|\operatorname{out}(Z)| < |\operatorname{out}(C)| \leq k + k'$.
If $\alpha \geq 1/11$, then $\alpha' = \alpha/(2+\alpha) \geq 1/23$, and we return $C' = Z$. Otherwise, if $\alpha < 1/11$, then $\alpha' = \alpha/(2+\alpha) \geq 1/64$, since $\alpha \geq 1/23$, and:

$$|\Gamma(Z)| \leq \frac{3|\Gamma(C)|}{4} + |E(Z, Z')| \leq \frac{3|\Gamma(C)|}{4} + \frac{\alpha}{1-\alpha}|\Gamma(C)| \leq \frac{17|\Gamma(C)|}{20} \leq \frac{17(k + k' + 1)}{20} \leq \frac{7k}{8},$$

since $k' = k/64$. If Z is a perfect cluster, then from Theorem 4.2, we obtain a 2-cluster chain. Otherwise, Z is a good cluster, and we obtain the 2-cluster chain by applying Theorem 4.4 to cluster Z.
Subcase 1c. This case happens if $|E(Z, Z')| \leq \frac{\alpha}{1-\alpha}|\Gamma(C)|$, but Z is not a good or a perfect cluster. From Observation 4.1, $|\operatorname{out}(Z') \cap \operatorname{out}(C)| \geq 27k/80$ must hold, and, since every vertex of $\Gamma(C)$ is incident on exactly one edge of $\operatorname{out}(C)$, we get that $|\Gamma(Z') \cap \Gamma(C)| \geq |\operatorname{out}(Z') \cap \operatorname{out}(C)| \geq 27k/80$.
In this case, we construct a 2-cluster chain in G. We let $X' = Z$. Since C has the $\alpha \geq 1/23$-bandwidth property, from Lemma 2.2, X' has the α'-bandwidth property, for $\alpha' = \alpha/(2 + \alpha) \geq \alpha^*$. Therefore, X' has the (k'', α^*)-bandwidth property. We then apply Theorem 4.5 to the partition (Z, Z'), to obtain a cluster $C' \subseteq Z'$, that has the (k'', α^*)-bandwidth property, and we set $Y' = C'$. In order to define the sets $\mathcal{P}_1$ and $\mathcal{P}_2$ of paths, observe that all but one vertices of $\Gamma(C)$ have a path of $\mathcal{P}(C)$ terminating at them, since $p(C) \geq |\Gamma(C)| - 1$. Since we have assumed that $|\Gamma(C)| \geq 7k/8$, $|\Gamma(Z) \cap \Gamma(C)| \geq |\Gamma(C)|/2 \geq 7k/16$. Therefore, at least $7k/16 - 1 \geq k/32 + k'$ paths of $\mathcal{P}(C)$ terminate at the vertices of $\Gamma(Z)$, and all paths in $\mathcal{P}(C)$ are internally disjoint from C. We let $\mathcal{P}_1$ be any subset of $k/32$ paths terminating at the vertices of Z, that originate from the vertices of T_1. Recall that $|\Gamma(C') \cap \Gamma(C)| > 3k/64 \geq k/32 + k' + 1$.

Therefore, at least $k/32$ paths of $\mathcal{P}(C)$ originate from the vertices of T_1 and terminate at the vertices of C'. We let $\mathcal{P}_2 \subseteq \mathcal{P}(C)$ be any set of $k/32$ such paths. It is now easy to see that $(X', Y', \mathcal{P}_1, \mathcal{P}_2)$ is a valid weak 2-cluster chain.

Case 2: $\alpha \geq 1/5$*, but* C *does not have the* $(k/4, 1)$*-bandwidth property.*

We say that a partition (Z, Z') of C is a *sparse cut*, iff the following condition holds:

$$|E(Z, Z')| < \min\left\{|\Gamma(C) \cap Z|, |\Gamma(C) \cap Z'|\right\}.$$

Since we have assumed that C does not have the $(k/4, 1)$-bandwidth property, there is some sparse cut (Z, Z') of C, with $|E(Z, Z')| < k/4$. Let r be the smallest value $|E(Z, Z')|$ among all sparse cuts (Z, Z') of C, so $r < k/4$, and let $\rho' = (r+1)/|\Gamma(C)|$. Finally, let (A', B') be the minimum ρ'-balanced cut of C with respect to $\Gamma(C)$, and assume w.l.o.g. that $|\Gamma(C) \cap A'| \geq |\Gamma(C) \cap B'|$. From the above discussion, $|E(A', B')| = r < k/4$. We need the following claim, whose proof is omitted from this extended abstract.

Claim 4.3 *Set* A' *has the* $1/11$*-bandwidth property.*

We now consider two subcases. The first subcase happens when A' is a good or a perfect cluster. It is easy to see that $|\operatorname{out}(A')| < |\operatorname{out}(C)| \leq k + k'$ in this case, and we return $C' = A'$. The second subcase happens when A' is not a good or a perfect cluster.

In this case, we construct a 2-cluster chain in G, similarly to Case 1c. We let $X' = A'$. From the above discussion X' has the $1/11 \geq \alpha^*$-bandwidth property, and so it has the (k'', α^*)-bandwidth property. From Observation 4.1, $|\operatorname{out}(B') \cap \operatorname{out}(C)| \geq 27k/80$ must hold, and, since every vertex of $\Gamma(C)$ is incident on exactly one edge of $\operatorname{out}(C)$, we get that $|\Gamma(B') \cap \Gamma(C)| \geq |\operatorname{out}(B') \cap \operatorname{out}(C)| \geq 27k/80$.

We then apply Theorem 4.5 to the partition (A', B'), to obtain a cluster $C' \subseteq B'$, that has the α^*-bandwidth property, and we set $Y' = C'$. In order to define the sets $\mathcal{P}_1$ and $\mathcal{P}_2$ of paths, observe that all but one vertices of $\Gamma(C)$ have a path of $\mathcal{P}(C)$ terminating at them, since $p(C) \geq |\Gamma(C)| - 1$. Since we have assumed that $|\Gamma(C)| \geq 7k/8$, $|\Gamma(A') \cap \Gamma(C)| \geq |\Gamma(C)|/2 \geq 7k/16$. Therefore, at least $7k/16 - 1 \geq k/32 + k'$ paths of $\mathcal{P}(C)$ terminate at the vertices of $\Gamma(A')$, and all paths in $\mathcal{P}(C)$ are internally disjoint from C. We let $\mathcal{P}_1$ be any subset of $k/32$ paths terminating at the vertices of A', that originate from the vertices of in T_1. Recall that $|\Gamma(C') \cap \Gamma(C)| > 3k/64 \geq k/32 + k' + 1$. Therefore, at least $k/32$ paths of $\mathcal{P}(C)$ originate from the vertices of T_1 and terminate at the vertices of C'. We let $\mathcal{P}_2 \subseteq \mathcal{P}(C)$ be any such set of $k/32$ paths. It is now easy to see that $(X', Y', \mathcal{P}_1, \mathcal{P}_2)$ is a valid weak 2-cluster chain.

Case 3: C *has the* $(k/4, 1)$*-bandwidth property.*

Observe that in Cases 1 and 2, whenever we did not construct the 2-cluster chain, we returned a good or a perfect cluster $C' \subsetneq C$ with the $1/23$-bandwidth property, such that $|\operatorname{out}(C')| < |\operatorname{out}(C)|$. We will use this fact later.

From Lemma 4.1, $p(C) = |\Gamma(C)|$, and so $|\Gamma(C)| \leq k + k'$. Since every vertex of $\Gamma(C)$ is incident on exactly one edge of $\operatorname{out}(C)$, we get that $|\operatorname{out}(C)| \leq k + k'$. We claim that there is some vertex $v \in \Gamma(C)$, such that v is a non-separating vertex for $G[C]$. Indeed, assume otherwise, and let $\mathcal{C}$ be the set of all connected components of $G[C] \setminus \Gamma(C)$. Then there must be some component $R \in \mathcal{C}$, so that exactly one vertex $v \in \Gamma(C)$ has an edge connecting v to a vertex of R. Let $u \in V(R)$ be any vertex. Since $V(R) \cap T = \emptyset$, and v separates R from T, it is easy to see that T_1 remains 1-well-linked, and (T_1, T_2) remain $\frac{1}{2}$-linked in $G \setminus \{u\}$, contradicting the minimality of G. Let v be any vertex in $\Gamma(C)$, such that v is not a separator vertex for $G[C]$. Consider the cluster $C' = C \setminus \{v\}$.

We start by observing that C' has the $1/23$-bandwidth property, in the following claim, whose proof uses standard techniques and is omitted from this extended abstract.

Claim 4.4 *Cluster* C' *has the* $1/23$*-bandwidth property.*

From Observation 4.1, cluster C' is either a good or a perfect cluster. Moreover, it is easy to see that $|\operatorname{out}(C')| \leq |\operatorname{out}(C)| + 1 \leq k + k' + 1$, and $p(C') \geq |\Gamma(C')| - 1$ (there are at most two vertices that belong to $\Gamma(C') \setminus \Gamma(C)$ - the neighbors of v in C; we can extend the path of $\mathcal{P}(C)$ terminating at v to terminate at one of these vertices). If C' is a perfect cluster, then from Theorem 4.2, G contains a 2-cluster chain. Therefore, we assume that C' is a good cluster. We now consider three subcases.

The first sub-case happens when C' does not have the $(k/4, 1)$-bandwidth property, and every vertex of $\Gamma(C')$ is incident on exactly one edge of $\operatorname{out}(C')$. In this case, cluster C' is a valid input to Theorem 4.3, where it falls under Case 1 or Case 2. In each of these cases, we either showed that G contains a 2-cluster chain, or produced a good or a perfect cluster $C'' \subsetneq C'$, that has the $1/23$-bandwidth property, and $|\operatorname{out}(C'')| < |\operatorname{out}(C')| \leq |\operatorname{out}(C)| + 1$, so $|\operatorname{out}(C'')| \leq |\operatorname{out}(C)|$. We can then return the cluster C''.

The second sub-case happens when at least one vertex u of $\Gamma(C')$ is incident on two edges of $\operatorname{out}(C')$. In this case, we consider the cluster $C'' = C' \setminus \{u\}$. It is easy to see that C'' still has the $1/23$-bandwidth property, and from Observation 4.1, it is a good or a perfect cluster. Moreover, $C'' \subsetneq C$, and $|\operatorname{out}(C'')| < |\operatorname{out}(C')| \leq |\operatorname{out}(C)| + 1$, so $|\operatorname{out}(C'')| \leq |\operatorname{out}(C)|$. We then return cluster C''.

The third sub-case happens when every vertex of $\Gamma(C')$ is incident on exactly one edge of $\operatorname{out}(C')$, and C' has the $(k/4, 1)$-bandwidth property. In this case, from Lemma 4.1, $p(C') = |\Gamma(C')| = |\operatorname{out}(C')|$. However, since $C' \subsetneq C$, $p(C') \leq |\Gamma(C)|$ must hold. Therefore, $|\operatorname{out}(C')| \leq |\Gamma(C)| \leq |\operatorname{out}(C)|$, and we return C'.

5. CONCLUSION

In this paper we have introduced a simple framework for proving the Excluded Grid Theorem. The proof of the theorem reduces to proving Theorem 3.2, and any improvement in the parameters of the resulting 2-cluster chain will immediately improve the bounds on $f(g)$ in Theorem 1.1. Another interesting open direction is improving the parameters of Theorem 2.3. Currently, in order to obtain the $(g \times g)$-grid minor, we need to start from a path-of-sets system, whose width and height are both $\Omega(g^2)$. Can we obtain a $(g \times g)$-grid minor from a path-of-sets system of width $\Theta(g)$ (at the cost of slightly increasing its height)? If the answer to this question is positive, then this can significantly improve the upper bound on $f(g)$, possibly almost halving its exponent. We leave these possible avenues for improvements as open questions.

Acknowledgement. The author thanks Chandra Chekuri for many extensive discussions.

6. REFERENCES

[1] Matthew Andrews. Approximation algorithms for the edge-disjoint paths problem via Raecke decompositions. In *Proceedings of IEEE FOCS*, pages 277–286, 2010.

[2] Chandra Chekuri and Julia Chuzhoy. Polynomial bounds for the grid-minor theorem. In *Proceedings of the 46th Annual ACM Symposium on Theory of Computing*, STOC '14, pages 60–69, New York, NY, USA, 2014. ACM.

[3] Chandra Chekuri and Julia Chuzhoy. Degree-3 treewidth sparsifiers. In Piotr Indyk, editor, *Proceedings of the Twenty-Sixth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2015, San Diego, CA, USA, January 4-6, 2015*, pages 242–255. SIAM, 2015.

[4] Chandra Chekuri and Alina Ene. Poly-logarithmic approximation for maximum node disjoint paths with constant congestion. In *Proceedings of the Twenty-Fourth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 326–341. SIAM, 2013.

[5] Chandra Chekuri, Sanjeev Khanna, and F. Bruce Shepherd. A deletable edge lemma for general graphs. Manuscript, 2004.

[6] Chandra Chekuri, Sanjeev Khanna, and F. Bruce Shepherd. Multicommodity flow, well-linked terminals, and routing problems. In *Proc. of ACM STOC*, pages 183–192, 2005.

[7] Chandra Chekuri, Sanjeev Khanna, and F. Bruce Shepherd. The all-or-nothing multicommodity flow problem. *SIAM Journal on Computing*, 42(4):1467–1493, 2013.

[8] Chandra Chekuri and Nitish Korula. A graph reduction step preserving element-connectivity and applications. In *Proc. of ICALP*, pages 254–265, 2009.

[9] Chandra Chekuri, Guyslain Naves, and F. Bruce Shepherd. Maximum edge-disjoint paths in k-sums of graphs. In Fedor V. Fomin, Rusins Freivalds, Marta Z. Kwiatkowska, and David Peleg, editors, *Automata, Languages, and Programming - 40th International Colloquium, ICALP 2013, Riga, Latvia, July 8-12, 2013, Proceedings, Part I*, volume 7965 of *Lecture Notes in Computer Science*, pages 328–339. Springer, 2013.

[10] Julia Chuzhoy. Routing in undirected graphs with constant congestion. In *Proc. of ACM STOC*, pages 855–874, 2012.

[11] Julia Chuzhoy and Shi Li. A polylogarithimic approximation algorithm for edge-disjoint paths with congestion 2. In *Proc. of IEEE FOCS*, 2012.

[12] E.D. Demaine and M Hajiaghayi. The Bidimensionality Theory and Its Algorithmic Applications. *The Computer Journal*, 51(3):292–302, November 2007.

[13] Erik Demaine, MohammadTaghi Hajiaghayi, and Ken-ichi Kawarabayashi. Algorithmic graph minor theory: Improved grid minor bounds and Wagner's contraction. *Algorithmica*, 54:142–180, 2009.

[14] Erik D Demaine and MohammadTaghi Hajiaghayi. Quickly deciding minor-closed parameters in general graphs. *European Journal of Combinatorics*, 28(1):311–314, January 2007.

[15] Reinhard Diestel. *Graph Theory, 4th Edition*, volume 173 of *Graduate texts in mathematics*. Springer, 2012.

[16] Reinhard Diestel, Tommy R. Jensen, Konstantin Yu. Gorbunov, and Carsten Thomassen. Highly connected sets and the excluded grid theorem. *J. Comb. Theory, Ser. B*, 75(1):61–73, 1999.

[17] Fedor V. Fomin, Saket Saurabh, and Dimitrios M. Thilikos. Strengthening Erdos-Pósa property for minor-closed graph classes. *Journal of Graph Theory*, 66(3):235–240, 2011.

[18] H. R. Hind and O. Oellermann. Menger-type results for three or more vertices. *Congressus Numerantium*, 113:179–204, 1996.

[19] K. Kawarabayashi and Y. Kobayashi. Linear min-max relation between the treewidth of H-minor-free graphs and its largest grid minor. In *Proc. of STACS*, 2012.

[20] Rohit Khandekar, Satish Rao, and Umesh Vazirani. Graph partitioning using single commodity flows. *J. ACM*, 56(4):19:1–19:15, July 2009.

[21] Stephan Kreutzer and Siamak Tazari. On brambles, grid-like minors, and parameterized intractability of monadic second-order logic. In *Proc. of ACM-SIAM SODA*, pages 354–364, 2010.

[22] Alexander Leaf and Paul Seymour. Tree-width and planar minors. *Journal of Combinatorial Theory, Series B*, 2014.

[23] W. Mader. A reduction method for edge connectivity in graphs. *Ann. Discrete Math.*, 3:145–164, 1978.

[24] Harald Räcke. Minimizing congestion in general networks. In *Proc. of IEEE FOCS*, pages 43–52, 2002.

[25] Satish Rao and Shuheng Zhou. Edge disjoint paths in moderately connected graphs. *SIAM J. Comput.*, 39(5):1856–1887, 2010.

[26] Bruce Reed. *Surveys in Combinatorics*, chapter Treewidth and Tangles: A New Connectivity Measure and Some Applications. London Mathematical Society Lecture Note Series. Cambridge University Press, 1997.

[27] Bruce A Reed and David R Wood. Polynomial treewidth forces a large grid-like-minor. *European Journal of Combinatorics*, 33(3):374–379, April 2012.

[28] N Robertson, P Seymour, and R Thomas. Quickly Excluding a Planar Graph. *Journal of Combinatorial Theory, Series B*, 62(2):323–348, November 1994.

[29] Neil Robertson and P D Seymour. Graph minors. V. Excluding a planar graph. *Journal of Combinatorial Theory, Series B*, 41(1):92–114, August 1986.

[30] Neil Robertson and Paul D Seymour. Graph minors. XIII. the disjoint paths problem. *Journal of Combinatorial Theory, Series B*, 63(1):65–110, 1995.

[31] Mohit Singh and Lap Chi Lau. Approximating minimum bounded degree spanning trees to within one of optimal. In David S. Johnson and Uriel Feige, editors, *STOC*, pages 661–670. ACM, 2007.

[32] C Thomassen. On the presence of disjoint subgraphs of a specified type. *Journal of Graph Theory*, 12(1):101–111, 1988.