Crossing number, pair-crossing number, and expansion

PETR KOLMAN, JIŘÍ MATOUŠEK

Department of Applied Mathematics and
Institute of Theoretical Computer Science*(ITI)
Charles University
Malostranské nám. 25, 118 00 Praha 1
Czech Republic

Abstract

The *crossing number* $\mathrm{cr}(G)$ of a graph G is the minimum possible number of edge crossings in a drawing of G in the plane, while the *pair-crossing number* $\mathrm{pcr}(G)$, introduced by Pach and Tóth, is the smallest number of pairs of edges that cross in a drawing of G in the plane. While $\mathrm{cr}(G) \geq \mathrm{pcr}(G)$ holds trivially, it is not known whether a strict inequality can ever occur. We aim at bounding $\mathrm{cr}(G)$ in terms of $\mathrm{pcr}(G)$. Using the methods of Leighton and Rao, Bhatt and Leighton, and Even, Guha and Schieber, we prove that $\mathrm{cr}(G) = O(\log^3 n(\mathrm{pcr}(G) + \mathrm{ssqd}(G)))$, where $n = |V(G)|$ and $\mathrm{ssqd}(G) = \sum_{v \in V(G)} \deg_G(v)^2$. One of the main steps is an analogy of the well-known lower bound $\mathrm{cr}(G) = \Omega(b(G)^2) - O(\mathrm{ssqd}(G))$, where $b(G)$ is the *bisection width* of G, that is, the smallest number of edges that have to be removed so that no component of the resulting graph has more than $\frac{2}{3}n$ vertices. We show that $\mathrm{pcr}(G) = \Omega(b(G)^2/\log^2 n) - O(\mathrm{ssqd}(G))$.

We also prove by similar methods that a graph G with $k = \mathrm{cr}(G) > C\sqrt{\mathrm{ssqd}(G)}\, m \log^2 n$ has a nonplanar subgraph on at most $O(\frac{\Delta n m \log^2 n}{k})$ vertices, where m is the number of edges, Δ is the maximum degree in G, and C is a suitable sufficiently large constant.

*Supported by the Ministry of Education of the Czech Republic as project LN00A056.

1 Introduction

By a drawing of a (multi)graph G, we mean a drawing in the plane such that every edge is represented by an arc. The arcs are allowed to cross, but they may not pass through vertices (except for their endpoints) and no point is an internal point of three or more arcs. A *crossing* is a common internal point of two arcs.

The *crossing number* $\mathrm{cr}(G)$ is the minimum possible number of crossings in a drawing of G. The *pair-crossing number* $\mathrm{pcr}(G)$ is the minimum possible number of (unordered) pairs of edges that cross in a drawing of G. Pach and Tóth [13] were the first to formulate explicitly the definition of $\mathrm{pcr}(G)$, and they raised the problem of whether $\mathrm{cr}(G) = \mathrm{pcr}(G)$ for all G, which had previously been overlooked in papers on the crossing number of graphs.

Surprisingly, this problem appears quite challenging. A natural approach to proving equality is, given a drawing witnessing $\mathrm{pcr}(G)$, to modify it locally so that multiple crossings of pairs of edges are eliminated. An example of Kratochvíl and Matoušek [5] shows that in general, given a drawing, it need not be possible to eliminate multiple crossings of pairs without introducing new crossing pairs. Namely, there is a graph G on n vertices and a drawing D_0 of G such that if D is any drawing of G for which every pair of crossing edges also crosses in D_0, then some two edges cross at least $2^{\Omega(n)}$ times in D. In this example, the drawing D_0 is not one minimizing the pair-crossing number, so it might be still possible to modify an optimal drawing locally, but at least this does not appear straightforward.

In view of these difficulties, it is natural to seek upper bounds on $\mathrm{cr}(G)$ as a function of $\mathrm{pcr}(G)$ (and possibly of other parameters of G). Pach and Tóth [13] proved a quadratic bound: $\mathrm{cr}(G) \le 2\,\mathrm{pcr}(G)^2$. They actually prove a stronger result, involving the *odd crossing number* (the minimum number of pairs of edges in a drawing that cross an odd number of times), and their proof is rather involved. Valtr [18] recently improved this bound to $\mathrm{cr}(G) = O(k^2/\log k)$ for every graph with $\mathrm{pcr}(G) = k$, with a simple proof. The crossing number and pair-crossing number for random graphs was studied by Pach, Spencer, and Tóth [11].

In the first part of this paper, we combine known techniques for bounding the crossing number of graphs, due to Leighton and Rao [8] and Bhatt and Leighton [2] (with a recent improvement by Even et al. [3]), with some additional observations, and we prove an upper bound on $\mathrm{cr}(G)$ in terms of $\mathrm{pcr}(G)$, which is interesting for graphs with $\mathrm{pcr}(G)$ large compared to

$\sum_{v\in V(G)} \deg_G(v)^2$. The last quantity will appear many times in our considerations, and so we introduce the notation $\mathrm{ssqd}(G)$ for it. The letter n will denote the number of vertices of G throughout this paper.

Theorem 1 *For every graph G we have*

$$\mathrm{cr}(G) \le O\left(\log^3 n\ (\mathrm{pcr}(G) + \mathrm{ssqd}(G))\right).$$

In particular, if G has maximum degree bounded by a constant and pair-crossing number at least n, then $\mathrm{cr}(G) = O(\mathrm{pcr}(G)\log^3(\mathrm{pcr}(G)))$.

The main step in the proof is a nontrivial lower bound on the pair-crossing number. Several methods are known for bounding below the crossing number of a given graph; see Shahrokhi et al. [16] for a survey. The well-known lower bound in terms of the number of edges,

$$\mathrm{cr}(G) \ge \Omega\left(\frac{m^3}{n^2}\right) \tag{1}$$

for all G with n vertices and $m \ge 4n$ edges, proved by Ajtai, Chvátal, Newborn, and Szemerédi [1] and independently by Leighton [7], is also valid for the pair-crossing number, as is easily checked.

Another important lower bound is

$$\mathrm{cr}(G) \ge \Omega(b(G)^2) - O(\mathrm{ssqd}(G)), \tag{2}$$

where $b(G)$ denotes the *bisection width* of G, that is, the smallest number of edges between V_1 and V_2, where (V_1, V_2) is a partition of $V(G)$ with $|V_1|, |V_2| \ge \frac{1}{3}|V(G)|$. This bound was proved by Leighton [6] for graphs of bounded degree and by Pach, Shahrokhi, and Szegedy [10], and independently by Sýkora and Vrťo [17], for general graphs. The usual proof fails miserably if one tries to replace the crossing number by the pair-crossing number: In the first step of the proof, one considers a drawing with the minimum crossing number and replaces every crossing by a new vertex of degree 4, obtaining a planar graph and applying a separator theorem. For the pair-crossing number, we have almost no control over the total number of crossings (and thus the size of the resulting planar graph). However, the following weaker substitute of the lower bound (2) can be proved for the pair-crossing number using a low-congestion path embeddings [7, 8]:

Theorem 2 *For every graph G, we have*

$$\mathrm{pcr}(G) \geq \Omega\left(\frac{b(G)^2}{\log^2 n}\right) - O(\mathrm{ssqd}(G)).$$

This almost solves (up to the $\log^2 n$ factor) Problem 11 of Pach and Tóth [12].

A related problem is to find an optimal drawing of G in the plane, with respect to $\mathrm{cr}(G)$ or $\mathrm{pcr}(G)$. The best known algorithm is by Even et al. [3] and it approximates $\mathrm{cr}(G)+n$ (not just $\mathrm{cr}(G)$!) within a multiplicative factor of $O(\log^3 n)$. The procedure is to recursively draw G on a circle arc, that is, to put vertices on the arc and to draw all edges as straight line segments. Their bound is an improvement of an earlier result by Bhatt and Leighton [2]. A corollary of the analysis is that for any graph G, there exists a drawing of G on the circle arc with at most $O(\log n(\mathrm{cr}(G)+n))$ crossings. Although we are not concerned about drawing algorithms, the outlined procedure will be used in our proofs.

In Section 5 we investigate "locality" of the crossing number. That is, if $\mathrm{cr}(G)$ is large, must G necessarily have small nonplanar subgraphs? By a detour via edge expansion, in analogy to the preceding section, we prove the following upper bound on the size of a nonplanar subgraph:

Theorem 3 *Let G be a graph with crossing number $\mathrm{cr}(G) > C\sqrt{\mathrm{ssqd}(G)}\, m \log^2 n$, where m is the number of edges in G and C is a sufficiently large absolute constant. Then G has a nonplanar subgraph on at most*

$$O\left(\frac{\Delta m n \log^2 n}{\mathrm{cr}(G)}\right)$$

vertices, where Δ is the maximum degree in G.

In particular, if the maximum degree Δ is bounded by a constant, the assumption becomes $\mathrm{cr}(G) > C' n \log^2 n$, and the bound for the size of the nonplanar subgraph becomes $O((n \log n)^2/\mathrm{cr}(G))$.

For graphs with maximum degree bounded by a constant and with $k = \Omega(n^2)$, this result is nearly optimal, up to a factor of $\log n$. Namely, a constant-degree expander of girth $g = \Omega(\log n)$ (i.e., with minimal length of a cycle $\Omega(\log n)$) has crossing number $\Omega(n^2)$, and all subgraphs on fewer than g vertices are planar (even trees).

2 Preliminaries

An (edge) *cut* $e(V_1, V_2)$ of a graph G is the set of edges connecting V_1 and V_2, for $V_1 \cup V_2 = V$ and $V_1 \cap V_2 = \emptyset$. The pair (V_1, V_2) is called a *partition* of G, and its *size* is the number of edges in the cut $e(V_1, V_2)$.

The *bisection width* $b(G)$ was introduced in the previous section, as the size of a minimal partition (V_1, V_2) with $|V_1|, |V_2| \geq \frac{1}{3}|V(G)|$. (Note that we do not insist on partitioning the vertices into two parts of equal size; we consider an *approximate* bisection.) The *hereditary bisection width* $\mathrm{hb}(G)$ is the maximum of $b(H)$ over all subgraphs H of G. The *edge expansion* of G is

$$\beta(G) = \min_{A \subseteq V} \frac{e(A, V \setminus A)}{\min\{|A|, |V \setminus A|\}}.$$

An *embedding* of a graph H in a graph G maps vertices of H to vertices of G and edges of H to paths in G. More formally, an embedding is a pair (f, φ), where $f\colon V(H) \to V(G)$ is an injective mapping, and φ is a mapping that assigns to each edge $e = \{u, v\} \in E(H)$ a path $\varphi(e)$ in G connecting the vertices $f(u)$ and $f(v)$. The *congestion* of the embedding is the maximum number of paths in the embedding passing through an edge of G, and the *dilation* is the maximum length of a path $\varphi(e)$, $e \in E(H)$.

The following theorem is one of our main tools; it will be used in the proof of both Theorem 1 (relating the crossing number and the pair-crossing number; the bound on the dilation is not needed here) and Theorem 3.

Theorem 4 *Let G be a graph on n vertices with edge expansion β and maximum degree Δ. Then there exists an embedding of the complete graph K_n in G with congestion $O(\beta^{-1} n \log n)$ and dilation $O(\Delta\beta^{-1} \log n)$.*

As a tool for proving this theorem, we will use *concurrent multicommodity flows*, namely a *uniform* multicommodity flow: there is a commodity with demand one for each (unordered) pair of vertices. A *feasible solution* of such a multicommodity flow problem is a system of flows in G, one flow for every pair of vertices, with the total flow through each edge at most one. The *flow* of the feasible solution is the maximum f such that at least f units are transfered for each commodity. The objective is to find a feasible solution with a maximal flow, called the *max-flow* of the problem.

The *min-cut* of the uniform multicommodity flow problem is

$$\varphi = \min_{A \subseteq V} \frac{e(A, V \setminus A)}{|A| \cdot |V \setminus A|}.$$

Observe that φ depends only on the graph G and that it is closely related to the expansion β of the graph:

$$\frac{1}{n} \cdot \beta \leq \varphi \leq \frac{2}{n} \cdot \beta. \tag{3}$$

Proof of Theorem 4. Leighton and Rao [8, Theorem 18] proved that on any graph G there exists a solution of the uniform multicommodity flow problem with flow of size $\Omega(\varphi/\log n)$, for which every flow path has length at most $O(\Delta\beta^{-1}\log n)$, where Δ is the maximum degree of G. (Later, Kolman and Scheideler proved an analogous result for a general multicommodity flow.)

We consider this solution, and we individually rescale each flow so that one unit flows between every pair of vertices. The largest scaling factor is $O(\varphi^{-1}\log n)$, and so the total flow through each edge after the rescaling is $O(\varphi^{-1}\log n) = O(\beta^{-1}n\log n)$ (using the relation (3) between φ and β).

The flow between every pair of vertices is at least one, all flow paths have the desired length, and also the maximal flow through an edge is as desired. It remains to turn each of the unit-capacity flows into a path (that is, to make the flows integral). Observing that $\beta^{-1}n\log n \geq \log n$, this can be accomplished by the randomized rounding of Raghavan and Thompson [14, Theorem 3.1], which increases the maximal flow through an edge only by another constant factor. $\square$

In the proof of Theorem 4 we first produced unit-capacity flows, and then we turned them into paths by randomized rounding. Let us remark that this rounding step is not essential for the forthcoming proofs. The integrality of the flows only simplifies some later arguments but it is not crucial for them.

Remark. For some classes of graphs, the bounds in Theorem 4 can be improved. A useful parameter, which to some extent measures the possibility of such an improvement, is the *flow number* $F = F(G)$ [4]. Let I_0 denote the instance of the concurrent multicommodity flow problem in which there is a commodity with demand $\deg(u) \cdot \deg(v)/2|E(G)|$ for each pair of vertices (u, v). For a feasible solution $\mathcal{S}$, let $D(\mathcal{S})$ be the length of the longest flow path in $\mathcal{S}$ and let $C(\mathcal{S})$ be the inverse of the flow (i.e., the maximum over all commodities of flow divided by demand) of $\mathcal{S}$. Then $F(G)$ is the minimum of $\max\{D(\mathcal{S}), C(\mathcal{S})\}$ over all feasible solutions $\mathcal{S}$ of I_0. The congestion bound in Theorem 4 can be replaced by $O(nF)$ and the dilation bound by $O(F)$ (cf. [4]). We always have $F = O(\Delta\beta^{-1}\log n)$, where Δ is the maximum degree of G, but sometimes F can be smaller by a factor Δ or $\log n$. For example, for the hypercube on $n = 2^m$ vertices we have $F = O(\log n)$, and for the 2-dimensional $\sqrt{n}\times\sqrt{n}$ mesh we have $F = O(\sqrt{n})$.

3 Pair-crossing number and bisection width: Proof of Theorem 2

We begin with a simple lemma showing that a graph with large bisection width contains a large subgraph with large expansion.

Lemma 5 *Every graph on n vertices with bisection width b contains a subgraph on at least $\frac{2}{3}n$ vertices with edge expansion at least b/n.*

Proof. If $\beta(G) \geq \frac{b}{n}$, we are done, and otherwise, there is a subset A_1, $1 \leq |A_1| \leq \frac{n}{2}$, that can be cut off by removing at most $|A_1|\frac{b}{n}$ edges. Moreover, $|A_1| < \frac{n}{3}$, for otherwise, we would get a contradiction to $b(G) = b$. We look at the subgraph induced by $V \setminus A_1$; if it has edge expansion at least $\frac{b}{n}$, we can finish, and otherwise, we can cut off a subset A_2, etc.

At each step, we cut off at most half of the current number of vertices, and so if we do not finish earlier, we must reach a situation when the current graph, induced by $V \setminus (A_1 \cup A_2 \cup \cdots \cup A_k)$, has between $\frac{n}{3}$ and $\frac{2}{3}n$ vertices. The set $A_1 \cup \cdots \cup A_k$ can be separated from this subgraph by removing at most $(|A_1|+\cdots+|A_k|)\frac{b}{n} < b$ edges. This contradiction shows that we obtain the desired subgraph in some of the earlier steps. $\square$

To prove Theorem 2, it suffices, by the lemma just proved, to show that

$$\mathrm{pcr}(G) \geq \Omega\left(\frac{n^2\beta(G)^2}{\log^2 n}\right) - O(\mathrm{ssqd}(G)). \tag{4}$$

Proof of Theorem 2. By Theorem 4, there exists a set of paths $\mathcal{P}$ such that

- for each pair $\{u, v\} \in \binom{V}{2}$ there is a path $p_{uv} \in \mathcal{P}$ connecting u and v, and
- for each edge $e \in E$, there are at most $O(\beta^{-1} n \log n)$ paths of $\mathcal{P}$ going through it.

Let us fix a drawing of G witnessing $\mathrm{pcr}(G)$. Using the paths from $\mathcal{P}$, we draw the complete graph K_V on the vertex set V: The edge $\{u, v\}$ of K_V is drawn along the path p_{uv}.

Crossings in this drawing of K_V come from crossings in the drawing of G and from crossings near vertices of G. The number of crossing pairs in the drawing of K_V induced by a crossing pair (e_1, e_2) of edges of G is at most $O(\beta^{-2}n^2 \log^2 n)$, and the number of crossing pairs caused by crossings of the paths from $\mathcal{P}$ near a vertex w is at most $O(\deg^2(w)\beta^{-2}n^2 \log^2 n)$. Thus,

$$\mathrm{pcr}(K_V) \leq O\left(\frac{n^2 \log^2 n}{\beta^2}(\mathrm{pcr}(G) + \mathrm{ssqd}(G))\right).$$

Since $\mathrm{pcr}(K_V) = \Omega(n^4)$, for example by (1), the proof of (4), and thus also of the theorem, is completed. $\square$

As was noted after the proof of Theorem 4, the rounding of the (non-integral) multicommodity flow to an integral one (i.e., a system of paths) can easily be avoided. Given arbitrary unit flows as in Theorem 4, we can again draw K_V using the optimal drawing of G. In this case, the edge $\{u, v\}$ of K_V is drawn along a path that is chosen at random from all paths that constitute the flow between u and v in the solution, with the random choice made according to the sizes of flows along these paths. Then we can estimate the expected pair-crossing number of the resulting drawing of K_V and compare it with the pair crossing number of K_V.

4 Drawing by recursive bisection: Proof of Theorem 1

We start with a proof of a slightly weaker version of Theorem 1:

$$\mathrm{cr}(G) \leq O\left(\log^4 n(\mathrm{pcr}(G) + \mathrm{ssqd}(G))\right). \tag{5}$$

We follow the procedure of Bhatt and Leighton [2] for drawing G, in a slightly simplified form similar to the version in Shahrokhi et al. [15]. The procedure is recursive. It places the vertices of a given graph on a circle arc, and the edges are drawn as straight segments.

For a given graph G, the procedure finds a bisection (V_1, V_2) of G witnessing $b(G)$, divides the given arc into two subarcs, and recursively places the vertices of G_1 on one of the arcs and the vertices of G_2 on the other arc (here G_i is the subgraph induced by V_i).

Let $\ell(G)$ denote the maximum number of edges going "over" a vertex in the resulting drawing of G (an edge $\{u, v\}$ is said to go over a vertex w if u and v lie on the arc on opposite sides of w). We have the recurrence

$$\ell(G) \leq b(G) + \max(\ell(G_1), \ell(G_2)),$$

and induction then shows that $\ell(G) \leq C_1 \mathrm{hb}(G) \log n$, where C_1 is a suitable constant.

Now we can prove $\mathrm{cr}(G) \leq C \log^4 n(\mathrm{pcr}(G) + \mathrm{ssqd}(G))$, for a suitable constant C, by induction on n. Using the drawing produced by the algorithm, we obtain

$$\mathrm{cr}(G) \leq \binom{b(G)}{2} + b(G) \cdot (\ell(G_1) + \ell(G_2)) + \mathrm{cr}(G_1) + \mathrm{cr}(G_2).$$

By estimating $\ell(G_1)$ and $\ell(G_2)$ as above and using the induction hypothesis for G_1 and G_2, we have

$$\begin{aligned} \mathrm{cr}(G) &\leq C_1 \mathrm{hb}(G)^2 \log n + \\ &\qquad C(\log(\tfrac{2}{3}n))^4 \big(\mathrm{pcr}(G_1) + \mathrm{pcr}(G_2) + \mathrm{ssqd}(G_1) + \mathrm{ssqd}(G_2)\big) \\ &\leq C_1 \mathrm{hb}(G)^2 \log n + C(\log n - \log \tfrac{3}{2}) \log^3 n \big(\mathrm{pcr}(G) + \mathrm{ssqd}(G)\big). \end{aligned}$$

The induction step is finished by using $\mathrm{hb}(G)^2 \leq C_2 \log^2 n(\mathrm{pcr}(G)+\mathrm{ssqd}(G))$ from Theorem 2. This completes the proof of the weaker bound (5). □

The stronger bound in Theorem 1 is again based on the recursive drawing of the graph on the circle arc, with two additional ideas. The first idea is to better split the graph into two parts: Rather then partitioning the graph into two parts of approximately the same *size*, it is more appropriate to partition the graph into two parts with approximately equal *pair-crossing number* (see Lemma 6 below). The other improvement is a better method for counting the crossings in the recursive drawing of G, based on a recent algorithm for crossing number approximation by Even et al. [3]. Even et al. actually improved the analysis of the recursive procedure of Bhatt and Leighton for drawing G.

Lemma 6 *For every graph G on n vertices there exists a partition (V_1, V_2) of $V(G)$ with size $e(V_1, V_2) = O\left(\log n \sqrt{\operatorname{pcr}(G) + \operatorname{ssqd}(G)}\right)$ such that for $i = 1, 2$,*

$$\operatorname{pcr}(G_i) + \operatorname{ssqd}(G_i) \leq \frac{2}{3}(\operatorname{pcr}(G) + \operatorname{ssqd}(G)),$$

where G_i is the subgraph of G induced by V_i.

Proof. The idea is to transform the given graph G into a new graph G' in such a way that the number of vertices in G' captures both $\operatorname{ssqd}(G)$ and $\operatorname{pcr}(G)$. Then we get the desired partition of G by applying Theorem 2 to the new graph G'.

We consider a drawing of G optimal with respect to the pair crossing number, and for an edge e, let $p(e)$ denote the number of pair crossings of this edge. We set a weight $w(e)$ to $p(e)/2$, for every edge. For every vertex $u \in V$, we increase the weight of every adjacent edge but one by $\deg(v)$, and the weight of the last adjacent edge is increased only by $\deg(v) - 1$. We get the new graph G' by replacing every edge $e = \{u, v\}$ by a path of length $w(e)-1$ (in other words, we add roughly $\deg(u)+\deg(v)+p(e)/2$ new vertices on an edge $e = \{u, v\}$). It is easy to see that $\operatorname{ssqd}(G') \leq 5 \operatorname{ssqd}(G) + 4 \operatorname{pcr}(G)$, $|V(G')| = \operatorname{pcr}(G) + \operatorname{ssqd}(G)$, $\log n' = O(\log n)$, and moreover, we can add the new vertices in such a way that $\operatorname{pcr}(G') = \operatorname{pcr}(G)$. Also, a bisection of G' of size m naturally induces a partition of G of size m.

For a subgraph G'_i of G', let V_i denote the set of original vertices in G'_i (that is, those coming from G), and let G_i be the subgraph of G induced by V_i. The important observation is that the number of vertices in G'_i is an upper bound on $\operatorname{pcr}(G_i) + \operatorname{ssqd}(G_i)$. By Theorem 2 there is a bisection of size

$$s = O\left(\log n' \sqrt{\operatorname{pcr}(G') + \operatorname{ssqd}(G')}\right)$$

that separates G' into G'_1 and G'_2, and thus we have a partition (V_1, V_2) of G of size $O(s)$ such that $\operatorname{pcr}(G_i) + \operatorname{ssqd}(G_i) \leq \frac{2}{3}(\operatorname{pcr}(G) + \operatorname{ssqd}(G))$, for $i = 1, 2$. $\square$

Proof of Theorem 1. Let T denote a binary tree representing a recursive decomposition of G according to Lemma 6: The root of T corresponds to the set V, and two children of a vertex $t \in T$ associated with a set V_t correspond to the two sets $V_{t_1}, V_{t_2} \subseteq V_t$ constituting the partition of V_t given by Lemma 6.

An edge $e = \{u, v\}$ of G is *split* in a tree vertex t if $u, v \in V_t$ and $|\{u, v\} \cap V_{t_i}| = 1$ for the two children t_1, t_2 of t. Let G_t denote the subgraph of G induced by V_t, and let $n_t = |V_t|$, for a tree vertex t.

Consider the drawing on the circle arc that is based on the recursive partitioning by Lemma 6. To bound the number of edge crossings in this drawing, we charge a crossing of e and e' to the first edge among e and e' that was split first by the partitioning procedure. It is easy to observe that for any two crossing edges e, e', the tree vertex in which the edge e was split is an ancestor of the tree vertex in which e' was split, or the other way round (by definition, a vertex is an ancestor to itself). In other words, for any tree vertex t with children t_1, t_2, the edges in G_{t_1} do not cross with edges in G_{t_2}.

Observation 7 (Even et al. [3]) *Let $P(u, v)$ denote the set of vertices in T on the path from the leaf corresponding to u to the leaf corresponding to v. The number of crossings that are charged to the edge $e = \{u, v\}$ is bounded by $\sum_{t \in P(u,v)} |e(V_{t_1}, V_{t_2})|$.*

By a combination of Observation 7 and Lemma 6, an edge $e = \{u, v\}$ is charged for at most

$$\sum_{t \in P(u,v)} |e(V_{t_1}, V_{t_2})| = O\Big(\sum_{t \in P(u,v)} \log n_t \sqrt{\mathrm{pcr}(G_t) + \mathrm{ssqd}(G_t)}\Big)$$

crossings. Since the partitioning procedure guarantees an exponential decrease of $\mathrm{pcr}(G_t) + \mathrm{ssqd}(G_t)$, an *edge* e is charged for at most

$$O\left(\log n \sqrt{\mathrm{pcr}(G_t) + \mathrm{ssqd}(G_t)}\right) \tag{6}$$

crossings, where t is the vertex in which e was split. Recalling that the size of the partition of G_t is

$$O\left(\log n_t \sqrt{\mathrm{pcr}(G_t) + \mathrm{ssqd}(G_t)}\right),$$

the number of crossings for which a *tree vertex* t is charged is at most

$$O\left(\log^2 n (\mathrm{pcr}(G_t) + \mathrm{ssqd}(G_t))\right).$$

Since the tree vertices in the same layer form a partition of V and the number of layers is $O(\log n)$, all tree vertices are charged for at most $O(\log^3 n(\mathrm{pcr}(G) + \mathrm{ssqd}(G)))$ crossings. Theorem 1 is proved. □

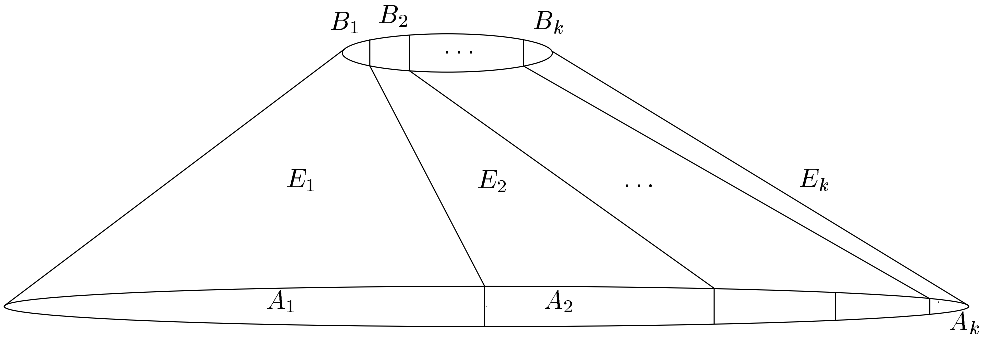

Figure 1: A bad graph for vertex-balanced partitions.

It is worth mentioning that balancing the partitions by Lemma 6, that is, with respect to the pair-crossing number as opposed to the number of vertices, is crucial in the above proof. We aim at upper bounding the number of crossings in our arc-drawing of G in terms of $\mathrm{pcr}(G)$. To do so, we rely on the relation between $\mathrm{pcr}(G)$ and $b(G)$ by Theorem 2, namely on the relation $b(G) = O(\log n\sqrt{\mathrm{pcr}(G) + \mathrm{ssqd}(G)})$. If we simply used a bisection that is balanced with respect to the number of vertices but not with respect to $\mathrm{ssqd}(G_i) + \mathrm{pcr}(G_i)$, the exponential decrease of cut sizes would not be guaranteed and the bound (6) would increase by a $\log n$ factor.

The following example demonstrates this difficulty. It shows that balancing the partitions with respect to the number of vertices while upper bounding the size of the cuts by the bound $O(\log n\sqrt{\mathrm{pcr}(G) + \mathrm{ssqd}(G)})$ can really yield a long sequence of cuts with nondecreasing size: Let k be such that $n/2^k = \sqrt{n}\log n$ (we have $k = \Theta(\log n)$). Let A_i be a set of $n/2^i$ vertices, for $i = 1, 2, \ldots, k$, and let $B_1, B_2, \ldots, B_k$ be sets of $\sqrt{n}$ vertices each. Let G_i be a bipartite graph on the sets A_i, B_i with n edges chosen in such a way that $\mathrm{pcr}(G_i) \le n^2/\log n$. Let E_i denote the edge set of G_i (see Fig. 1). Let E' be the set of edges of a complete graph on the set B_k. Consider the graph

$$G = \left(\bigcup_{i=1}^{k}(A_i \cup B_i), \bigcup_{i=1}^{k}(E_i) \cup E'\right).$$

It is easy to check that G has $\Theta(n)$ vertices and that $\mathrm{pcr}(G) = \Theta(n^2)$. We observe that E_1 defines a bisection of G, and moreover, that even $\mathrm{pcr}(G \setminus A_1) = \Theta(n^2)$. Similarly, $E_2, E_3, \cdots, E_k$ define bisections in next levels of the recursive partitioning such that one of the remaining parts of the graph still has pair crossing number $\Theta(n^2)$ and each of the partitions has size n.

5 Small nonplanar subgraphs in graphs with large crossing number

First we relate the existence of small nonplanar subgraphs to edge expansion.

Theorem 8 *Let G be a graph with edge expansion β and maximum degree Δ such that* $\mathrm{ssqd}(G)\cdot\beta^{-2}\cdot\log^2 n < c\cdot n^2$*, for a sufficiently small absolute constant $c > 0$. Then there exists a nonplanar subgraph in G of size $O(\Delta\beta^{-1}\log n)$.*

In particular, a nonplanar subgraph of size $O(\beta^{-1}\log n)$ exists in bounded degree graphs with $\beta \geq c' \cdot \log n/\sqrt{n}$, for a sufficiently large absolute constant $c' > 0$.

Proof. Let $\mathcal{P}$ be the system of paths from the embedding of K_n in G guaranteed by Theorem 4. That is, there is a path of length at most $L = O(\Delta\beta^{-1}\log n)$ between each pair of vertices in G, and the maximal number of paths passing through an edge is $C = O(\beta^{-1}n\log n)$.

Let us choose an ordered sixtuple $U = (u_1, u_2, u_3, v_1, v_2, v_3)$ of distinct vertices from $V(G)$ at random, all ordered sixtuples having the same probability. Let $\mathcal{F} = \{p_{u_i,v_j} \in \mathcal{P} : i, j = 1, 2, 3\}$. Let H be the subgraph induced by the union of these paths. Clearly, H has $O(L)$ vertices. We want to show that with a positive probability H is a nonplanar subgraph in G. An obstacle that we have to overcome is that the paths in $\mathcal{F}$ may cross at vertices and/or share edges, and thus that we need not always get a subdivision of $K_{3,3}$.

We introduce the following types of pairs of paths from $\mathcal{P}$ (see Fig. 2): A pair (p, q) is called

- *disjoint* if p and q are vertex disjoint, with the possible exception of a common terminal vertex;
- *crossing* if p and q have four different terminal vertices and they have at least one common vertex;
- *arching* if p and q have a common terminal vertex, and the other terminal vertex of one of the paths is an internal vertex of the other path; and

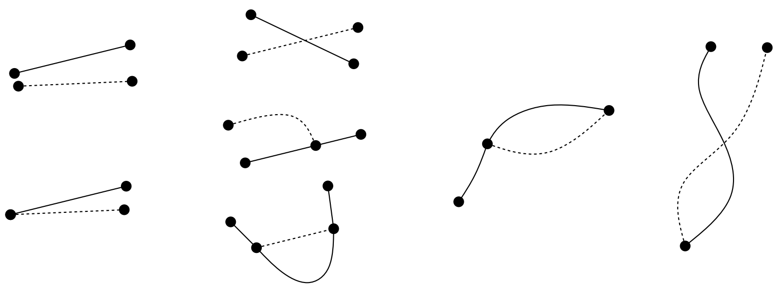

Figure 2: Types of pairs of paths.

- *tangled* if p and q have a common terminal vertex, the other terminal vertex of p does not lie on q and vice versa, and p and q cross in at least one other vertex.

We claim: *With a positive probability, there are no crossing pairs and no arching pairs in $\mathcal{F}$.*

To prove the claim, we show that the expected number of crossing and arching pairs in $\mathcal{F}$ is strictly smaller than one. The number of paths of $\mathcal{P}$ passing through a vertex v is at most $C\deg(v)$, and hence the total number of crossing pairs (p,q) with $p,q \in \mathcal{P}$ is at most $\sum_{v\in V(G)}(C\deg(v))^2 = C^2\mathrm{ssqd}(G)$. Since a fixed pair of paths with four distinct terminal vertices appears in $\mathcal{F}$ with probability $O(n^{-4})$, the expected number of crossing pairs in $\mathcal{F}$ is $O(C^2\mathrm{ssqd}(G)n^{-4}) = O(\mathrm{ssqd}(G)\cdot\beta^{-2}n^{-2}\log^2 n)$. This can be made smaller than any prescribed constant, say smaller than $1/4$, by choosing the constant c in the assumption of the theorem sufficiently small.

Next, we consider the arching pairs. To choose an arching pair, we can first select the vertex v that is terminal for one of the paths, say q, and internal for the other one, p. Then p can be chosen in at most $C\deg(v)$ ways, and there are only two possibilities of choosing q (one of the terminal vertices of q is v and the other one is one of the terminal vertices of p). Hence there are $O(C\cdot\sum_{v\in V(G)}\deg(v))$ arching pairs, and each of them has probability $O(n^{-3})$ of appearing in $\mathcal{F}$. The expected number of arching pairs in $\mathcal{F}$ is thus $O(\beta^{-1}n^{-2}\log n\sum_{v\in V(G)}\deg(v))$. Since $\beta \leq \delta(G)$, while $\mathrm{ssqd}(G) \geq \delta(G)\cdot\sum_{v\in V(G)}\deg(v))$, where $\delta(G)$ denotes the minimum degree of G, the above estimate for the expected number of arching pairs is dominated by the earlier bound for the expected number of crossing pairs. We conclude that the expected number of crossing and arching pairs is smaller than one.

We can thus choose a sixtuple U whose paths form only disjoint and tangled pairs. If there is no tangled pair in $\mathcal{F}$, then H is a subdivision of $K_{3,3}$. It remains to check that even if tangled pairs appear in $\mathcal{F}$, H still is nonplanar.

Indeed, suppose that H is planar and $\mathcal{F}$ contains tangled pairs. Consider a planar drawing of H. It defines a drawing of $K_{3,3}$: the vertices of $K_{3,3}$ are placed to the vertices of U, and each edge of $K_{3,3}$ is drawn along the corresponding path in the drawing of H. This drawing is not necessarily planar, but no two vertex-disjoint edges cross in it. But it is well known that every drawing of $K_{3,3}$ in the plane has two vertex-disjoint edges that cross (see, e.g., [9]). Hence H is nonplanar and Theorem 8 is proved. □

Remark. Similarly as in Theorem 4, the terms $\beta^{-1}\log n$ and $\Delta\beta^{-1}\log n$ can be replaced by the flow number F of G. Then, for bounded degree graphs, the condition $F < c\sqrt{n}$ guarantees a nonplanar minor of size $O(F)$ in G. In a way, this is the best possible in general: a two-dimensional $\sqrt{n}\times\sqrt{n}$ mesh is a planar graph with flow number $\Omega(\sqrt{n})$.

Lemma 9 *For every graph G,*

$$\mathrm{cr}(G) \leq 2\cdot \mathrm{hb}(G)\cdot m\cdot \log n \ ,$$

where m is the number of edges in G.

Proof. Consider the recursive drawing procedure of G on an arc: in each level use the minimal bisection to divide the current part H into H_1 and H_2 and recursively draw H_1 on one side of the arc and H_2 on the other. The depth of the recursion is at most $2\log n$, and thus each edge crosses at most $4\cdot \mathrm{hb}(G)\cdot\log n$ other edges, which sums into $2\cdot\mathrm{hb}(G)\cdot m\cdot\log n$ over all edges. □

Proof of Theorem 3. Let us consider a graph G with n vertices, m edges, and $\mathrm{cr}(G) = k$. By Lemma 9, there is a subgraph G_1 with $b(G_1) = \Omega(\frac{k}{m\log n})$. Lemma 5 then yields a subgraph G_2 on $n_2 \geq \frac{2}{3}|V(G_1)|$ vertices with edge expansion

$$\beta(G_2) = \Omega\left(\frac{b(G_1)}{n_2}\right) = \Omega\left(\frac{k}{n_2 m\log n}\right).$$

Applying Theorem 8 to G_2, we obtain a nonplanar subgraph of size

$$O\left(\frac{\Delta(G_2)mn_2\log n\cdot\log n_2}{k}\right)=O\left(\frac{\Delta mn\log^2 n}{k}\right).$$

It remains to check that the assumption of Theorem 8 holds for G_2, namely that $\mathrm{ssqd}(G_2)\beta(G_2)^{-2}\log^2 n_2 < cn_2^2$. Using the above lower bound for $\beta(G_2)$, it is sufficient to check that $\mathrm{ssqd}(G_2)m^2(\log^4 n)/k^2 < c$, and this follows from the assumption $k \geq C\sqrt{\mathrm{ssqd}(G)}\, m\log^2 n$ in Theorem 3. □

6 Open problems

An obvious open problem is to decide whether $\mathrm{cr}(G) = \mathrm{pcr}(G)$ for all G. We currently do not see any good reason why this equality should always hold, and so we believe that it makes sense to continue the investigation of upper bounds on $\mathrm{cr}(G)$ in terms of $\mathrm{pcr}(G)$.

A nice problem concerns the pair-crossing number of a constant-degree expander G. While $\mathrm{cr}(G) = \Omega(n^2)$ (since $b(G) = \Omega(n)$), the method of Theorem 2 cannot yield such lower bound for $\mathrm{pcr}(G)$, since the flows simply cannot be made sufficiently large. Still, it is very natural to conjecture that $\mathrm{pcr}(G) = \Omega(n^2)$.

In Section 5 we proved a lower bound on the edge expansion (of a subgraph) in terms of the crossing number, using the recursive drawing procedure. Although the resulting bound is almost tight in some cases (for bounded-degree graphs with quadratic crossing number, say), perhaps it can be improved for smaller crossing numbers. For example, is it true that for all k, every graph with maximum degree bounded by a constant contains a subgraph with edge expansion $\Omega(n/\sqrt{(k+n)\mathrm{polylog}\, n})$?

Finally, the crossing number is much less understood for graphs with large degrees than for those with degrees bounded by a constant. The term $\mathrm{ssqd}(G)$ appears very often in various bounds and, if some degrees are large, it usually makes the bounds uninteresting. One of the main tools for bounding the crossing number, the recursive procedure of drawing on an arc by recursive bisection, no longer works in the presence of high degrees: For example, while $K_{2,n}$ is planar, any straight-edge drawing with vertices on an arc has $\Omega(n^2)$ crossings. Further, the bisection width of $K_{2,n}$ is $\Omega(n)$, of the same order as for $K_{3,n}$; the former graph is planar, while the latter has crossing number $\Omega(n^2)$. So the bisection width is no longer suitable for estimating the crossing number. It seems that substantial new ideas are needed for, say, a good approximation algorithm for the crossing number of general graphs.

Acknowledgment

We would like to thank Robert Babilon, Helena Nyklová, Pavel Valtr, and other participants of a seminar in Prague for useful discussions. We also thank Fahrad Shahrokhi for pointing out the paper [3] to us, which led to an improvement of Theorem 1.

References

[1] M. Ajtai, V. Chvátal, M. Newborn, and E. Szemerédi. Crossing-free subgraphs. *Ann. Discrete Math.*, 12:9–12, 1982.

[2] S. N. Bhatt and F. T. Leighton. A framework for solving VLSI graph layout problems. *J. Comput. Syst. Sci.*, 28:300–343, 1984.

[3] G. Even, S. Guha, and B. Schieber. Improved approximations of crossings in graph drawings. In *Proceedings of the Thirty Second Annual ACM Symposium on Theory of Computing*, pages 296–305, 2000.

[4] P. Kolman and C. Scheideler. Improved bounds for the unsplittable flow problem. In *Proceedings of the 13th ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 184–193, 2002.

[5] J. Kratochvíl and J. Matoušek. Intersection graphs of segments. *J. Combin. Theory Ser. B*, 35(2):317–339, 1994.

[6] F. T. Leighton. *Complexity issues in VLSI.* MIT Press, Cambridge, MA, 1983.

[7] F. T. Leighton. New lower bound techniques for VLSI. *Math. Systems Theory*, 17:47–70, 1984.

[8] F. T. Leighton and S. Rao. Multicommodity max-flow min-cut theorems and their use in designing approximation algorithms. *J. Assoc. Comput. Machin.*, 46:787–832, 1999.

[9] B. Mohar and C. Thomassen. *Graphs on Surfaces.* Johns Hopkins University Press, Baltimore, MD, 2001.

[10] J. Pach, F. Shahrokhi, and M. Szegedy. Applications of the crossing number. *Algorithmica*, 16:111–117, 1996. (special issue on Graph Drawing, edited by G. Di Battista and R. Tamassia).

[11] J. Pach, J. Spencer, and G. Tóth. New bounds for crossing numbers,. *Discrete Comput. Geom.*, 24:623–644, 2000.

[12] J. Pach and G. Tóth. Thirteen problems on crossing numbers. *Geombinatorics*, 9:194–207, 2000.

[13] J. Pach and G. Tóth. Which crossing number is it anyway? *J. Combin. Theory Ser. B*, 80:225–246, 2000.

[14] P. Raghavan and C. D. Thompson. Randomized rounding: A technique for provably good algorithms and algorithmic proofs. *Combinatorica*, 7:365–374, 1987.

[15] F. Shahrokhi, L. Székely, O. Sýkora, and I. Vrťo. The book crossing number of a graph. *J. Graph Theory*, 21(4):413–424, 1996.

[16] F. Shahrokhi, L. A. Székely, and I. Vrťo. Crossing numbers: bounds and applications. In I. Bárány and K. Böröczky, editors, *Intuitive Geometry*, Bólyai Society Mathematical Studies 6, pages 179–206. Akadémiai Kiadó, Budapest, 1997.

[17] O. Sýkora and I. Vrťo. On VLSI layouts of the star graph and related networks. *Integration, The VLSI Journal*, 17(1):83–93, 1994.

[18] P. Valtr. 2002. Personal communication.