

ON 3-PUSHDOWN GRAPHS WITH LARGE SEPARATORS

Z. GALIL*, R. KANNAN and E. SZEMERÉDI

Received March 17, 1987

For an integer s let $I^s(n)$, the s -iterated logarithm function, be defined inductively: $I^0(n)=n$, $I^{s+1}(n)=\log_2(I^s(n))$ for $s \geq 0$. We show that for every fixed s and all n large enough, there is an n -vertex 3-pushdown graph whose smallest separator contains at least $\Omega(n/I^s(n))$ vertices.

1. Introduction

Let $S: N \rightarrow N$ be a monotone function. An n -vertex graph $G=(V, E)$ (directed or undirected) has an S -separator C if there is a partition $V=A \cup B \cup C$, $|A|, |B| \geq n/3$, $|C| \leq S(n)$ and $E \cap (A \times B) = \emptyset$. A family of graphs is S -separable if every graph in the family has an S -separator. A family is separable if it is S -separable for some $S(n)=o(n)$.

For convenience, we restrict attention to nice functions S . A function S is nice if for every a , $0 < a < 1$, there is b , $a < b < 1$, such that $aS(n) < S(an) < bS(n)$.

Remark 1. The planar separator theorem [10] can be restated as follows: the family of planar graphs is $O(\sqrt{n})$ -separable.

Remark 2. The nonexistence of separators is closely related to expansion property in graphs. For a graph $G=(V, E)$ and $A \subseteq V$, $\Gamma_G(A)$ is the set of neighbours of A . A family of graphs is expanding (with expansion constant d) if for every n -vertex graph $G=(V, E)$ in the family and every $A \subseteq V$ we have $|\Gamma_G(A) - A| \geq d|A||V-A|/n$. It follows that every family of expanding graphs is nonseparable. Since expanding graphs “expand” also small sets (that contain less than one third of the vertices), the converse is not necessarily true.

Outerplanar graphs are graphs that can be embedded on the plane so that all vertices lie on the outer face; equivalently, such a graph can be embedded on the plane so that all vertices lie on one straight line and all edges can be embedded on one of the half planes defined by the line. Formally, an outerplanar graph is a graph $G=(V, E)$ where V is the ordered set $\{1, 2, \dots, n\}$ for some n and $E=S \cup R$, where the spine $S \subseteq \{(i, i+1) | i=1, \dots, n-1\}$ and in R edges do not cross; specifically for each pair of edges $(i_1, j_1), (i_2, j_2)$ in R with $i_1 < i_2 < j_1$ we have $j_2 \leq j_1$. A k -page

* The work of the first author was supported in part by NSF Grants MCS—83—03139, DCR—85—11713 and CCR—86—05353.

** The work of the second author was supported in part by NSF Grants MCS—84—16190. AMS subject classification (1980): 05C35, 68C25, 68C40.

graph is a graph which consists of k outerplanar graphs sharing the same ordered vertex set V . k -page graphs can be considered as undirected graphs or as directed graphs where an edge always goes from a small numbered to a large numbered vertex. If every vertex has at most one incident edge in each page of a k -page graph, the graph is called a k -pushdown graph (or k -pd graph in short).

Obviously, a 2-page graph is planar. Conversely, it was shown in [4] that every planar graph can be embedded in nine pages. The number has been improved to seven [7] and very recently Yannakakis [16] improved it to four and showed that four pages are necessary. Computation graphs of Turing machines are k -pd graphs, where k depends on the number of tapes of the Turing machine. This has been the reason for substantial interest in such graphs. In [12] it was shown that k -pd graphs (considered as directed graphs) contain nontrivial *segregators*. This graph property was used to show that nondeterministic multitape Turing machines are strictly more powerful than their deterministic counterparts, settling a longstanding open problem. A family of directed graphs contains a non-trivial segregator if every n -vertex graph $G=(V, E)$ in the family contains an $o(n)$ set of vertices S (the segregator) such that if we delete S and the edges incident with S from G , each vertex in $V-S$ has at most $o(n)$ (not necessarily immediate) predecessors in the remaining graph. It is quite easy to show that if a family of directed graphs that is closed under containment (i.e. if G is in the family then all the subgraphs of G are) has a nontrivial separator then it has a nontrivial segregator.

k -page graphs also arise in connection with embedding of VLSI circuits [5] and fault tolerant arrays of processors [13]. Intuitively, k -page graphs can be drawn on a “book” with k “pages” with all vertices placed on the “binding”, all edges placed on the pages, and no two edges on a page crossing. For this reason, the minimum k for which a graph is k -page embeddable is called the page number of the graph [3]. See also [1], [14] for a discussion of outerplanar graphs.

The following problems are open:

Problem 1. *Is the family of 3-pd graphs separable?*

Problem 1*. *Is the family of k -page graphs separable for any $k \geq 3$?*

In [9], the second author showed that for any fixed $k \geq 3$, the family of k -page graphs is separable if and only if the family of 3-page graphs is. The equivalence of problems 1 and 1* (derived below) is slightly stronger since in Problem 1 we have 3-pd graphs, which are special 3-page graphs.

Remark 3. One can similarly ask if there is an expanding family of 3-pd graphs (or k -page graphs).

We relate these problems to an open problem from an entirely different domain. Consider a real-time nondeterministic Turing machine with two working tapes and a separate input tape. By “real time” we mean that the machine reads a new symbol each step. We want to simulate it by an on-line one-tape nondeterministic Turing machine. By “on line” we mean that the additional input tape is one way. We refer to this as *the simulation*. It is well known that the simulation can be done in time $O(n^2)$. But the following problem is still open:

Problem 2. *Can the simulation be done in subquadratic ($o(n^2)$) time?*

In [6] we showed that if 3-pd graphs have small separators then one can derive a fast simulation:

Theorem A. *If the family of 3-pd graphs is S -separable, then there is a simulation of time $t(n) = O(S(n)/\log n)n \log n$.*

Corollary 1. *If $S(n) = o(n)$ then $t(n) = o(n^2)$.*

Corollary 1 states that if the answer to Problem 1 is positive, then so is the answer to Problem 2.

In [6] we introduced families F_k of graphs. Informally, an n -vertex graph G in F_k is defined by a string x_G of length at most kn . x_G is a sequence of instructions for a 3-pushdown machine M . M manipulates the symbols $1, \dots, n$ (the vertices of G). Initially, the three pushdown stores of M contain $1, \dots, n$. Each symbol in x_G is a triple which indicates from which pd to pop, on which pd's to push and whether the popped symbol is queried or not. If $i_1, i_2, i_3, i_4, \dots$ is the sequence of vertices queried during the execution of x_G , the edges of G are $(i_1, i_2), (i_3, i_4), \dots$. One can easily show that every k -page graph is in F_{6k} . In [6] we defined a language L which consists of strings associated with graphs in $\cup_k F_k$. L is recognizable by a real-time nondeterministic Turing machine with two working tapes and a separate input tape. We then showed:

Theorem B. *Assume M' is an on-line one-tape nondeterministic Turing machine that accepts L in time $t(n)$ and $k > 0$. Then there is a constant $c = c(M')$ such that if $t(n) \leq cn^2/k$, then F_k is S_k -separable, where $S_k(n) = O(k^2 t(n) \log(n^2/kt(n))/n)$.*

Corollary 2. *If $t(n) = o(n^2)$ then $S_k(n) = o(n)$ for all k .*

In particular ($k=3$) if the answer to Problem 1 is negative so is the answer to Problem 2. Hence,

Corollary 3. *Problems 1 and 2 are equivalent.*

Corollary 4. *If the answer to Problem 1 is negative, then L requires time $\Omega(n^2)$.*

In the case that S in Theorem A (S_k in Theorem B) is not nice, the theorem should be slightly modified (the corresponding expressions are uglier). But Corollaries 1–4 still hold.

Several examples where a graph property implies a theorem concerning computation are known. One example of such a property is the existence of nontrivial segregators mentioned above. Another well known example is from [8], where it was first shown how to pebble an n -vertex directed acyclic graph of constant in-degree with $O(n/\log n)$ pebbles; this then was used to prove that "space is better than time". In both examples an upper bound on a graph property implied an upper bound on time or space complexity. In both cases it is unlikely that the converse theorem holds. (The difference between space and time, and between nondeterminism and determinism is believed to be exponential.) On the other hand, there are examples where lower bounds on the sizes of graphs satisfying certain connectivity properties imply a lower bound on time for certain types of computations [15]. However, our results may constitute the first example where a graph problem is shown to be *equivalent* to a problem in computational complexity.

The following corollaries are easily obtained by using Theorems A and B.

They state properties of graphs and they are proved via a detour through Turing machines. All of them probably have direct proofs.

Corollary 5. *For every k , F_k is separable if and only if the family of k -page graphs is separable if and only if the family of 3-pd graphs is separable.*

The next corollary deals with different definitions of a separator and separability. For $0 < \alpha < 1$, let us define an (α, S) -separator as we defined an S -separator except that we require $|A|, |B| \leq \alpha n$. Thus, an S -separator is a $\left(\frac{1}{3}, S\right)$ -separator.

Corollary 6. *The results above hold if we replace S -separators by (α, S) -separators. In particular, k -page graphs are S -separable if and only if they are (α, S) -separable for some $0 < \alpha < 1$.*

The *bandwidth* of a graph $G=(V, E)$ with respect to the naming $V=\{1, \dots, n\}$ is $\sum_{(i, j) \in E} |i-j|$. The bandwidth of a graph is the minimum bandwidth with respect to all possible namings. The proof of Theorem A uses only the fact that the existence of an $o(n)$ separator implies that the bandwidth is $o(n^2)$. The latter is used to derive a fast simulation. Consequently,

Corollary 7. *k -page graphs have nontrivial separators if and only if they have a sub-quadratic bandwidth.*

Let $l^s(n)$, the s -iterated logarithm function, be defined inductively: $l^0(n)=n$, and $l^{s+1}(n)=\log_2(l^s(n))$ for $s \geq 0$. In this paper we define n -vertex graphs G_k^n in $F_{6(k+2)}$ and derive a lower bound on the sizes of their separators:

Theorem 1. $S_{6(k+2)}(n) = \Omega(n/k10^{2k(k+1)} l^k(n))$.

Theorem 1 is proved by induction on k . Towards this end, let $\{G_k^n\}_{n=1}^\infty$ be a family of graphs in $F_{6(k+2)}$ that establish the lower bound for k . G_{k+1}^n is constructed as follows: We choose the largest prime number p smaller than $n/\log^2 n$ and let $m = \lceil \log^2 p \rceil$. We arrange pm of the n vertices in a rectangular array with p rows and m columns. The remaining vertices never enter the picture—they are left as isolated vertices. In each row, we connect up the graph to be a copy of G_k^n . Each pair of adjacent columns of vertices is connected by a shifting graph, i.e., there are numbers t_1, t_2, \dots, t_{m-1} so that the i -th vertex of column j is adjacent to the $i+t_j \pmod{p}$ -th vertex of column $j+1$. We show by a counting argument that the shifts t_j can be chosen to be integers modulo p so that they satisfy a number theoretic property. We then show that this property of the shifts together with the fact that in each row there is a copy of G_k^n implies the required lower bound on the separator size. Establishing membership in $F_{6(k+3)}$ is easy. Since the doubling graphs provide the basis of the induction ($k=1$), this completes the proof.

By using Theorem 1 together with Theorem B we derive the following corollary:

Corollary 8. *For any positive integer s , the time of the simulation is at least $\Omega(n^2/l^s(n))$.*

It is possible to let s grow as a function of n and derive slightly stronger versions of Theorem 1 and Corollary 8. These results yield the currently best lower bound

for the simulation. The graphs G_k^n are not k -page graphs. But one can modify them to get k -pd graphs that have large separators.

Corollary 9. *There are n -vertex k -pd graphs whose smallest separators have at least $\Omega(n/(2^k k 10^{2k(k+1)} l^k(n)))$ vertices.*

A stronger version of Corollary 9 can be proved by combining Theorems A, B and Theorem 1.

Corollary 10. *For every fixed positive integer s and all n large enough, there is an n -vertex 3-pd graph whose smallest separator contains at least $\Omega(n/l^s(n))$ vertices.*

Proof. Fix $k > 0$. All the statements in this proof are valid for sufficiently large n . By Theorem 1, $S_{6(k+2)}(n) \cong c(k)n/l^k(n)$ and by Theorem B, $S_{6(k+2)}(n) \cong \leq d(k) \frac{t(n)}{n} \log \frac{n^2}{t(n)}$. Combining the two we get $\frac{n^2}{t(n)} \cong \leq f(k)l^{k-1}(n)$. By Theorem A, $\frac{n}{S(n)} \cong \leq \frac{n^2 \log^2 n}{t(n \log n)} \cong \leq g(k)l^{k-1}(n)$ from which the corollary follows. ■

In Section 2 we define the family F_k , in Section 3 we define the graphs G_k^n and in Section 4 we prove Theorem 1. In Section 5 we list some open problems.

2. The family F_k

We define a graph G_x of n vertices by the operation of a 3-pd machine M defined by a "program" x . Initially the three pd's contain $1, \dots, n$. Let $I = \{1, 2, 3\}$ be the index set of the pd's and let $\Gamma = \{(\text{pop}, \text{push}, \text{query}) \mid \text{pop} \in I, \text{push} \in 2^I, \text{query} \in \{T, F\}\}$. A symbol $\gamma = (\gamma_1, \gamma_2, \gamma_3) \in \Gamma$ can be interpreted as follows: pop a symbol from pd number γ_1 , push it on top of the pd's listed in γ_2 , and if $\gamma_3 = T$, then γ is a query symbol.

Consider a string $x \in \Gamma^*$. x is executable by a machine M that pushes and pops the symbols $1, \dots, n$ in its pd's. Let i_1, i_2, i_3, \dots be the sequence of symbols popped by a query symbol $\gamma \in \Gamma$. The graph $G_x = (V, E)$ is defined as follows. $V = \{1, \dots, n\}$ and $E = \{(i_1, i_2), (i_3, i_4), \dots\}$. We denote the set $\{G_x : x \in \Gamma^*, |x| \leq kn\}$ by F_k .

In the proofs below we consider "restricted programs" $x \in \Gamma^*$ in which (i) whenever a symbol is pushed on a pd it is not the pd it was popped from; (ii) a query symbol pops without pushing; and (iii) in x query symbols appear in pairs (corresponding to edges) popping different pd's. Formally, if $\gamma = (\gamma_1, \gamma_2, \gamma_3)$ appears in x , then (i) $\gamma_1 \neq \gamma_2$; (ii) if $\gamma_3 = T$, then $\gamma_2 = \emptyset$; and (iii) if γ is the $(2i-1)$ -th query symbol in x for some $i \geq 1$, then the next symbol in x , $\gamma' = (\gamma'_1, \gamma'_2, \gamma'_3)$ satisfies $\gamma'_3 = T$ and $\gamma'_1 \neq \gamma_1$. Moreover, we assume that (iv) at the end of the "execution" of x the three pd's are empty. We need the following definition for stating the last restriction. Consider a query symbol γ in x . It pops some pd symbol which was pushed l times before. We define the *parity of γ* to be the parity of l and assume that (v) the parities of two query symbols that correspond to an edge are the same. We denote by F'_k the corresponding family of graphs. Obviously $F'_k \subseteq F_k$.

We will need the following graphs: A *doubling graph* is $G_n = (V_n, E_n)$, n is a power of 2, $V_n = \{0, \dots, n-1\}$, $E_n = \{(i, 2i \pmod n), (i, (2i+1) \pmod n) \mid i \in V_n\}$. A *shifting graphs* is $G_n^t = (V_n, E_n^t)$, V_n as before and $E_n^t = \{(i, (i+t) \pmod n) \mid i \in V_n\}$,

where t is called the *shift* of the graph. The following proposition follows from the definitions:

Proposition 1. *k page-graphs are in F_{6k} , doubling graphs are in F'_9 , and shifting graphs are in F'_3 .*

Remark. As can be seen by a simple counting argument, for $ck < n$, most k degree graphs are not in F_{ck} . Hence in some sense F_{2k} contains graphs with average degree at most k that have some simple structure.

3. The Graphs G_k^n

In this section we construct, for every k and all large enough n , the graph G_k^n of n vertices and maximum degree $2k+7$. The graphs G_k^n are constructed inductively in k .

Base ($k=1$): The graphs $G_1^n = (V_n, E_n^1 \cup E_n^2 \cup E_n^3)$ are obtained from the doubling graphs of Section 3. V_n is as before. Let n' be the largest power of 2 not larger than n , and let V_n^1 (V_n^2) be the first (last) n' vertices of V_n . For $i=1, 2$, we choose E_n^i so that the graph (V_n^i, E_n^i) is a doubling graph and

$$E_n^3 = \{(i, n'+i) | i=1, 2, \dots, n-n'\}.$$

(E_n^3 connects the corresponding vertices in $V_n^1 - V_n^2$ and $V_n^2 - V_n^1$.)

Inductive step: To construct G_{k+1}^n we need the following number theoretic lemma.

Lemma 1. *Let p be a prime number, and let $m = \lceil \log^2 p \rceil$. There is a set of $m-1$ integers $T \subset \{1, \dots, p\}$ such that for every set $A \subset \{1, \dots, p\}$ with $|A| = cp$ (c may depend on p), the number of triples (a, t, a^*) , $a \in A$, $t \in T$ and $a^* \in \bar{A} = \{1, \dots, p\} - A$ with $a+t \equiv a^* \pmod{p}$ is at least $c(1-c)pm - O(\sqrt{c(1-c)}pm/(\log \log p)^2)$. (The constant in the O notation does not depend on c or p .)*

We now define G_{k+1}^n . Let p be the largest prime smaller than $n/\log^2 n$ and let $m = \lceil \log^2 p \rceil$. Since for sufficiently large n , $p \cong n/\log^2 n - \frac{1}{10}n^{2/3}$ (cf. [2]), $pm \cong \left(n/\log^2 n - \frac{1}{10}n^{2/3}\right) \log^2 \left(n/\log^2 n - \frac{1}{10}n^{2/3}\right) = n(1-o(n))$. We arrange the pm vertices in a pm -rectangle. The remaining $o(n)$ vertices are isolated vertices. In each of the p rows of the rectangle we connect the vertices by a copy of G_k^m . Each pair of consecutive columns are connected by a shifting graph. The $m-1$ different shifts are the elements of the set T of Lemma 1.

Lemma 2. $G_k^n \in F'_{6(k+2)}$.

Proof Sketch. We use induction on k . $G_1^n \in F'_{18}$ since doubling graphs are in F'_9 (Proposition 1). We now describe the restricted program x that "generates" G_{k+1}^n . We number the vertices of G_{k+1}^n by numbering vertices in each column consecutively. The first $3n$ symbols of x copy the three pd's. The top copies are used to generate the column connections and the bottom copies the row connections. By Proposition 1 each shifting graph is in F'_3 and so is the union of the $m-1$ shifting graphs. So the next substring of x of length $3n$ generates the column connections. Now, by

the induction assumption each copy of G_k^m is in $F'_{6(k+2)}$. But so is the union of p copies of G_k^m because in the latter each vertex (each edge) is replaced by p consecutive vertices (by p edges). The program handles each such p vertices as blocks. Each non query symbol in the program that generates G_k^m is replicated p times. Consequently, the p copies in each block are popped and pushed successively. Each pair of query symbols corresponding to an edge is also replicated p times. As a result x pops alternatively from each block. Finally, restriction (v) guarantees that the two blocks are popped (and queried) in the same order (from the first to the last or vice versa). It is easy to see that x is a restricted program that generates G_{k+1}^n . Since the row connections are generated by a string of length at most $6(k+2)n$, it follows that $|x| \leq 6(k+3)n$ and $G_{k+1}^n \in F'_{6(k+3)}$. ■

Proof of Lemma 1. For two given sets $T, A \subset \{1, \dots, p\}$, we define the following values for $k=1, \dots, p$: $f\left(\frac{k}{p}\right) = \sum_{t \in T} e^{2\pi i t k/p}$, $g\left(\frac{k}{p}\right) = \sum_{a \in A} e^{2\pi i a k/p}$, $g^*\left(\frac{k}{p}\right) = \sum_{a^* \in \bar{A}} e^{2\pi i a^* k/p}$.

Claim 1. The number of triples (a, t, a^*) , $a \in A, t \in T$ and $a^* \in \bar{A}$ with $a+t \equiv a^* \pmod{p}$ is $\frac{1}{p} \sum_{k=1}^p f\left(\frac{k}{p}\right) g\left(\frac{k}{p}\right) \bar{g}^*\left(\frac{k}{p}\right)$.

Proof. Each triple (a, t, a^*) , $a \in A, t \in T$ and $a^* \in \bar{A}$ contributes p to the sum if $a+t \equiv a^* \pmod{p}$ and 0 otherwise. ■

Hence, we have to estimate the sum of Claim 1. In this sum, the term for $k=p$ gives $\frac{1}{p} |A||T||\bar{A}| = c(1-c)pm$. The other terms are estimated by $\left| \frac{1}{p} \sum_{k=1}^{p-1} f\left(\frac{k}{p}\right) g\left(\frac{k}{p}\right) \bar{g}^*\left(\frac{k}{p}\right) \right| \leq \frac{1}{p} \sum_{k=1}^{p-1} \left| f\left(\frac{k}{p}\right) \right| \left| g\left(\frac{k}{p}\right) \right| \left| g^*\left(\frac{k}{p}\right) \right| \leq \max_{1 \leq k < p} \left| f\left(\frac{k}{p}\right) \right| \frac{1}{p} \sum_{k=1}^p \left| g\left(\frac{k}{p}\right) \right| \left| g^*\left(\frac{k}{p}\right) \right| \leq$ (by the Cauchy-Schwarz inequality) $\max_{1 \leq k < p} \left| f\left(\frac{k}{p}\right) \right| \frac{1}{p} \left[\sum_{k=1}^p \left| g\left(\frac{k}{p}\right) \right|^2 \right]^{1/2} \left[\sum_{k=1}^p \left| g^*\left(\frac{k}{p}\right) \right|^2 \right]^{1/2}$.

Note that $\sum_{k=1}^p \left| g\left(\frac{k}{p}\right) \right|^2 = cp^2$ and $\sum_{k=1}^p \left| g^*\left(\frac{k}{p}\right) \right|^2 = (1-c)p^2$.

$\left(\sum_{k=0}^p \left| g\left(\frac{k}{p}\right) \right|^2 \right) = \sum_{k=1}^p g\left(\frac{k}{p}\right) \bar{g}\left(\frac{k}{p}\right) = \sum_{a,b \in A} \sum_{k=0}^p e^{2\pi i(a-b)k/p} = p|A|$ because the

inner sum is p if $a=b$ and 0 otherwise. Similarly $\sum_{k=0}^p \left| g^*\left(\frac{k}{p}\right) \right|^2 = p|\bar{A}|$.

It follows that the contribution of the other terms is $\leq \sqrt{c(1-c)} \cdot p \max_{1 \leq k < p} \left| f\left(\frac{k}{p}\right) \right|$.

We show that if we choose the elements of T uniformly at random, then with probability greater than 0, $\max_{1 \leq k < p} \left| f\left(\frac{k}{p}\right) \right| = O(m/(\log \log p)^2)$, which completes the proof of Lemma 1.

Let $m' = m - 1$, $T = \{t_1, \dots, t_{m'}\}$ and $x_j^k = e^{2\pi i t_j k/p}$. Take $q = [(\log \log p)^2]$, and divide the unit circle into q arcs S_1, \dots, S_q . (S_r goes from $e^{2\pi i(r-1)/q}$ to $e^{2\pi i r/q}$.)

Let $\lambda = \log p / (\log \log p)^4$, and fix k , $1 \leq k < p$. For $1 \leq r \leq q$, let A_r^k be the event $\left| \left\{ j: x_j^k \in S_r \right\} - \frac{m'}{q} \right| \leq \lambda \sqrt{\frac{m'}{q}}$. Since $x_j^k \in S_r$ with probability $\frac{1}{q}$, Probability $(\neg A_r^k) \leq e^{-\lambda^2/2}$ by Chernoff's bound on the tail of the binomial distribution (cf. [3] Corollary 4). (For given k and r denote by $X_j^{k,r}$ the random variable which takes the value 1 if $x_j^k \in S_r$ and 0 otherwise. Then $\neg A_r^k$ is the event that the sum (over j) of $X_j^{k,r}$ differs from its mean $\left(\frac{m'}{q}\right)$ by at least λ standard deviations $\left(\sqrt{\frac{m'}{q}\left(1-\frac{1}{q}\right)}\right)$, and λ is not too large $\left(\lambda < \sqrt{\frac{m'}{q}\left(1-\frac{1}{q}\right)}\right)$)

Claim 2. For $1 \leq k < p$, if A_r^k holds for all $1 \leq r \leq q$, then $\left| f\left(\frac{k}{p}\right) - \left| \sum_{j=1}^m x_j^k \right| \right| \leq O(m/(\log \log p)^2)$.

Proof. Let $n_r^k = \left| \{j: x_j^k \in S_r\} \right|$. By our assumption $\left| n_r^k - \frac{m'}{q} \right| \leq \lambda \sqrt{\frac{m'}{q}}$. Also, for $x_j^k \in S_r$, $|x_j^k - e^{2\pi i r/q}| = O\left(\frac{1}{q}\right)$. Now, $\left| \sum_{j=1}^{m'} x_j^k \right| \leq \left| \sum_{r=1}^q n_r^k e^{2\pi i r/q} \right| + O\left(\frac{m}{q}\right) \leq \frac{m'}{q} \cdot \left| \sum_{r=1}^q e^{2\pi i r/q} \right| + q\lambda \sqrt{\frac{m'}{q}} + O\left(\frac{m}{q}\right) \leq O(m/(\log \log p)^2)$. ■

But the probability that A_r^k does not hold for some $1 \leq r \leq q$ and $1 \leq k < p$ is at most $qpe^{-\lambda^2/4} = o(1)$. Hence, with probability greater than 0, A_r^k holds for all $1 \leq r \leq q$ and $1 \leq k < p$, and by Claim 2, $\max_{1 \leq k < p} \left| f\left(\frac{k}{p}\right) - \left| \sum_{j=1}^{m'} x_j^k \right| \right| \leq O(m/(\log \log p)^2)$. ■

4. The Proof of Theorem 1

We prove that the separators of G_n^k satisfy the lower bound of Theorem 1. We will need the following definition: For a set A of vertices in a graph we denote by $\#(A, \bar{A})$ the number of edges (u, v) with $u \in A$ and $v \in \bar{A}$. Theorem 1 follows immediately from the following lemma.

Lemma 3. For $K > k$ and sufficiently large n , let A be any set of vertices of G_n^k such that $10^{k-k} \leq |A|/n \leq 1/2$. Then $\#(A, \bar{A}) \geq 10^{-2kk} |A|/l^k(n)$.

Doubling graphs satisfy a property stronger than the one needed in Lemma 3. Lemma 4 below is due to Maass [11].

Lemma 4. Any set A of vertices of an n -vertex doubling graph satisfies $\#(A, \bar{A}) \geq |A| |\bar{A}| / (n \log n)$.

Proof of Lemma 4. Assume $n = 2^l$. For all pairs of vertices e, d , let $(d_0^e = e, \dots, d_l^e = d)$ be the unique path from e to d that changes the bits of vertex name one at a time. We define $G(e, r) = \{b \in A \mid b = d_r^e \text{ for some } d \text{ with } d_{r+1}^e \in \bar{A}\}$.

Claim 3. $|\bar{A}| \geq \sum_{r=0}^{l-1} |G(e, r)| 2^{l-r}$ for every $e \in A$.

Proof. To each d in \bar{A} associate \bar{r} , the last index r with $d_r^e \in A$. Since d and d_r^e have the same \bar{r} bits, at most $2^{l-\bar{r}}$ d 's in \bar{A} have the same \bar{r} and the same d_r^e . ■

Claim 4. For every index r and vertex b we have $|\{e|b \in G(e, r)\}| \cong 2^r$.

Proof. Each e in this set has $l-r$ bits in common with b . ■

By summing Claim 3 over A and by Claim 4,

$$\begin{aligned} |A||\bar{A}| &\cong \sum_{e \in A} \sum_{r=0}^{l-1} |G(e, r)| 2^{l-r} \cong \sum_{r=0}^{l-1} 2^{l-r} \sum_{e \in A} |G(e, r)| \cong \\ &\cong \sum_{r=0}^{l-1} 2^{l-r} 2^r |\bigcup_{e \in A} G(e, r)| = n \sum_{r=0}^{l-1} |\bigcup_{e \in A} G(e, r)|. \end{aligned}$$

Therefore there is an index r such that $|\bigcup_{e \in A} G(e, r)| \cong |A||\bar{A}|/(n \log n)$.

This completes the proof of Lemma 3 since each vertex in $\bigcup_{e \in A} G(e, r)$ corresponds to a unique edge of the doubling graph with one endpoint in A and one in \bar{A} . ■

Proof of Lemma 3. We fix K , and use induction for $k < K$.

Base ($k=1$): Let A be any subset of vertices of $G_n^1 = (V_n^1, E_n^1 \cup E_n^2 \cup E_n^3)$ such that $10^{1-K} \cong |A|/n \cong 1/2$. Let $a_i = |A \cap V_n^i|$, $i=1, 2$ and $a = |A|$. Without loss of generality $a_1 \cong a_2$.

Claim 5. Either (1) there is $i \in \{1, 2\}$ with $\frac{a}{8} \cong a_i \cong \frac{7}{8} n'$ or (2) $a_1 - a_2 \cong \frac{a}{8}$.

Proof. Assume (1) and (2) do not hold. We must have $a_1 > \frac{7}{8} n'$ and thus $|\bar{A} \cap V_n^1| \cong \frac{n'}{8}$. Now by (2) $a_2 \cong a_1 - \frac{a}{8} \cong \frac{7}{8} n' - \frac{a}{8} \cong \frac{3}{4} a$ and by (1) $a_2 \cong \frac{7}{8} n'$ and thus $|\bar{A} \cap V_n^2| \cong \frac{n'}{8}$. It follows that $|\bar{A}| \cong \frac{n'}{4}$, which is impossible. ■

If (1) holds, then since (V_n^i, E_n^i) is a doubling graph, we have by Lemma 4 and Claim 5: $\#(A, \bar{A}) \cong a_i(n' - a_i)/(n' \log n') \cong \frac{a}{8} \frac{n'}{8} / (n' \log n') \cong |A|/(100 \log n)$.

If (2) holds, then since there are $a_1 - a_2$ edges in E_n^3 which contribute to $\#(A, \bar{A})$, we have $\#(A, \bar{A}) \cong |A|/10$.

Induction step: We assume the lemma holds for k and prove it for $k+1$. We color the vertices in A black and those in \bar{A} white. A row is called *balanced* if it has at least $10^{k-K} m$ vertices of each color, and it is called *unbalanced* otherwise.

First assume that at least $\frac{1}{10} 10^{k-K} p$ of the rows are balanced. We apply the induction hypothesis to each of the balanced rows, choosing the set A in each row to be the smaller of the two sets (of the black vertices or the white vertices) so that its size is $\cong m/2$. It follows that for sufficiently large n , $\#(A, \bar{A}) \cong 10^{k-K-1} p 10^{-2k} 10^{k-K} m/l^k(m) \cong 10^{-2(k+1)K} n/l^{k+1}(n)$, which completes the proof in this case.

We now assume that less than $\frac{1}{10} 10^{k-K} p$ of the rows are balanced. Call a row black if it is not balanced and most of its vertices are black, and call a row white otherwise. Let $f_b(f_w)$ be the fraction of black (white) rows and let f'_w be the fraction of white balanced rows.

In a black row, at least $(1 - 10^{k-K})m \cong \frac{9}{10}m$ vertices are black. Hence, since $|A| \cong n/2$, $\frac{9}{10}f_b pm \cong \frac{5}{10}n$, $f_b \cong 0.56$, and $1 - f_b \cong 0.44$ for n large enough.

In a white unbalanced row, at least $(1 - 10^{k-K})m$ vertices are white. Hence, since $|\bar{A}|/n \cong 1 - 10^{k+1-K}$, $f'_w(1 - 10^{k-K}) \cong 1 - 10^{k+1-K}$. But, since for $0 < y < x < 1$, $(1-x)/(1-y) < 1 - (x-y)$, $1 - f'_w > 9 \times 10^{k-K}$ and $f_b = 1 - f'_w > 8.9 \times 10^{k-K}$.

We now consider only edges connecting different columns. Since these connections were chosen according to the set T of Lemma 1, the number of edges connecting a vertex in a black row to one in a white row is at least $pm(f_b(1 - f_b) - o(1)) \cong 3.5 \times 10^{k-K}n$ for n large enough. Subtracting edges connecting vertices in black rows to vertices in balanced (white) rows ($\cong 0.2 \times 10^{k-K}n$), and edges connecting vertices of the wrong color (white vertices in black rows or black vertices in white unbalanced rows), ($\cong 2 \times 10^{k-K}n$) we get that the number of edges connecting black vertices to white vertices $= \#(A, \bar{A}) > 10^{k-K}n \cong 10^{-2kK}|A|/l^{k+1}(n)$. ■

5. Open Problems

The main open problem left is Problem 1 (or equivalently Problem 2). Theorem 1 and Corollary 10 imply that even if the answer to the problems is positive the smallest separator of 3-pd graphs must be almost linear and the fastest simulation must be almost quadratic. Two other problems are the following

— Can the family of 3-pd graphs be replaced by a simpler family of graphs F such that F is separable if and only if the answers to Problems 1 and 2 is positive?

— Theorems A and B provide upper and lower bounds (for the size of the separators or the time of the simulation) that are tight only if the answer to problems 1 and 2 is negative. Can the gap be closed (even without settling these problems, or if the answer is positive)?

Acknowledgment. This paper was conceived during a visit by the three authors at the Department of Computer Science of the University of Chicago in May 1985. The authors would like to thank Bob Soare and the members of the department for their kind hospitality. Thanks also to the referees for not letting us get away with some "shortcuts" in the proofs.

References

- [1] B. S. BAKER (1983), Approximation algorithms for NP-complete problems, in *Proc. 24th IEEE Annual Symp. on the Foundations of Computer Science*, 265—273.
- [2] B. BOLLOBÁS (1978), *Extremal Graph Theory*, Prentice Hall, London, New York, pp. xx.
- [3] B. BOLLOBÁS (1985), *Random Graphs*, Prentice Hall, London, New York, pp. 11.
- [4] J. BUSS and O. SHOR (1984), On the pagenumber of planar graphs, in *Proc. 16th ACM Symp. on the Theory of Computing*, 98—101.
- [5] F. CHUNG, F. T. LEIGHTON and A. ROSENBERG (1984), *Embedding graphs in books: A layout problem with applications to VLSI design*, manuscript.
- [6] Z. GALIL, R. KANNAN and E. SZEMERÉDI, On nontrivial separators for k -page graphs and simulations by nondeterministic one-tape Turing machines, to appear in *JCSS*.
- [7] L. HEATH (1984), Embedding planar graphs in seven pages, in *Proc. 25th IEEE Symp. on the Foundations of Computer Science*, 74—83.

- [8] J. HOPCROFT, W. PAUL and L. VALIANT (1977), On time versus space, *J. Assoc. Comput. Math.* **24**, 332—337.
- [9] R. KANNAN (1985), Unraveling k page graphs, *Info. & Control* **66**, 1—5.
- [10] R. J. LIPTON and R. E. TARJAN (1980), Applications of a planar graph separator theorem, *SIAM J. Comput.* **9**, No. 3, 615—627.
- [11] W. MAASS (1985), Combinatorial lower bound arguments for deterministic and non-deterministic Turing machines, *Trans. Amer. Math. Soc.* **292**, No. 2, 675—693.
- [12] W. I. PAUL, N. PIPPENGER, E. SZEMERÉDI and W. T. TROTTER (1983), On determinism versus non-determinism, in *Proc. 24th Symp. on the Foundations of Computer Science*, 429—238.
- [13] A. L. ROSENBERG (1983), The Diogenes approach to testable fault-tolerant arrays of processors, *IEEE Trans. Comput.* **C-32**, pp. 902—910.
- [14] M. M. SYSLO (1979), Characterisations of outerplanar graphs, *Discrete Math.* **26**, 47—53.
- [15] L. G. VALIANT, (1976), Graph-theoretic properties in computational complexity, *J. Comput. System Sci.* **13**, 278—285.
- [16] M. YANNAKAKIS (1986), Four pages are necessary and sufficient, in *Proc. 18th ACM Symp. on the Theory of Computing*, 104—108.

Zvi Galil

*Columbia University
and Tel-Aviv University*

Ravi Kannan

Carnegie Mellon University

Endre Szemerédi

*University of Chicago
and the Hungarian Academy of Sciences*