North-Holland

DENSE EXPANDERS AND PSEUDO-RANDOM BIPARTITE GRAPHS

Andrew THOMASON

Dept. of Pure Mathematics and Mathematical Statistics, University of Cambridge, 16 Mill Lane, Cambridge, CB2 1SD, U.K.

1. Introduction

There has been considerable interest over the last few years in bipartite graphs wherein any set of vertices is guaranteed to have quite a large set of neighbours, without the maximum degree being large. Formally, it is customary to define a bipartite graph, with vertex classes X and Y, to be (n, a, b)-*expanding* if $|X| = |Y| = n$ and every subset $A \subset X$ with $|A| = a$ has at least b neighbours in Y; b is usually a function of a. The aim is to find graphs with b as large as possible, given some constraint on the maximum degree. It is not hard to show that random bipartite graphs will have more or less best possible expansion properties, but explicit constructions have proved hard to find.

Most attention has been focussed on *linear* expanders, that is, where the maximum degree is regarded as constant whilst n grows large (see for instance Margulis [10], Chung [6] and Gabber and Galil [7]). However, *dense* expanders, where the maximum degree grows with n, have also found applications, most noticeably to parallel sorting algorithms. The algorithm of Ajtai, Komlós and Szemerédi [1] for sorting n objects in time $O(\log n)$ using $n/2$ processors uses expanders explicitly, and the proofs of the two round sorting algorthms of Häggkvist and Hell [8] and of Bollobás and Thomason [5] make implicit use of the expanding properties of certain dense graphs, in the latter case random graphs. (An algorithm for sorting n objects in r rounds, using m parallel processors, makes m pairwise comparisons in the first round, deduces as much as possible from the result via transitivity, then makes m further comparisons in the second round, and so on. The rth and final round consists of comparing all pairs whose relative order remains in doubt. We denote by $f_r(n)$ the least value of m for which there is such an algorithm. Moreover $f_r(n, d)$ denotes the least m for which there is an algorithm employing only d-step transitivity between rounds; this means if we know $x_0 < x_1 < \cdots < x_k$ we may deduce $x_0 < x_k$ only if $k \le d$.) It was shown in [5] that

$$(2/\sqrt{3} + o(1))n^{\frac{3}{2}} \le f_2(n) \le (1/2)n^{\frac{3}{2}} \log n$$

and

$$2^{-7}n^{1+d/(2d-1)} \le f_2(n, d) \le (1/2d)n^{1+d/(2d-1)}(\log n)^{1/(2d-1)}$$

for fixed d and large n. The proofs were non-constructive. Alon [2], developing an eigenvalue technique of Tanner [11], was able to prove that a certain graph (which we describe later) was $(n, x, v - n^{1+1/d}/x)$-expanding for all x, and showed how, in the case $d = 4$, such a graph would yield a constructive proof that $f(n, 2) = O(n^{\frac{7}{4}})$. For a good survey of the use of graphs in parallel sorting, see Bollobás and Hell [4].

Our purpose in this note is to point out how checking a very simple condition often suffices to show that a dense bipartite graph is a good expander. The check is much easier to apply than the eigenvalue method, though in the cases where both methods are feasible both will give much the same results. The sufficient condition we offer is derived from a study of *pseudo-random graphs* [12]; for a survey see Thomason [13] and for an extension to hypergraphs see Haviland and Thomason [9]. The check involves merely the degrees of vertices and the number of common neighbours of pairs of vertices. In fact it is sufficient to imply the bipartite graph is "pseudo-random" in a sense analogous to that of [12] and [9], but we shall not develop the point.

2. Sufficient conditions

Our principal result is Theorem 2. The proof is very straightforward and is the bipartite analogue of Theorem 1. The latter theorem appeared in [12] but the proof was slightly deficient, so we give a correct version here.

Theorem 1. *Let G be a graph of order n, with minimum degree at least pn, where $0 < p < 1$. Let $\mu \geq 0$ be such that no two vertices of G have more than $p^2 n + \mu$ common neighbours. Then, for every induced subgraph H of G,*

$$\left| e(H) - p\binom{|H|}{2} \right| \leq \alpha\, |H|$$

holds. Here $2\alpha = \varepsilon + \sqrt{pn + \mu\, |H|}$, and $\varepsilon = 1$ if $p\, |H| < 1$, $\varepsilon = 0$ otherwise.

Proof. Let H be a subgraph of G of order $k \leq n$, and let the average degree in H be d. Let $a_1, \ldots, a_k$ be the degree sequence of H, and let $b_1, \ldots, b_{n-k}$ be the number of edges between H and each of the $n - k$ vertices of $G - H$. Then

$$\sum_{i=1}^{k} a_i = kd \quad \text{and} \quad \sum_{i=1}^{n-k} b_i \geq \sum_{i=1}^{k} (pn - a_i) = k(pn - d).$$

Moreover, since no two vertices have more than $p^2 n + \mu$ common neighbours, we have

$$\sum_{i=1}^{k} \binom{a_i}{2} + \sum_{i=1}^{n-k} \binom{b_i}{2} \leq \binom{k}{2}(p^2 n + \mu),$$

so, if $pn \geqslant d$,

$$k\binom{d}{2} + (n-k)\binom{k(pn-d)/(n-k)}{2} \leqslant \binom{k}{2}(p^2n+\mu).$$

Rearranging gives

$$\begin{aligned}(d-p(k-1))^2 &\leqslant [(n-k)/n][(k-1)\mu + np(1-p)] + 2p(d-pk) + p^2 \\ &\leqslant k\mu + (n-k)p(1-p) + 2p(d-pk) + p^2 \\ &\leqslant pn + \mu k \quad \text{if } pn \geqslant d,\end{aligned}$$

which yields the result claimed.

We must now check the result in the case $pn \leqslant d$. Since $b_i \geqslant 0$ the above gives

$$k\binom{d}{2} \leqslant \binom{k}{2}(p^2n+\mu) \quad \text{or} \quad d \leqslant \sqrt{m(p^2n+\mu)+\tfrac{1}{4}} + \tfrac{1}{2}, \tag{1}$$

where $m = k-1$. Our aim is to show $|d-pm| \leqslant 2\alpha$, and since $d \geqslant pn$, we must show $d \leqslant pm + 2\alpha$. So by (1) it is enough to demonstrate $\sqrt{m(p^2n+\mu)+\frac{1}{4}} \leqslant pm + 2\alpha - \frac{1}{2}$; squaring both sides (note $pm + \varepsilon \geqslant \frac{1}{2}$) this is equivalent to proving

$$p(n-m)(pm-1) - \mu \leqslant (2pm + 2\varepsilon - 1)\sqrt{pn+\mu k} + 2\varepsilon pm.$$

This is immediate if $pm \leqslant 1$, the left hand side being negative. Otherwise $\varepsilon = 0$ and $2pm - 1 \geqslant pm$, so it suffices to show $p(n-m) \leqslant \sqrt{pn+\mu m}$. In this case $pn \geqslant pm \geqslant 1$, and we compute from (1) and $pn \leqslant d$ that $\mu m \geqslant pn(p(n-m)-1)$. So we need only verify $p(n-m) \leqslant p\sqrt{n(n-m)}$, and this is clearly true. □

For the bipartite analogue of this theorem it is convenient to use the notation $e(A, B)$ to denote the number of edges between the vertex subsets A and B.

Theorem 2. *Let G be a bipartite graph with vertex classes X and Y, where $|X| = |Y| = n$. Let each vertex in X have degree at least pn, where $0 < p < 1$, and let $\mu \geqslant 0$ be such that no two vertices of X have more than $p^2n + \mu$ common neighbours. Then for every subset $A \subset X$ and every subset $B \subset Y$, with $|A| = a$ and $|B| = b$,*

$$|e(A, B) - pab| \leqslant \varepsilon b + \sqrt{ab(pn+\mu a)},$$

where $\varepsilon = 1$ if $pa < 1$ and $\varepsilon = 0$ otherwise.

Proof. The proof is very similar to that of the previous theorem. Define d by $e(A, B) = bd$. Then $e(A, Y-B) \geqslant apn - e(A, B) = b(rpn - d)$, where $r = a/b$. By estimating common neighbours of pairs of vertices of A, we have, if $rpn \geqslant d$,

$$b\binom{d}{2} + (n-b)\binom{b(rpn-d)/(n-b)}{2} \leqslant \binom{a}{2}(p^2n+\mu).$$

Rearranging gives

$$(d-pa)^2 \leqslant [r(n-b)/n][(a-1)\mu + np(1-p)]$$
$$\leqslant r(pn+\mu a),$$

as desired.

Otherwise $d \geqslant rpn$, and we obtain instead

$$b\binom{d}{2} \leqslant \binom{a}{2}(p^2n+\mu) \quad \text{or} \quad d \leqslant \sqrt{rm(p^2n+\mu)+\tfrac{1}{4}}+\tfrac{1}{2}, \tag{2}$$

where $m = a-1$. Our aim is to show $|d-pa| \leqslant \beta$, where $\beta = \varepsilon + \sqrt{rpn + r\mu a}$. Since $d \geqslant rpn \geqslant pa$ this means we show $d \leqslant pa + \beta$. In fact by (2) it is enough to demonstrate $\sqrt{rm(p^2n+\mu)+\frac{1}{4}} \leqslant pm + \beta - \frac{1}{2}$; squaring both sides (note $pm + \varepsilon \geqslant \frac{1}{2}$) this is equivalent to

$$p(rn-m)(pm-1) - r\mu \leqslant (2pm + 2\varepsilon - 1)\sqrt{rpn + r\mu a} + 2\varepsilon pm.$$

The right hand side is positive so we need only consider the case $pm \geqslant 1$. But then $\varepsilon = 0$ and $2pm - 1 \geqslant pm$, so it suffices to show $p(rn-m) \leqslant \sqrt{rpn + r\mu m}$. Moreover $pm \geqslant 1$ implies $rpn \geqslant pm \geqslant 1$, whence (2) and $rpn \leqslant d$ yield $r\mu m \geqslant rpn(rpn - pm - 1)$. So it need only be shown that $p(rn-m) \leqslant p\sqrt{rn(rn-m)}$, which is manifest. □

We remark that in each of the proofs approximations were made for the sake of obtaining a cleanly stated inequality, designed for ease of general application. In any particular circumstance it would be possible to obtain a better, though likely not significantly better, result.

The expansion properties of the graph described in the last theorem can be stated more explicitly.

Corollary 3. *Let G be a bipartite graph with vertex classes X and Y, where $|X| = |Y| = n$. Let each vertex in X have degree at least pn, where $0 < p < 1$, and let $\mu \geqslant 0$ such that no two vertices of X have more than $p^2n + \mu$ common neighbours. Then G is $(n, a, n - n/pa - \mu/p^2)$-expanding for every $a \leqslant n$.*

Proof. If $pa \leqslant 1$ there is nothing to prove. Otherwise Theorem 2 shows G is $(n, a, n-b)$-expanding, where $pab \leqslant \sqrt{ab(pn + \mu a)}$, as claimed. □

3. Some examples

Here are just a few examples of how Corollary 3 may be used to check the expansion properties of a graph.

(1) Let G be a random bipartite graph, with edge probability $n^{-\delta}$ where $\delta < \frac{1}{2}$. Standard estimates for the binomial distribution show that G satisfies the conditions of the corollary with $p = n^{-\delta}(1 + o(1))$ and $\mu = n^{\frac{1}{2}-\delta}(1 + o(1))$. So G is

$(n, a, n - [n^{1+\delta}/a + n^{\frac{1}{2}+\delta}](1+o(1)))$-expanding. Of course, in this case it would be better to use the usual methods of random graph theory.

(2) Take a projective geometry of dimension d over the finite field of order q. Let X be the points and Y the hyperplanes of the geometry. Form G by joining $x \in X$ to $y \in Y$ if $x \in y$. Then it is straightforward to compute $n = (q^{d+1} - 1)/(q - 1)$, that G is $(q^d - 1)/(q - 1)$-regular, and every two points lie in $(q^{d-1} - 1)/(q - 1)$ hyperplanes. Calculation then reveals that G satisfies the condition of the corollary with $p = n^{-\frac{1}{d}}(1 + o(1))$ and $\mu = 0$, so G is $(n, a, n - (1 + o(1))n^{1+1/d}/a)$-expanding. The $o(1)$ term can be removed by following through the proof of Theorem 2 for this particular graph and being less wasteful. This is the graph mentioned earlier, used by Alon [2].

(3) Take a symmetric block design with parameters v, k, λ. Let X be the set of v objects and Y the set of v blocks, and join $x \in X$ to $y \in Y$ if $x \in y$. This gives us a pn-regular graph, where $n = v$ and $pn = k$. Since each pair of objects occurs together in λ blocks, we have $\lambda = k(k-1)/(v-1) < p^2 n$, so we obtain a $(n, a, n - n/pa)$-expanding graph. The previous example is a special case of this one.

(4) Let X and Y be two copies of an orthogonal geometry of dimension $2d$ over the field of order q, equipped with a quadratic form of minimal Witt index (see Artin [3] for definitions). Form G by joining $x \in X$ to $y \in Y$ if $g(x + y) = 0$. It can be shown that G is $q^{2d-1} - q^d + q^{d-1}$-regular, and no pair of vertices has more than $q^{2d-2} - q^{d-1}$ neighbours in common. On choosing appropriate p and μ, Corollary 3 implies G is $(n, a, n - [n^{1+1/d} + n^{\frac{1}{2}+1/d}]/a - n^{\frac{1}{2}+1/d} - 2n^{1/d})$-expanding.

References

[1] M. Ajtai, J. Komlós and E. Szemerédi, Sorting in $C \log n$ parallel steps, Combinatorica 3 (1983) 1–19.

[2] N. Alon, Eigenvalues, geometric expanders, sorting in rounds and ramsey theory, Combinatorica 6 (1986) 207–219.

[3] E. Artin, Geometric Algebra (Interscience Publishers, New York, 1957).

[4] B. Bollobás and P. Hell, Sorting and graphs, in Graphs and Order, I. Rival, ed., NATO ASI series (Reidel, 1985) 169–184.

[5] B. Bollobás and A. Thomason, Parallel Sorting, Discrete Appl. Math. 6 (1983) 1–11.

[6] F.R.K. Chung, On concentrators, superconcentrators, generalizers and non-blocking networks, Bell Syst. Tech. J. 58 (1978) 1765–1777.

[7] O. Gabber and Z. Galil, Explicit constructions of linear superconcentrators, J. Comp. and Sys. Sci. 22 (1981) 407–420.

[8] R. Häggkvist and P. Hell, Parallel sorting with constant time for comparisons, SIAM J. Comput. 10 (1981) 465–472.

[9] J. Haviland and A. Thomason, Pseudo-random hypergraphs, Discrete Math. 75 (1989) 255–278.

[10] G.A. Margulis, Explicit constructions of concentrators, Problemy Peredachi Informatsii 9(4) (1973) 71–80 (in Russian). English translation in Problems Info. Transmission (Plenum Press, 1975) 325–332.

[11] R.M. Tanner, Explicit construction of concentrators from generalized N-gons, SIAM J. Alg. Discr. Meth. 5 (1985) 287–293.
[12] A. Thomason, Pseudo-random graphs, in Proceedings of Random Graphs, Poznań 1985, M. Karonski, ed., Annals of Discrete Math. 33 (North-Holland, 1987) 307–331.
[13] A. Thomason, Random graphs, strongly regular graphs and pseudo-random graphs, Surveys in Combinatorics, 1987, C. Whitehead, ed., LMS Lecture Note Series 123 (1987) 173–195.