An algorithmic Friedman–Pippenger theorem on tree embeddings and applications*

Domingos Dellamonica Jr.

Department of Mathematics and Computer Science, Emory University
400 Dowman Dr., Atlanta, GA, 30322, USA
ddellam@mathcs.emory.edu

Yoshiharu Kohayakawa

Instituto de Matemática e Estatística, Universidade de São Paulo
Rua do Matão 1010, 05508–090 São Paulo, Brazil
yoshi@ime.usp.br

Submitted: Apr 29, 2008; Accepted: Oct 2, 2008; Published: Oct 13, 2008
Mathematics Subject Classification: 05C05, 05C85

Abstract

An (n, d)-*expander* is a graph $G = (V, E)$ such that for every $X \subseteq V$ with $|X| \leq 2n - 2$ we have $|\Gamma_G(X)| \geq (d + 1)|X|$. A tree T is *small* if it has at most n vertices and has maximum degree at most d. Friedman and Pippenger (1987) proved that any (n, d)-expander contains every small tree. However, their elegant proof does not seem to yield an efficient algorithm for obtaining the tree. In this paper, we give an alternative result that does admit a polynomial time algorithm for finding the immersion of any small tree in subgraphs G of (N, D, λ)-graphs Λ, as long as G contains a positive fraction of the edges of Λ and λ/D is small enough. In several applications of the Friedman–Pippenger theorem, including the ones in the original paper of those authors, the (n, d)-expander G is a subgraph of an (N, D, λ)-graph as above. Therefore, our result suffices to provide efficient algorithms for such previously non-constructive applications. As an example, we discuss a recent result of Alon, Krivelevich, and Sudakov (2007) concerning embedding nearly spanning bounded degree trees, the proof of which makes use of the Friedman–Pippenger theorem.

We shall also show a construction inspired on Wigderson–Zuckerman expander graphs for which any sufficiently dense subgraph contains all trees of sizes and maximum degrees achieving essentially optimal parameters.

Our algorithmic approach is based on a reduction of the tree embedding problem to a certain on-line matching problem for bipartite graphs, solved by Aggarwal et al. (1996).

*Research partially supported by FAPESP and CNPq through a Temático-ProNEx project (Proc. FAPESP 2003/09925-5) and by CNPq (Proc. 485671/2007-7 and 486124/2007-0). The first author is supported by a CAPES/Fulbright Scholarship and the second author was partially supported by CNPq (Proc. CNPq 308509/2007-2).

1 Introduction

Let $G = (V, E)$ be a graph. We say that G is an (n, d)-*expander* if for every $X \subseteq V$ with $|X| \leq 2n-2$ we have $|\Gamma_G(X)| \geq (d+1)|X|$, where, as usual, we write $\Gamma_G(X)$ for the neighborhood of X in G. A tree T is called (n, d)-*small*, or simply *small*, if $|V(T)| \leq n$ and $\Delta(T) \leq d$, that is, if it has at most n vertices and it has maximum degree $\Delta(T) = \max_{v \in V(T)} d_T(v)$ at most d. An *embedding* of a graph H in a graph G is simply an edge-preserving injection $f\colon V(H) \hookrightarrow V(G)$. The following elegant theorem was discovered by Friedman and Pippenger [8].

Theorem 1.1. *Any (n, d)-expander graph contains every small tree.*

Theorem 1.1 generalizes to trees the essence of a well known result of Pósa [13], concerning the existence of long paths in expanders.

Organization of the Paper

In Section 2 we show how one can obtain an algorithmic analogue of Theorem 1.1 when the graph is a constant density subgraph of a pseudorandom graph; see Theorem 2.3. In Section 3, by making use of Theorem 2.3 instead of the original Friedman–Pippenger Theorem in the argument of Alon, Krivelevich and Sudakov [3], we can turn their result—the embedding of nearly spanning bounded degree trees—into an algorithmic result. Section 4 deals with the explicit construction of tree-universal graphs where tree embeddings can be computed efficiently.

2 On-line Games and Tree Embeddings

We reduce the problem of finding a tree embedding to an on-line matching game that we will call $\mathcal{G}$. The game $\mathcal{G}$ can be described as follows. Let $H = (U, W; E)$ be a bipartite graph. At each step there is a matching M (initially empty), an adversary *requests* a vertex $u \in U$, where u is not covered by M, and we are supposed to extend M (by adding a single edge) in order to cover u.

Aggarwal et al. [1, 2] provided a polynomial time algorithm that can find a matching of size n against any adversary provided that the graph H is such that, for every $X \subseteq U$ with $|X| \leq n$, even after at most half the edges incident to every vertex of X are removed, there remains at least $2|X|$ neighbors of X. This result is given by the theorem below.

Definition 2.1. *Let G be a graph and $v \in V(G)$. We denote by $\partial_G(v)$ the set of edges incident to v in G, that is $\partial_G(v) = \{e \in E(G) \mid v \in e\}$.*

Theorem 2.2 ([1]). *Let H be a bipartite graph with classes U, W. Suppose H is such that for every $X \subseteq U$, with $|X| \leq n$, and $F \subseteq E(H)$ such that $|F \cap \partial(x)| \leq d_H(x)/2$ for every $x \in X$, we have $|\Gamma_{H-F}(X)| \geq 2|X|$. There is a polynomial time algorithm $\mathcal{A}$ that can find a matching of size n in the game $\mathcal{G}$ against any adversary.*

In what follows we show how the problem of finding an embedding of a given tree can be reduced to the on-line game $\mathcal{G}$. We also describe some sufficient conditions for a graph G that guarantee that the on-line algorithm $\mathcal{A}$ can be used in order to find successfully an embedding of any (n, d)-small tree in G.

We now introduce a key property that synthesizes a sufficient condition for the success of the reduction of the tree embedding problem. Roughly speaking, we require that for any small set X, after at most half the edges incident to each vertex of X are removed, X still has many neighbors.

Property $\mathcal{P}(n, d)$. For all $X \subseteq V(G)$ with $|X| \leq n-1$ and for any set $F \subseteq E(G)$ such that $|F \cap \partial_G(x)| \leq d_G(x)/2$ for every $x \in X$, we have $|\Gamma_{G-F}(X)| \geq 2d|X| + 1$.

Theorem 2.3. *Let G be a graph satisfying Property $\mathcal{P}(n, d)$. There is a polynomial time algorithm to find an embedding of any (n, d)-small tree T into G.*

A preliminary version of this paper appeared in [7]. In that version, there was a flaw in the proof which forced us to change the embedding strategy. Although we have lost much of the generality in the result, for all applications considered in [7], corollaries derived from Theorem 2.3 are applicable.

Proof. First, define a bipartite graph $H = H(G)$ with vertex classes U and W as follows. Let U consist of d copies of each vertex of G — formally, we let $U = V(G) \times \{1, \ldots, d\}$ — and let $W = \{v' \mid v \in V(G)\}$ consist of single copies of the vertices of G. For each $\{u, v\} \in E(G)$ and $j \in \{1, \ldots, d\}$ we put both $\{(u, j), v'\}$ and $\{(v, j), u'\}$ in $E(H)$.

Let $T_1 \subsetneq T_2 \subsetneq \cdots \subsetneq T_k = T$ be a sequence of trees such that $T_{i+1} = T_i + w_i$ with w_i being a leaf of T_{i+1} for all $i = 1, \ldots, k-1$, where $k = |V(T)| \leq n$. Let $v_i \in V(T_i)$ be the (unique) vertex adjacent to w_i in $V(T_{i+1})$ for $i = 1, \ldots, k-1$. Note that $w_1, \ldots, w_{k-1}$ are all distinct; indeed, these are all the vertices of T, except for the vertex $r = v_1$ in T_1, which we could think of as the "root" of T. Let us moreover observe that, if $v \in V(T)$, then the number of indices j for which we have $v_j = v$ is precisely the number of children that v has in the rooted tree (T, r).

Our procedure consists in building an embedding $f\colon V(T) \hookrightarrow V(G)$ by a sequence of 1-vertex extensions of any trivial one-vertex embedding. Let $f_1\colon v_1 \mapsto x_1$ be an embedding of T_1. Let H' be the graph obtained from H after the deletion of the vertex $x_1' \in W$. We now play the game $\mathcal{G}$ against the algorithm $\mathcal{A}$ on the graph H'. Therefore, we give a sequence of requests $u_1, u_2, \ldots \in U$ and $\mathcal{A}$ tries to fulfill our requests returning edges $\{u_1, z_1'\}, \{u_2, z_2'\}, \ldots \in E(H')$ with all these edges independent.

First, let us define our strategy in the game. Suppose in step i we have an embedding $f_i\colon V(T_i) \hookrightarrow V(G)$. Let $x_j = f_i(v_j)$ for $1 \leq j \leq i$. We proceed by requesting the vertex $u_i = (x_i, \#\{j \mid j \leq i,\ x_j = x_i\}) \in U$. Let $\{u_i, z_i'\}$ be the edge selected by algorithm $\mathcal{A}$ to cover u_i. Extend f_i by setting $f_{i+1}(w_i) = z_i$ and $f_{i+1}|_{V(T_i)} \equiv f_i$. Proceed to step $i + 1$ (or stop, if the embedding of T is complete).

We have to show that the above procedure actually produces embeddings and that algorithm $\mathcal{A}$ always finds an edge to cover the given request and succeeds in extending the matching. The proof will follow by induction. The invariants holding before step i

are: (1) $w' \in W$ is covered by the matching given by $\mathcal{A}$ if and only if $w \in f_i(T_i) \setminus \{x_1\}$; (2) f_i is an embedding of T_i into G. Let us assume, for now, that $\mathcal{A}$ always provides an edge to extend the matching. Clearly, the invariants hold for $i = 1$.

Suppose that, for all $j < i$ the invariants hold. Notice that, since $\Delta(T) \leq d$, the vertex u_i is well defined. Given that $\{u_i, z_i'\}$ was selected by $\mathcal{A}$ to extend the matching, the map f_{i+1} is defined so that $f_{i+1}(T_{i+1}) = f_i(T_i) \cup \{z_i\}$ (with $z_i \neq x_1$) and thus (1) does follows. By the definition of H', we have $\{x_i, z_i\} \in E(G)$, but that is the same as saying that $\{f_{i+1}(v_i), f_{i+1}(w_i)\} \in E(G)$, which implies that f_{i+1} preserves all the edges $E(T_{i+1}) = E(T_i) \cup \{v_i w_i\}$, since f_i preserves $E(T_i)$. Note also that invariant (1) implies that f_{i+1} is injective. Therefore (2) does follow.

We conclude that, as long as the algorithm can provide an edge that extends the matching, we are able to extend the tree embedding. By assumption, Property $\mathcal{P}(n,d)$ holds for G and we shall prove that this property implies that, for every set $X \subseteq U$ with $|X| \leq n-1$, even after at most half of the edges incident to each $x \in X$ are removed from the graph H', the set X still has at least $2|X|$ neighbors in H'.

Suppose $F \subseteq E(H')$ is such that $|F \cap \partial_{H'}(x)| \leq d_{H'}(x)/2$ for all $x \in X$. Let $\pi\colon X \to V$ be the projection onto the first coordinate and set

$$F' = \bigcup_{v \in \pi(X)} \big\{\{v, w\} \in E(G) \;\big|\{z, w'\} \in F \text{ for all } z \in \pi^{-1}(v)\big\}.$$

We claim that $|\Gamma_{H'-F}(X)| \geq |\Gamma_{G-F'}(\pi(X))| - 1$ and that $|F' \cap \partial_G(v)| \leq d_G(v)/2$ for all $v \in \pi(X)$. The first assertion follows from the fact that if we have $w \in \Gamma_{G-F'}(\pi(X)) \setminus \{x_1\}$ then $w' \in \Gamma_{H'-F}(X)$, since if $v \in \pi(X)$ and $\{v, w\} \in E(G) \setminus F'$ we must have some $z \in \pi^{-1}(v)$ such that $\{z, w'\} \notin F$ and thus $w' \in \Gamma_{H'-F}(X)$. The second assertion follows from the assumption on F since $d_{H'}(x) \leq d_G(\pi(x))$ and $|F \cap \partial_{H'}(x)| \geq |F' \cap \partial_G(\pi(x))|$ for all $x \in X$. By Property $\mathcal{P}(n,d)$ we have $|\Gamma_{H'-F}(X)| \geq 2d|\pi(X)| \geq 2|X|$. □

From the proof of Theorem 2.3 above one can actually get a stronger result, as follows.

Corollary 2.4. *Let T be a rooted, (n,d)-small tree and let G be a graph satisfying $\mathcal{P}(n,d)$. Let $r \in V(T)$ be the root of T and let $v \in V(G)$ be any given vertex of G. The embedding f obtained by Theorem 2.3 can be forced to be such that $f\colon r \mapsto v$.*

Proof. Just take the sequence $T_1 \subsetneq T_2 \subsetneq \cdots \subsetneq T_k = T$ such that T_1 is just the root r and fix $x_1 = v$ in the proof of Theorem 2.3. □

We next give a central definition.

Definition 2.5. *An (N, D, λ)-graph is a regular graph with N vertices, degree D, and with all the eigenvalues except the largest $\leq \lambda$ in absolute value.*

We denote by $e_G(X, Y)$ (or simply $e(X, Y)$ when G is clear from the context) the number of pairs $(x, y) \in X \times Y$ such that $\{x, y\} \in E(G)$. Note that $e_G(X, Y)$ counts the number of edges with one endpoint in X and the other endpoint in Y, with edges induced by $X \cap Y$ counted twice. We let $e(X) = e(X, X)/2$, so that $e(X)$ counts the number of edges induced by X.

Lemma 2.6 (Edge Distribution Estimate). *Let Λ be an (N, D, λ)-graph. For any two sets of vertices $X, Y \subseteq V(\Lambda)$, we have*

$$\left| e(X,Y) - \frac{|X|\,|Y|D}{N} \right| \leq \lambda\sqrt{|X|\,|Y|}.$$

We now state a technical lemma that will be useful in the proof of the main theorem of this section and also in the application that follows.

Lemma 2.7. *Let Λ be an (N, D, λ)-graph. Let $\beta > 0$ be a constant. Suppose*

$$\frac{\beta}{2} > \frac{n(1+4d)}{2N} + \frac{\lambda}{D}(1+\sqrt{2d}). \tag{1}$$

Let $G \subseteq \Lambda$ be such that $\delta(G) = \min_{v\in V(G)} d_G(v) \geq \beta D$. Then G satisfies Property $\mathcal{P}(n, d)$.

Proof. Let $X \subseteq V(G)$ such that $r = |X| \leq n-1$. Suppose that Property $\mathcal{P}(n, d)$ fails for this particular set. This means that there is a set $F \subseteq E(G)$ such that $|F \cap \partial_G(v)| \leq d_G(v)/2$ for all $v \in X$ and $|\Gamma_{G-F}(X)| \leq 2d|X|$. Let $T = \Gamma_{G-F}(X) \setminus X$. By assumption, we have $e_G(X) + e_G(X, T) \geq |X|\delta(G)/2 \geq r\beta D/2$. From the edge estimates in Λ we have

$$e_\Lambda(X) + e_\Lambda(X,T) \leq \frac{r^2D(1+4d)}{2N} + \lambda r\bigl(1+\sqrt{2d}\bigr).$$

But, by (1), the right side above is smaller than $r\beta D/2$, a contradiction. □

Theorem 2.8. *Let Λ be an (N, D, λ)-graph. Let $\alpha > 0$ be a constant. Suppose*

$$\lambda < \frac{\alpha D}{12(1+\sqrt{2d})} \tag{2}$$

and

$$N > \frac{6n}{\alpha}(4d+1). \tag{3}$$

Let G be any subgraph of Λ with $|E(G)| \geq \alpha DN/2 = \alpha|E(\Lambda)|$. We can obtain, in polynomial time, a subgraph of G satisfying Property $\mathcal{P}(n, d)$.

Proof. Let G be a graph as above. We will obtain a graph G' from G by deleting vertices until $\delta(G') \geq \alpha D/3$. The procedure is quite simple: if there is a vertex v in G' with $d_{G'}(v) < \alpha D/3$ we delete it.

We prove that this procedure stops before G' has too few vertices (or no vertex at all). Suppose that in some step we deleted γN vertices. Then G' still has $(1-\gamma)N$ vertices and at least $\alpha DN/2 - \gamma\alpha DN/3$ edges. By the above edge distribution estimate for (N, D, λ)-graphs (Lemma 2.6) we have

$$\begin{aligned}\frac{1}{2}\alpha DN - \frac{1}{3}\gamma\alpha DN &\leq \frac{(1-\gamma)^2N^2D}{2N} + \frac{\lambda}{2}(1-\gamma)N \\ &\leq \frac{(1-\gamma)^2DN}{2} + \lambda N.\end{aligned}$$

Dividing both sides by DN and using that $\gamma \leq 1$, we get

$$\frac{\alpha}{6} \leq \frac{(1-\gamma)^2}{2} + \frac{\lambda}{D}.$$

In particular, by our choice of parameters, $\lambda/D \leq \alpha/12$ and hence G' has size $\geq \sqrt{\alpha/6} \cdot N$.

Since, by (2) and (3),

$$\frac{\alpha}{12} > \frac{\lambda}{D}(1+\sqrt{2d})$$

and

$$\frac{\alpha}{12} > \frac{n(1+4d)}{2N},$$

taking $\beta = \alpha/3$ in Lemma 2.7 proves that G' satisfies Property $\mathcal{P}(n,d)$ and we are done. $\square$

By Theorems 2.3 and 2.8 it follows that for any graph G satisfying the conditions of Theorem 2.8 we have a polynomial time algorithm that finds an embedding of any small tree in G, even if the tree is given on-line. Let us state this result formally.

Corollary 2.9. *Let T be a rooted, (n,d)-small tree and let G be a subgraph of an (N,D,Λ)-graph Λ with $|E(G)| \geq \alpha|E(\Lambda)|$. Suppose further that (2) and (3) hold. Let $r \in V(T)$ be the root of T and let $v \in V(G)$ be any given vertex of G. There is a polynomial time algorithm that finds an embedding f of T in G with $f\colon r \mapsto v$.*

3 Embedding almost spanning trees

We mention that the Friedman–Pippenger theorem plays a fundamental rôle in [3], where it is proved that, roughly speaking, random and pseudorandom graphs G with n vertices contain all trees of bounded degree with $(1-o(1))n$ vertices, even if G is rather sparse. To turn the proofs in [3] algorithmic, one may use Theorem 2.8. (The local lemma is also used in [3], but fortunately this does not present difficulties — powerful enough algorithmic versions of that lemma have been developed and they may be simply invoked; see, e.g., [11, 12]).

The embedding plan used in [3] consists in first obtaining a sequence of trees $T_1, \ldots, T_s$ such that, for each $1 \leq j < s$, the tree T_{j+1} is connected to $\bigcup_{i=1}^{j} T_i$ by a single edge. The sizes $|V(T_i)|$ are also controlled. Then, each tree T_i is embedded sequentially, together with the edges connecting T_i with $T_{i+1}, \ldots, T_s$. These extra edges are used to define the roots of the trees that will be embedded later (we use Corollary 2.4 for this).

In order to replace Friedman–Pippenger Theorem with our Corollary 2.4 in this embedding plan, we also need to control the sizes of the trees. Suppose the tree we want to embed has size $(1-\varepsilon)N$, and that the (N,D,λ)-graph Λ under consideration satisfies[1]

$$\begin{aligned} &D \geq 240d^2\varepsilon^{-1}, \\ &\frac{D}{\lambda} > (1+\sqrt{2d})\varepsilon^{-2}. \end{aligned} \tag{4}$$

[1]This is somewhat stronger than the requirement of [3].

We require that the obtained sequence $\{T_i\}_{i=1}^s$ is such that, for all $i > 1$,

$$\frac{\varepsilon^2 N}{30d^2} \leq |V(T_i)| \leq \frac{\varepsilon^2 N}{30d}, \tag{5}$$

and, for $i = 1$, the upper bound holds. This sequence can be obtained by consecutive applications of [3, Proposition 4.2].

From the lower bound in (5), we get that $s \leq 30d^2/\varepsilon^2$. Now we use [3, Lemma 4.4] — an application of the Local Lemma — to obtain a collection of s sets $S_1, \dots, S_s$ such that $|\bigcup_{i=1}^s S_i| \leq \varepsilon N$ and every vertex of Λ has at least $\varepsilon^3 D/(60d^2)$ neighbors in each of the sets S_i.

At the ith step, we are to embed the tree T_i (with some pre-assigned root) and all the edges crossing T_i to yet unembedded trees $T_{i+1}, \dots, T_s$ — we will name T_i' the union of the tree T_i and those crossing edges — into the graph induced by the vertices of Λ that were not already used in the embedding of the previous trees and vertices that do not belong to any S_j with $j > i$. Since $s = o(N)$ and the size of T_i is bounded by (5), it follows that the size of T_i' is bounded (for sufficiently large N) by $\varepsilon^2 N/(20d)$.

Notice that S_i is contained in the graph into which we want to embed T_i'. By the construction of S_i, this implies the minimum degree of such graph is at least $\varepsilon^3 D/(60d^2)$. Since, by (4),

$$\frac{\varepsilon^3 D}{120d^2} > \varepsilon^2 \frac{5d}{40d} + \varepsilon^2 > \varepsilon^2 \frac{(1+4d)}{40d} + \frac{\lambda}{D}(1 + \sqrt{2d}),$$

it follows by Lemma 2.7 that such graph satisfies $\mathcal{P}(\varepsilon^2 N/(20d), d)$ and, by Corollary 2.4, we can find an embedding of T_i' in polynomial time. The members of $V(T_i') \setminus V(T_i)$ will be special vertices that are reserved as roots of the trees that will be embedded later.

We remark that (5) may be replaced by more sophisticated conditions in order to decrease s, the number of trees obtained in the decomposition. This would allow us to soften conditions (4) imposed on D and λ.

4 Constructive lossless expanders and tree universality

In this section we shall construct tree-universal graphs with relatively few edges using the machinery of expander graphs. Beck [5] considered, for an arbitrary tree T, the problem of obtaining a Ramsey graph G with few edges such that no matter how one colors the edges of G with blue and red, there is always a monochromatic copy of T. He showed that any Ramsey graph for T must contain $\Omega(\beta(T))$ edges, where β is a tree invariant, and conjectured that there exist Ramsey graphs for T with $O(\beta(T))$ edges. This conjecture was nearly confirmed by Haxell and the second author [9], who applied random graph methods and an argument very similar to the one used in the proof of the Friedman–Pippenger theorem to show the existence of such Ramsey graphs with $O\big(\beta(T) \log \Delta(T)\big)$ edges.

In recent years, several sophisticated constructions of expander graphs have been discovered. Some of them are algebraic in nature (for instance, [14] which is based on Cayley graphs) and many are explicit in the sense that there is a specified polynomial-time algorithm that completely describes the expander (e.g., [6, 15, 16]). We shall present a construction which is strongly based on *lossless expanders* and bears similarity with the construction of Wigderson–Zuckerman expanders that beat the eigenvalue bound in [17]. This graph construction seems interesting because any sufficiently dense subgraph contains all trees in which both the number of vertices and maximum degree are close to the best possible (see Theorem 4.8).

Definition 4.1. *A* (K,ε)-lossless expander *of left-degree* D *is a bipartite graph* $G = (U,W;E)$ *where every vertex of* U *has degree* D *and every set* $X \subseteq U$ *with* $|X| \leq K$ *is such that* $|\Gamma_G(X)| \geq (1-\varepsilon)D\,|X|$.

Theorem 4.2 (Theorem 7.3, [6]). *For every* $n, t \leq n$, $\varepsilon > 0$ *there is an explicit* $(K = 2^{k_{\max}}, \varepsilon)$*-lossless expander* $G = (U,W;E)$ *of left-degree* 2^d *and with* $|U| = 2^n$, $|W| = 2^{n-t}$, *where*

- $d = O\big(\log^3 \frac{t}{\varepsilon}\big)$, *and*
- $k_{\max} = n - t - d - \log\frac{1}{\varepsilon} - O(1)$.

Moreover, the ith neighbor of any vertex of U *can be found in time* $\mathrm{poly}(n, \log\frac{1}{\varepsilon})$.

It will be simpler to start with a digraph construction and later derive an undirected construction. The following digraph essentially inherits lossless expansion from the expander of Theorem 4.2, but the degree now is much larger.

Theorem 4.3. *For every* $m \leq n$ *and* $\alpha > 0$ *there is an explicit directed graph* Λ *on* 2^n *vertices of out-degree* 2^{d+n-m}, *where* $d = O\big(\log^3 \frac{n-m}{\alpha}\big)$. *Furthermore, there is an absolute constant* $c > 0$ *such that if* $G \subseteq \Lambda$ *has* $\deg^+_G(v) \geq \alpha 2^{d+n-m}$ *for every* $v \in V(G)$ *then* G *contains all trees*[2] T *with at most* $c\alpha^2 2^n$ *vertices and* $\Delta(T) \leq \alpha 2^{d+n-m-3}$.

Furthermore, there is a polynomial time algorithm that finds an embedding of T *into* G *even if* T *is given on-line by single leaf extensions.*

Let $t = n - m$, $\varepsilon = \alpha/4$. Let H_1 be the (K,ε)-lossless expander $(K = 2^{k_{\max}})$ of left-degree $D = 2^d$ obtained from Theorem 4.2. Identify the vertex classes of H_1 with $\{0,1\}^n$ and $\{0,1\}^m$. Denote by H_2 a bipartite graph on $\{0,1\}^m \times \{0,1\}^n$ where $y \in \{0,1\}^m$ is adjacent to (y,w) for every $w \in \{0,1\}^{n-m}$. Let Λ be a digraph on $\{0,1\}^n$ where

$$\Gamma^+_\Lambda(x) = \Gamma_{H_2}\big(\Gamma_{H_1}(x)\big).$$

Clearly, every out-degree in Λ is 2^{d+n-m}.

For simplicity, set $N = 2^n$ and $M = 2^m$. We shall consider numbers like $\alpha ND/M$ to be integer by possibly adjusting the value of α to some power 2^{-a}.

[2] The trees are directed with edges leaving the root and going to the leaves.

Claim 4.4. *Let $\beta \geq \varepsilon$ and suppose that $T, X \subseteq \{0,1\}^n$ are such that $\deg^+_\Lambda(v, X) \geq \beta DN/M$ for every $v \in T$ and $|T| \leq K$. Then $|X| \geq (\beta - \varepsilon)|T|\, DN/M$.*

Proof. We have $|\Gamma^+_\Lambda(T)| = \bigl|\Gamma_{H_2}\bigl(\Gamma_{H_1}(T)\bigr)\bigr| = |\Gamma_{H_1}(T)|\, DN/M$. Hence, by our assumption over H_1, we have $|\Gamma^+_\Lambda(T)| \geq (1-\varepsilon)|T|\, DN/M$. On the other hand

$$|\Gamma^+_\Lambda(T)| \leq |X| + (1-\beta)|T|\, DN/M.$$

The claim follows. □

Consider the following on-line embedding algorithm for trees inside G. For simplicity, we shall assume that $V(G) = [|V(G)|]$. Let r be the root of the tree T. We assume that there is a sequence of on-line requests $(r = v_0, -1), (v_1, j_1 = 0), (v_2, j_2), \dots, (v_{|T|-1}, j_{|T|-1})$ where every v_i, $i \geq 1$, is the child of v_{j_i} in T and $j_i < i$. The algorithm is supposed to extend a partial embedding f_{i-1} of $T_{i-1} = T[\{v_0, \dots, v_{i-1}\}]$ by defining $f_i(v_i) \in \Gamma^+_G\bigl(f_{i-1}(v_{j_i})\bigr)$. The initial embedding is given by $f_0 \colon r \mapsto 1$.

Invariant 4.5. *At the beginning of step i, the following invariants hold:*

i. $|Z| \leq i + |C|\, \alpha DN/(8M)$*;*

ii. $|C| < K$*;*

iii. for every $w \notin C$, we have $\deg^+_G\bigl(w, V(G) \setminus Z\bigr) > \alpha DN/(2M)$*.*

Recall that $K = c\,\varepsilon M/D$ for some absolute constant $c > 0$ (which we may and will assume that is of the form 2^{-j} for some positive integer j).

Claim 4.6. *Given an r-rooted tree T with $|T| \leq c\alpha^2 N/32$ and $\Delta(T) \leq \alpha DN/(8M)$, Algorithm 1 finds an embedding of T into G in polynomial-time.*

Proof. It is clear that, if the algorithm does not abort, it produces a valid embedding. It is also simple to check that the algorithm runs in polynomial time, except perhaps for line 1.20, for which we have the analysis in Claim 4.7.

Claim 4.7. *Given a family of sets $S_1, \dots, S_m$ and some $b \in \mathbb{N}^m$ such that, for every $I \subseteq [m]$, we have $\bigl|\bigcup_{i\in I} S_i\bigr| \geq \sum_{i\in I} b_i$, there exists a disjoint family $\mathcal{F} = \{S'_i \subseteq S_i\}_{i=1}^m$ with each $|S'_i| = b_i$. This family can be found in* $\mathrm{poly}(n)$*-time algorithm, where $n = \sum_i |S_i|$.*

Moreover, if $\{S''_i \subseteq S_i\}_{i=1}^k$ is any disjoint family with $|S''_i| = b_i$, we may find $\mathcal{F}$ such that $\bigcup_{i=1}^k S''_i \subseteq \bigcup_{S\in\mathcal{F}} S$.

Proof. Let $H = (U, W; E)$ be the following bipartite graph. Set

$$U = \bigcup_{i=1}^m \{i\} \times [b_i] = \{(1,1), (1,2) \dots, (1, b_1), \dots, (m, 1), \dots, (m, b_m)\}$$

and $W = \bigcup_{i\in U} S_i$. If $w \in S_i$ we have $\{(i,j), w\} \in E$ for all $j \in [b_i]$. We claim that H satisfies the Hall condition and we may find a matching in H covering U. Indeed, given

Algorithm 1: Embedding trees

$M \leftarrow \{(r,1)\}$; `// initialize embedding`
$C \leftarrow \emptyset$; `// critical vertices`
$\mathcal{S} \leftarrow \emptyset$; `// reserved neighborhoods`
$Z \leftarrow \{1\}$; `// set of used or reserved vertices`
$i \leftarrow 1$;
while $i < |T|$ **do**
    $p \leftarrow f_{i-1}(v_{j_i})$; `// embedding of parent vertex`
    if $p \in C$ **then**
        $z_i \leftarrow \min(S_p)$; `// we have` $S_p \in \mathcal{S}$ `and` $S_p \subseteq \Gamma^+_G(p) \cap Z$
        $M \leftarrow M \cup \{(v_i, z_i)\}$;
        $S_p \leftarrow S_p \setminus \{z_i\}$;
    else
        $z_i \leftarrow \min\bigl(\Gamma^+_G(p) \setminus Z\bigr)$;
        $M \leftarrow M \cup \{(v_i, z_i)\}$;
        $Z \leftarrow Z \cup \{z_i\}$;
        $C^* \leftarrow \emptyset$;
        $S \leftarrow \emptyset$;
        while *there exists* $w \in V(G) \setminus (C \cup C^*)$ *having* $\deg^+_G\bigl(w, V(G) \setminus (Z \cup S)\bigr) \le \frac{\alpha DN}{2M}$ **do**
            $C^* \leftarrow C^* \cup \{w\}$;
            find disjoint family $\{S_v \subseteq \Gamma^+_G(v) \setminus Z\}_{v \in C^*}$ covering S and where each set has cardinality $\frac{\alpha DN}{8M}$;
            if *no such family exists* **then**
                abort ;
            $S \leftarrow \bigcup_{v \in C^*} S_v$;
        $Z \leftarrow Z \cup S$;
        $C \leftarrow C \cup C^*$;
        $\mathcal{S} \leftarrow \mathcal{S} \cup \{S_v\}_{v \in C^*}$;
    $i \leftarrow i + 1$;

any $X \subseteq U$, let $Y = \pi_1(X) \subseteq [m]$ be the projection onto the first coordinate. We have $\Gamma_H(X) = \bigcup_{i \in Y} S_i$ and $\left|\bigcup_{i \in Y} S_i\right| \geq \sum_{i \in Y} b_i = |\pi_1^{-1}(Y)| \geq |X|$. Given a matching covering U in H, let $S_i' \subseteq W$ be the set of b_i vertices matched to $\{i\} \times [b_i]$. Clearly, the sets S_i' form a disjoint family.

For the moreover part, let us assume the existence of a family $\{S_i'' \subseteq S_i\}_{i=1}^k$. If $x \in \bigcup_{i=1}^k S_i'' \setminus \bigcup_{S \in \mathcal{F}} S$, let $j \in [k]$ be such that $x \in S_j''$. Since $b_j = |S_j'| = |S_j''|$, there must be $y \in S_j' \setminus S_j''$. Set $S_j' \leftarrow S_j' \setminus \{y\} \cup \{x\}$. Note that this strictly decreases

$$\sum_{i=1}^{k} |S_i' \triangle S_i''|.$$

In particular, since this number is always non-negative, in at most $\sum_{i=1}^k |S_i' \triangle S_i''|$ steps, we can obtain the desired family $\mathcal{F}$. $\square$

Let us now prove that the algorithm never aborts. In particular, we shall show that Invariants 4.5.i–iii are preserved throughout the execution of the algorithm.

The base case $i = 1$ is trivially satisfied. Assume that the algorithm is at step $i \geq 1$ and that the invariants hold at the beginning of the step. If $p \in C$ holds (line 1.8) then the sets Z and C remain unchanged and thus the invariants are preserved after this step is finished. Notice that S_p is always non-empty (at line 1.9) as the number of elements from S_p that can be used by the algorithm is at most $\Delta(T)$.

Assume that $p \notin C$. Since Invariant 4.5.iii holds, line 1.13 is well defined. Suppose that C^* grew (line 1.19) so much that $|C \cup C^*| = K$. Since every $w \in C \cup C^*$ is such that $\deg^+_G(w, Z \cup S) \geq \alpha DN/(2M)$, by Claim 4.4, we have $|Z \cup S| \geq \alpha KDN/(4M) = c\alpha^2 N/16$. On the other hand, Invariant 4.5.i indicates $|Z \cup S| = |Z| + |C^*|\, \alpha DN/(8M) \leq i + |C \cup C^*|\, \alpha DN/(8M)$, a contradiction since $i < |T|$. It follows that Invariants 4.5.i–iii hold after the inner loop finishes (if it does not abort).

We finish the proof by showing that line 1.20 always succeeds and the algorithm does not abort. By the argument above, we always have $|C^*| < K$. Furthermore, by Invariant 4.5.iii, every $w \in C^*$ has $\deg^+_G\big(v, V(G) \setminus Z\big) \geq \alpha DN/(2M)$. For any $C' \subseteq C^*$, the sets C' and $X = \Gamma^+_G(C') \setminus Z$ satisfy the conditions of Claim 4.4. Hence, $|X| \geq \alpha\, |C'| DN/(4M)$, which is enough to ensure that the disjoint family of line 1.20 can be found. $\square$

Theorem 4.3 follows directly from Claim 4.6. We now turn to the undirected version of Theorem 4.3. In the result below, the $o(1)$ terms refer to $r \to \infty$. More specifically, we have $r^{o(1)} = \exp\{\mathrm{poly}(\log\log r)\}$ in both occurrences.

Theorem 4.8. *For every $\alpha > 0$, $r \geq r_0(\alpha)$, $\Delta \in \mathbb{N}$ ($\Delta < r^{1-o(1)}$) there exists an explicit graph G on $\Theta\big(\alpha^{-2} r\big)$ vertices, with maximum degree bounded by $r^{o(1)}\Delta/\alpha^2$ such that, for every $H \subseteq G$ with $e(H) \geq \alpha\, e(G)$, the graph H contains all (r, Δ)-small trees.*

Furthermore, there is a polynomial time algorithm that finds an embedding of any (r, Δ)-small tree T into G even if T is given on-line by single leaf extensions.

Proof. Set $\varepsilon = 2^{-8}\alpha$. Pick the smallest n such that $2^n \geq \frac{r}{2^{12}c\alpha^2}$, where c is the constant of Theorem 4.3. Hence, $2^n < \frac{r}{2^{11}c\alpha^2}$. Let $m \leq n$ be the greatest integer such that the value $d = d(m) = \Theta\big(\log^3\big\{\frac{n-m}{\varepsilon}\big\}\big)$ from Theorem 4.3 satisfies $\alpha 2^{n+d-m-9} \geq \Delta$. If $m = n$, then $2^{n+d-m} = r^{o(1)}$. If $m < n$, since $d(m+1) = \Theta(d(m))$, we have $\alpha 2^{n+d(m+1)-m-10} = 2^{n+d-m}(\alpha 2^{\Theta(d)-d-10}) < \Delta$, which means that $2^{n+d-m} = r^{o(1)}\Delta/\alpha$.

Now take the graph Λ of Theorem 4.3 with $\alpha \leftarrow 4\varepsilon$ (with this choice of parameters, the ε in the proof of Theorem 4.3 coincides with the current ε). Let $f\colon V(\Lambda)^2 \to \binom{V(\Lambda)}{2}$ be given by

$$f(u,v) = f(v,u) = \{u,v\}.$$

Slightly abusing the notation, say $G = f(\Lambda) = \big(V(\Lambda), f(E(\Lambda))\big)$. Given $H \subseteq G$, its corresponding digraph (that is, the digraph obtained by putting back the orientations on the edges of H) is $f^{-1}(H) = \big(V(H), f^{-1}(E(H))\big)$.

Claim 4.9. *Let $G = f(\Lambda)$. Suppose that $H \subseteq G$ is such that $e(H) \geq \beta e(G)$. There exists an induced subgraph $H' \subseteq H$ such that $\Lambda' = f^{-1}(H')$ satisfies $\deg^+_{\Lambda'}(v) \geq \frac{\beta}{8}2^{n+d-m}$ for every vertex $v \in V(\Lambda') = V(H')$.*

Proof. Let $\Lambda'' = f^{-1}(H)$. The number of edges in Λ'' is at least $\beta e(G) \geq \beta e(\Lambda)/2 = (\beta/4)2^{2n+d-m}$. While there exists a vertex in Λ'' of out-degree smaller than $\frac{\beta}{8}2^{n+d-m}$, remove that vertex from Λ''. When this procedure stops, we set $\Lambda' = \Lambda''$ and the number of edges incident to removed vertices is upper bounded by $2^n(\frac{\beta}{8}2^{n+d-m})$. It follows that there remains at least $(\beta/8)2^{2n+d-m}$ edges in Λ'. Take $H' = H[V(\Lambda')] = f(\Lambda')$. $\square$

Take $G' = f(\Lambda)$. We shall eliminate high degree vertices of G' in order to obtain the desired graph. While there exists a vertex v in G' with $\deg_{G'}(v) \geq 4(c\varepsilon)^{-1}2^{n+d-m}$, remove it from G'. When such procedure ends, let S be the set of removed vertices. Since the number of edges of G is at most 2^{2n+d-m}, we have $|S| \leq \frac{c\varepsilon}{4}2^n$. Let $x_1, \ldots, x_t$ be a maximal sequence of vertices of $V(\Lambda) \setminus S$ such that, for every $j \in [t]$, we have

$$\deg^+_\Lambda(x_j, S \cup \{x_i \mid i < j\}) \geq \frac{1}{2}2^{n+d-m}.$$

Note that, in particular, every vertex not in $S \cup \{x_i \mid i = 1, \ldots, t\}$ has out-degree at least $2^{n+d-m-1}$ in $\Lambda \setminus \{S\} \setminus \{x_i \mid i = 1, \ldots, t\}$. Suppose that $t \geq K$ (recall that $K = c\varepsilon 2^{m-d}$). Taking $T = \{x_1, \ldots, x_K\}$ and $X = S \cup T$ in Claim 4.4, we obtain a contradiction since $|X| = K + |S| \leq \frac{c\varepsilon}{3}2^n$. Define $G = G' \setminus \{x_i \mid i = 1, \ldots, t\} = f(\Lambda[V(\Lambda) \setminus S \setminus \{x_i \mid i = 1, \ldots, t\}])$.

By construction, the number of edges in the graph G is at least $\frac{1}{8}2^{2n+d-m}$ and the maximum degree is bounded by $4(c\varepsilon)^{-1}2^{n+d-m} \leq 2^{\Theta(d)}\Delta/\alpha^2 = r^{o(1)}\Delta/\alpha^2$. Given any $H \subseteq G$ with $e(H) \geq \alpha e(G) \geq \frac{\alpha}{8}|E(f(\Lambda))|$, Claim 4.9 ensures that there is a subgraph $H' \subseteq H$ such that $\Lambda' = f^{-1}(H')$ satisfies $\deg^+_{\Lambda'}(v) \geq \frac{\alpha}{64}2^{n+d-m}$ for every vertex $v \in V(\Lambda')$. Since Λ is constructed through Theorem 4.3, it follows that Λ' (and hence H') contains all trees with at most $c\alpha^2 2^{-12}2^n \geq r$ vertices and degree at most $\alpha 2^{n+d-m-9} \geq \Delta$. $\square$

5 Concluding Remarks

In this paper we have shown an algorithmic theorem for embedding trees in pseudorandom graphs. Our result can be used as a substitute for the Friedman–Pippenger theorem in many situations, including the applications in their original paper and, to our knowledge, in most (if not all) results that followed their work. However, our algorithm requires more than simple expansion and is therefore it does not quite imply the Friedman–Pippenger theorem. It is an interesting open problem to determine if there is an efficient algorithm that does generalize their original theorem (for work in this direction, see [10]).

Another related result in this paper deals with tree-universal graphs. For any r and $\Delta < r^{1-o(1)}$, we construct an explicit graph on $O(r)$ vertices with maximum degree bounded by $O(r^{o(1)}\Delta)$ that contains all (r, Δ)-small trees. Moreover, this graph is resilient with respect to tree-universality: even if a fraction, say 0.99, of its edges are removed by a malicious adversary, the graph still contains all (r, Δ)-small trees. This construction is optimal up to a multiplicative factor in all parameters except for the $o(1)$ appearing in the exponent of the upper bound for the maximum degree of the constructed graph. Perhaps a more careful analysis of our construction may yield a better bound for the maximum degree.

Acknowledgment

The authors are grateful to the referee for the comments and suggestions.

References

[1] A. Aggarwal, A. Bar-Noy, D. Coppersmith, R. Ramaswami, B. Schieber, and M. Sudan. Efficient routing in optical networks. *J. ACM*, 43(6):973–1001, 1996.

[2] A. Aggarwal, R. Ramaswami, A. Bar-Noy, B. Schieber, D. Coppersmith, and M. Sudan. Efficient routing and scheduling algorithms for optical networks. In *Proceedings of the 5th Annual ACM-SIAM Symposium on Discrete Algorithms (Arlington, VA, 1994)*, pages 412–423, New York, 1994. ACM.

[3] N. Alon, M. Krivelevich, and B. Sudakov. Embedding nearly-spanning bounded degree trees. *Combinatorica*, 27(6):629–644, 2007.

[4] S. Arora, F. T. Leighton, and B. M. Maggs. On-line algorithms for path selection in a nonblocking network. *SIAM J. Comput.*, 25(3):600–625, 1996.

[5] J. Beck. On size Ramsey number of paths, trees and circuits. II. In *Mathematics of Ramsey theory*, volume 5 of *Algorithms Combin.*, pages 34–45. Springer, Berlin, 1990.

[6] M. Capalbo, O. Reingold, S. Vadhan, and A. Wigderson. Randomness conductors and constant-degree lossless expanders. In *Proceedings of the 34th Annual ACM Symposium on Theory of Computing*, pages 659–668 (electronic), New York, 2002. ACM.

[7] D. Dellamonica, Jr. and Y. Kohayakawa. An algorithmic Friedman-Pippenger theorem on tree embeddings and applications to routing (extended abstract). In *Proceedings of the 17th Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1038–1044, New York, 2006. ACM.

[8] J. Friedman and N. Pippenger. Expanding graphs contain all small trees. *Combinatorica*, 7(1):71–76, 1987.

[9] P. E. Haxell and Y. Kohayakawa. The size-Ramsey number of trees. *Israel J. Math.*, 89(1-3):261–274, 1995.

[10] P. E. Haxell and M. Loebl. On defect sets in bipartite graphs (extended abstract). In *Algorithms and computation (Singapore, 1997)*, volume 1350 of *Lecture Notes in Comput. Sci.*, pages 334–343. Springer, Berlin, 1997.

[11] T. Leighton, C.-J. Lu, S. Rao, and A. Srinivasan. New algorithmic aspects of the local lemma with applications to routing and partitioning. *SIAM J. Comput.*, 31(2):626–641 (electronic), 2001.

[12] M. Molloy and B. Reed. *Graph colouring and the probabilistic method*, volume 23 of *Algorithms and Combinatorics*. Springer-Verlag, Berlin, 2002.

[13] L. Pósa. Hamiltonian circuits in random graphs. *Discrete Mathematics*, 14:359–364, 1976.

[14] E. Rozenman, A. Shalev, and A. Wigderson. Iterative construction of Cayley expander graphs. *Theory Comput.*, 2:91–120, 2006.

[15] A. Ta-Shma and C. Umans. Better lossless condensers through derandomized curve samplers. In *FOCS '06: Proceedings of the 47th Annual IEEE Symposium on Foundations of Computer Science*, pages 177–186, Washington, DC, USA, 2006. IEEE Computer Society.

[16] A. Ta-Shma, C. Umans, and D. Zuckerman. Loss-less condensers, unbalanced expanders, and extractors. In *Proceedings of the 33rd Annual ACM Symposium on Theory of Computing*, pages 143–152 (electronic), New York, 2001. ACM.

[17] A. Wigderson and D. Zuckerman. Expanders that beat the eigenvalue bound: explicit construction and applications. *Combinatorica*, 19(1):125–138, 1999.