Theoretical Computer Science 406 (2008) 207–214

Contents lists available at ScienceDirect

Theoretical Computer Science

journal homepage: www.elsevier.com/locate/tcs

The Steiner tree problem on graphs: Inapproximability results

Miroslav Chlebík [a,*], Janka Chlebíková [b]

[a] *Department of Mathematics, University of Sussex, Brighton, BN1 9RF, UK*

[b] *Faculty of Mathematics, Physics, and Informatics, Comenius University, Mlynská Dolina, 824 48 Bratislava, Slovakia*

ABSTRACT

The Steiner tree problem on weighted graphs seeks a minimum weight subtree containing a given subset of the vertices (terminals). We show that it is NP-hard to approximate the Steiner tree problem within a factor 96/95. Our inapproximability results are stated in a parametric way, and explicit hardness factors would be improved automatically by providing gadgets and/or expanders with better parameters.

1. Introduction

Consider a graph $G = (V, E)$ with weight function $w: E \to \mathbb{R}^+$ on the edges and a set of required vertices $T \subseteq V$, called the *terminals*. A *Steiner tree* $\mathcal{T}$ is a subtree of G that spans all vertices in T (using vertices in $V \setminus T$ as well) and its weight is defined by $w(\mathcal{T}) = \sum_{e \in E(\mathcal{T})} w(e)$.

The Steiner Tree problem (STP) is to find a Steiner tree of minimum weight. Steiner trees are important in various applications, for example, in VLSI design, wirelength estimation, and network routing.

An instance of the Steiner tree problem is called *quasi-bipartite* if there is no edge within the set $V \setminus T$, and *uniformly quasi-bipartite* if it is quasi-bipartite and all edges incident to the same non-terminal vertex have the same weight.

The Steiner Tree problem is among the 21 basic problems for which Karp has shown NP-hardness [11]. As we cannot expect to find polynomial time algorithms for solving it exactly (unless P = NP), strong research was done in the area of effective approximation algorithms. During the past years many approximation algorithms for the Steiner Tree problem were designed, see [9] for a survey. The currently best approximation algorithm of Robins and Zelikovsky [14] has an approximation factor of 1.550, and 1.279 for quasi-bipartite instances. In the case of uniformly quasi-bipartite instances, the best known algorithm has an approximation factor 1.217 [9].

It is a natural question how small the approximation factor of the polynomial time algorithm for the Steiner Tree problem can be. Unless P = NP, it cannot be arbitrarily close to 1. This follows from the PCP-Theorem [1] and from the fact that the problem is APX-complete [3].

The starting point of our inapproximability results for the Steiner Tree problem on graphs was the results given by Thimm [16]. Some errors of [16] were fixed in its journal version, and inapproximability within $\frac{163}{162}$ is claimed there under the slightly more restrictive assumption RP ≠ NP. It should be mentioned that some changes are still necessary, to make the proof in the journal version of [16] correct. The author should use two-sided expanders, instead of one-sided ones, along the lines of his proof. Otherwise the crucial assumption $|U_1| \leq |U_2|$ (line 14 from below, p. 394) is hardly "without loss of generality", as claimed there.

The main result of this paper improves the lower bounds on approximability of the Steiner Tree problem and reduces the gap between known approximability and inapproximability results: *It is* NP-*hard to approximate the* Steiner Tree

* Corresponding author.
E-mail addresses: m.chlebik@sussex.ac.uk (M. Chlebík), chlebikova@fmph.uniba.sk (J. Chlebíková).

0304-3975/$ – see front matter
doi:10.1016/j.tcs.2008.06.046

PROBLEM *within a factor* 1.01063 ($>\frac{96}{95}$). *For the case of (uniformly) quasi-bipartite instances approximation within a factor* 1.00791 ($>\frac{128}{127}$) *is* NP-*hard.*

Preliminaries

Our inapproximability results use a reduction from Håstad's NP-hard gap type result for MAX-E3-LIN-2, the maximum satisfiability problem for linear equations modulo 2 with exactly 3 unknowns per equation.

Definition 1. Let a system of linear equations over $\mathbb{Z}_2$ with exactly 3 variables in each equation be given. The goal of the MAX-E3-LIN-2 problem is to find an assignment to the variables that satisfies as many equations as possible.

To suit our purposes we state Håstad's tight inapproximability result in the following way (see [4] for a detailed description how it follows from results in [8], and [13], where it was firstly used in a similar context).

Theorem 2 ([8]). *For every* $\varepsilon \in \left(0, \frac{1}{4}\right)$ *and every fixed sufficiently large integer* $k \geq k(\varepsilon)$, *the following partial decision subproblem of* MAX-E3-LIN-2 *is* NP-*hard:*

$$P(\varepsilon, k) \quad \begin{cases} \text{Given an instance of MAX-E3-LIN-2 with } n \text{ equations} \\ \text{and exactly } 2k \text{ occurrences of each variable, to decide} \\ \text{if at least } (1-\varepsilon)n \text{ or at most } \left(\frac{1}{2}+\varepsilon\right)n \text{ equations are} \\ \text{satisfied by the optimal assignment.} \end{cases}$$

The same NP-hardness result holds on instances where all equations are of the form $x + y + z = 0$ (respectively, all equations are of the form $x+y+z = 1$), where literals x, y, z are variables or their negations, and each variable appears exactly k times negated and k times unnegated. This subproblem of the problem $P(\varepsilon, k)$ will be referred to as $P_0(\varepsilon, k)$ (respectively, $P_1(\varepsilon, k)$) in what follows.

2. NP-hard gap preserving reduction

We start with a set L of n linear equations over $\mathbb{Z}_2$, all of the form $x+y+z = 0$ (respectively, all of the form $x+y+z = 1$), where literals x, y, z are variables from the set $\mathcal{V}$ or their negations, and each variable $v \in \mathcal{V}$ appears in L exactly k times negated as $\overline{v}$ and k times unnegated.

For an assignment $\pi \in \{0, 1\}^{\mathcal{V}}$ to variables let $S(\pi)$ be the number of equations of L satisfied by π. We will reduce the problem of maximizing $S(\pi)$ over all assignments to an instance of the STEINER TREE PROBLEM. To make our reduction approximation preserving, we will use equation gadgets (one for each equation) and couple them properly using $|\mathcal{V}|$ copies of a graph with certain vertex-expansion properties, called the *expander graph*. This expander graph can be the same for all input instances as its choice depends on k only, that is assumed to be constant for the problem $P_0(\varepsilon, k)$ (resp., $P_1(\varepsilon, k)$) we reduce from.

The equation gadgets

Now we define an (α, β, γ)-*gadget*, where α, β, γ are non-negative real numbers, used for each equation from the equation system of the form $x + y + z = 0$ (respectively, of the form $x + y + z = 1$). This will be an instance $G = (V, E)$, $w: E \to \mathbb{R}^+$, $T \subseteq V$ of the STEINER TREE PROBLEM with the following properties:

1. One of the (possibly more) terminal vertices is distinguished and denoted by O.
2. Three of the (possibly more) non-terminal vertices are distinguished and denoted by x, y and z.
3. For any $u \in \{x, y, z\}$ there is a path from u to O of weight at most 1.
4. For any subset R of $\{x, y, z\}$ consider the instance of the STP with altered terminal set $T_R := T \cup R$. The weight of the corresponding minimum Steiner tree is denoted by s_R and it is required to depend only on the cardinality of the set R in the following way,

$$s_R = \alpha + |R|\beta + (|R| \bmod 2)\gamma.$$

(Respectively, if the system L is of the form $x + y + z = 1$, we require $s_R = \alpha + |R|\beta + (1 - |R| \bmod 2)\gamma$.)

An (α, β, γ)-gadget with no edges between non-terminal vertices is called a *quasi-bipartite* (α, β, γ)-*gadget*. A quasi-bipartite (α, β, γ)-gadget such that edges incident to the same non-terminal have the same weight and for vertices x, y, z the incident edges have weight 1, is called a *uniformly quasi-bipartite* (α, β, γ)-*gadget*.

Condition 3 above is just a proper normalization. Condition 4 on $s_l := s_R$, $l = |R| \in \{0, 1, 2, 3\}$, has the following interpretation in our construction: α is a *basic cost* per equation, β is an *extra payment* for connecting some of $\{x, y, z\}$ to the Steiner tree, and γ is a *penalty* for the failure in the parity check of the number of vertices of $\{x, y, z\}$ adjacent to the Steiner tree.

Example 3. For any $\gamma \in \left(0, \frac{1}{4}\right)$ there is a $(0, 1-\gamma, \gamma)$-gadget (for the system L of the form $x+y+z = 0$), depicted in Fig. 1. The vertex O is the only terminal. Clearly $s_0 = 0$, $s_1 = 1$, $s_2 = 2 - 2\gamma$, and $s_3 = 3 - 2\gamma$.

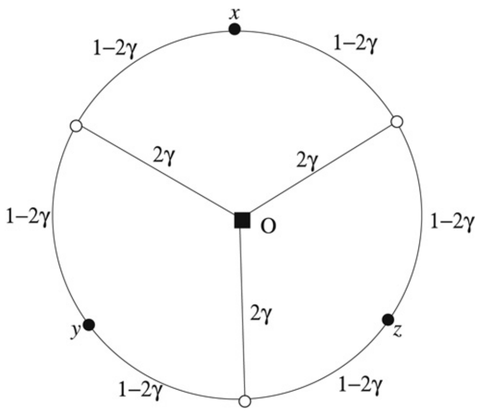

Fig. 1.

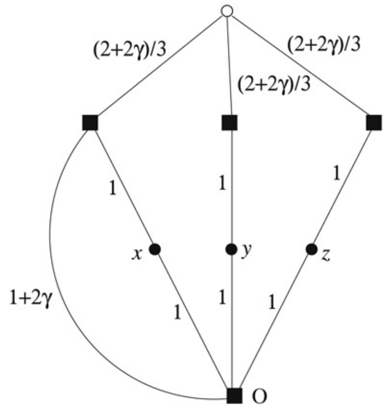

Fig. 2.

Example 4. For any $\gamma \in \langle 0, \frac{1}{2}\rangle$ there is a uniform quasi-bipartite $(3+3\gamma, 1-\gamma, \gamma)$-gadget (for the system L of the form $x+y+z = 1$), depicted in Fig. 2. There are 4 terminals in this gadget, all drawn as boxes. One can easily check that $s_0 = 3+4\gamma$, $s_1 = 4+2\gamma$, $s_2 = 5+2\gamma$, and $s_3 = 6$.

Expansion properties of graphs

Definition 5. An expander with parameters (c, τ, d) (or a (c, τ, d)-expander) is a d-regular bipartite multigraph with k by k bipartition (V_1, V_2), i.e. $|V_1| = |V_2| = k$, such that

$$\text{if } U \subseteq V_1 \quad \text{or} \quad U \subseteq V_2, \quad \text{and} \quad |U| \le \tau k, \text{ then } |N(U)| \ge c|U|,$$

where d is a natural number and c, τ are non-negative real numbers. Here $N(U)$ stands for the set of neighbors of U, $N(U) := \{y : y \text{ is a vertex adjacent to some } x \in U\}$.

We will recall in Section 3 that for any sufficiently large k, a (c, τ, d)-expander with k by k bipartition exists, provided that $0 < \tau < \frac{1}{c} < 1$ and $G_{c,d}(\tau) < 0$, where

$$G_{c,d}(\tau) := H(c\tau) + dc\tau H\left(\frac{1}{c}\right) + (1-d)H(\tau),$$

with $H(x) = -x\ln x - (1-x)\ln(1-x)$ being the entropy function. In fact, under the above assumption, almost all random d-regular bipartite (multi)graphs are (c, τ, d)-expanders, see Theorem 17 in Section 3.

Definition 6. We say that a d-regular bipartite multigraph with k by k bipartition (V_1, V_2) is a c-*good expander* provided the following implication holds:

$$\text{if } \subseteq V_1 \quad \text{or} \quad U \subseteq V_2, \text{ then } |N(U)| \ge \min\{c|U|, k+1-|U|\}.$$

The condition of being a c-good expander for a d-regular bipartite graph is just a bit stronger than the one of being a $\left(c, \frac{1}{c+1}, d\right)$-expander.

Remark 7. In particular, one can observe that a (c, τ, d)-expander with k by k bipartition is c-good, provided that $\tau > \frac{1}{c+1}$ and $k \ge \frac{c}{(c+1)\tau - 1}$.

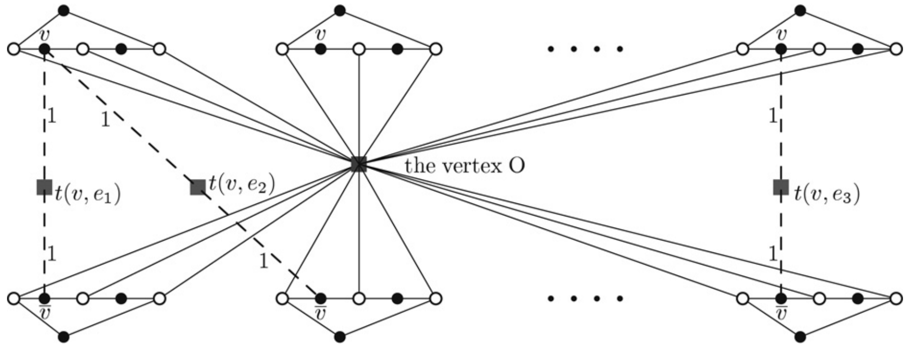

Fig. 3. Here the equation gadget from Example 3 was used. The part of the graph that corresponds to a fixed variable v is depicted. The full edges are those of equation gadgets; the dashed ones are examples of edges adjacent to coupling terminals $t(v, \cdot)$ that correspond to the edges of an expander chosen.

Consider $U \subseteq V_1$ or $U \subseteq V_2$. If $|U| \leq \tau k$, the statement is clear and hence suppose $|U| > \tau k$. The set U contains a subset of cardinality at most $\lfloor \tau k \rfloor$. Hence, $|N(U)| \geq c\lfloor \tau k \rfloor > c(\tau k - 1)$ and due to our assumptions about τ and k, $c(\tau k - 1) \geq k(1-\tau) > k - |U|$. This implies $|N(U)| \geq k + 1 - |U|$.

Consequently, for any sufficiently large k, a d-regular c-good expander with k by k bipartition exists, provided that $c > 1$ satisfies

$$G_{c,d}\left(\frac{1}{c+1}\right) < 0. \tag{1}$$

By continuity of $G_{c,d}$ the inequality (1) implies that $G_{c,d}(\tau) < 0$ also for some $\tau \in \left(\frac{1}{c+1}, \frac{1}{c}\right)$, and we can use the existence result for (c, τ, d)-expanders given in Section 3. For any integer $d \geq 3$ we introduce the constant $c(d)$ defined in the following way:

$$c(d) = \sup\{c : \text{there are infinitely many } d\text{-regular } c\text{-good expanders}\}. \tag{2}$$

Denote by $x(d)$ the unique $x \in (1, \infty)$ for which $G_{x,d}\left(\frac{1}{x+1}\right) = 0$. The existence of such x and its uniqueness can be easily proved and the value of $x(d)$ numerically approximated. To make this clear, notice that $(x+1)G_{x,d}(\frac{1}{x+1})$ simplifies to $(2-d)(x+1)\ln(x+1) + (2d-2)x\ln x - d(x-1)\ln(x-1)$, which is a strictly increasing concave function of variable x on $(1, \infty)$ with negative limit $2(2-d)\ln 2$ at 1, and with growth $2\ln x + 2 + o(1)$ for x approaching $+\infty$. Hence (1) holds for any c in $(1, x(d))$ and, consequently, $c(d) \geq x(d)$ for any integer $d \geq 3$. In particular, $c(6) > 1.76222$ and $c(7) > 1.94606$.

Construction

Now we are ready to describe the gap preserving reduction of instances like L to the instances of the Steiner Tree problem. For this purpose we will use one fixed (α, β, γ)-gadget, and one fixed k by k bipartite d-regular multigraph D which is supposed to be a $\frac{\beta+\gamma}{\beta-\gamma}$-good expander.

For each equation of the system L we take one copy of the fixed (α, β, γ)-gadget and then identify their vertices labeled by O. The resulting graph contains only one vertex labeled by O common to all n (α, β, γ)-gadgets, and it is connected. The x, y, z vertices in each *equation gadget* correspond to occurrences of literals in that equation and we re-label them by those literals. By assumption, each variable from $\mathcal{V}$ appears exactly k times negated and k times unnegated among these labels. Now we couple negated and unnegated occurrences of each variable using one fixed $\frac{\beta+\gamma}{\beta-\gamma}$-good expander D with bipartition (V_1, V_2), $V_1 = \{a_1, a_2, \ldots, a_k\}$, $V_2 = \{b_1, b_2, \ldots, b_k\}$ in the following way:
Assume that equations (and their equation gadgets) are numbered by $1, 2, \ldots, n$. Given a literal x, i.e., $x = v$ or $x = \overline{v}$ for some $v \in \mathcal{V}$, let $m_1(x) < m_2(x) < \cdots < m_k(x)$ be the numbers of equations in which that literal occurs.

Consider one fixed variable v of $\mathcal{V}$. For each edge e of the form a_ib_j from D $(1 \leq i, j \leq k)$ we add a new *coupling terminal vertex* $t(v, e)$. Now connect $t(v, e)$ with the v-vertex in the $m_i(v)$th equation gadget and with the $\overline{v}$-vertex in the $m_j(\overline{v})$th equation gadget, by the edges of weight 1 (see Fig. 3).

Making the above coupling for all variables from $\mathcal{V}$, one after another, we get an instance of the Steiner Tree problem, that corresponds to the system L. Keep this instance fixed, and denote by OPT the minimum weight of a Steiner tree.

Definition 8. We call a Steiner tree $\mathcal{T}$ simple, if each coupling terminal vertex $t(v, e)$ is a leaf of $\mathcal{T}$.

In the following claim we observe that there is also a simple Steiner tree $\mathcal{T}$ of minimum weight.

Claim 9. *OPT* $= \min\{w(\mathcal{T}) : \mathcal{T}$ *is a simple Steiner tree*$\}$.

Proof. It is sufficient to prove that it is possible to transform any given Steiner tree $\mathcal{T}$ with non-empty 'bad' set $\text{BAD}(\mathcal{T}) :=$ $\{$coupling terminals that are not leaves of $\mathcal{T}\}$ to another Steiner tree $\mathcal{T}'$ with $|\text{BAD}(\mathcal{T}')| < |\text{BAD}(\mathcal{T})|$ and $w(\mathcal{T}') \leq w(\mathcal{T})$. Fix one $\mathcal{T}$ with non-empty bad set and choose $t = t(v, e) \in \text{BAD}(\mathcal{T})$. Deleting one of the edges incident to t decreases both $|\text{BAD}(\mathcal{T})|$ and $w(\mathcal{T})$ by 1. In the forest with two components obtained from $\mathcal{T}$ choose a vertex labeled by v or $\overline{v}$ that belongs to a component which does not contain a vertex O. Connect this vertex with O in its equation gadget in the cheapest possible way to obtain the Steiner tree $\mathcal{T}'$.

By property 3 of the (α, β, γ)-gadget it increases the weight by at most 1, hence $w(\mathcal{T}') \leq w(\mathcal{T})$. □

Definition 10. We say that a simple Steiner tree $\mathcal{T}$ is *well-behaved* if it is locally minimal in the following sense: consider any equation of L, say the ith, $i \in \{1, 2, \ldots, n\}$. Let x, y, z be its literals, $T := T^i$ be the set of terminal vertices of its equation gadget, and $R := R^i$ be the set of vertices of this gadget labeled by x, y, or z, that belong to $\mathcal{T}$. The subgraph $\mathcal{T}^i$ of $\mathcal{T}$ induced by this equation gadget is supposed to be the local minimal Steiner tree (in this gadget) for the altered terminal set $T_R := T \cup R$.

Claim 11. OPT $= \min\{w(\mathcal{T}) : \mathcal{T}$ *is a well-behaved Steiner tree*$\}$.

Proof. Clearly, any simple Steiner tree $\mathcal{T}$ with $w(\mathcal{T}) = \text{OPT}$ has to be well-behaved, because otherwise one could create, by local change in some of its gadget, a Steiner tree with less weight. In particular, OPT $= \min\{w(\mathcal{T}) : \mathcal{T}$ is a well-behaved Steiner tree$\}$. □

By property 4 of the (α, β, γ)-gadget, the weight of subtree $\mathcal{T}^i$ is $\alpha + |R|\beta + (|R| \bmod 2)\gamma$ (respectively, $\alpha + |R|\beta + (1 - |R| \bmod 2)\gamma$). Hence, the weight of any well-behaved Steiner tree $\mathcal{T}$ can be expressed in the following way: denote by N the number of vertices corresponding to literals belonging to $\mathcal{T}$, and by M the number of equations for which $R := R^i$ above fails the parity check, i.e., $|R^i|$ is odd (respectively, $|R^i|$ is even). Then

$$w(\mathcal{T}) = \alpha n + \frac{3}{2}nd + N\beta + M\gamma. \tag{3}$$

Here $\frac{3}{2}nd$ edges of weight 1 connect all $\frac{3}{2}nd$ coupling terminals as leaves of the tree $\mathcal{T}$. Clearly, $N \geq \frac{3}{2}n$, as at least one from each coupled pair of vertices corresponding to variables has to belong to $\mathcal{T}$ in order to connect the corresponding coupling terminal to the tree $\mathcal{T}$.

Suppose we are given an assignment $\pi \in \{0, 1\}^{\mathcal{V}}$ to variables and let $S(\pi)$ be the number of equations satisfied by π. For the ith equation of L $(i = 1, 2, \ldots, n)$ let $R := R^i_\pi$ denote the set of vertices in its equation gadget labeled by literals with value 1 by the assignment π, and let $T := T^i$ denote the terminals of this equation gadget. Take one (of possibly more) locally minimum Steiner tree in this gadget with altered terminal set $T_R := T \cup R$ and connect each vertex of R to all d coupling terminals adjacent to it. Such a kind of well-behaved Steiner tree (denoted by $\mathcal{T}_\pi$), which is generated by some assignment π, will be called a *standard* Steiner tree.

The weight of a standard Steiner tree $\mathcal{T}_\pi$ can be expressed using (3), where we have now $N = \frac{3}{2}n$ (exactly half of vertices for variables correspond to literals assigned 1), and $M = n - S(\pi)$. Hence

$$w(\mathcal{T}_\pi) = \alpha n + \frac{3}{2}nd + \frac{3}{2}n\beta + (n - S(\pi))\gamma. \tag{4}$$

The challenge is to prove Lemma 12 below that OPT is achieved on a Steiner tree that is standard, i.e., of the form $\mathcal{T}_\pi$ for some assignment π. With this result in hand, using (4) it is easy to see that the hard-gap result of Håstad for the problem $\max S(\pi)$ implies the corresponding hard-gap and inapproximability results for the Steiner Tree problem.

Lemma 12. *If the* (α, β, γ)*-gadget has parameters* $\beta > \gamma \geq 0$*, and an expander graph used for the coupling is* $\frac{\beta+\gamma}{\beta-\gamma}$*-good, then*

$$OPT = \min\{w(\mathcal{T}) : \mathcal{T} \text{ is a standard Steiner tree}\}.$$

Proof. We already know that there exists a well-behaved Steiner tree $\mathcal{T}$ such that $w(\mathcal{T}) = \text{OPT}$. Thus it is sufficient to show that $\mathcal{T}$ can be transformed into a standard Steiner tree $\mathcal{T}^*$ without increasing the weight. In the following we describe such a construction of $\mathcal{T}^*$ from $\mathcal{T}$ in $|\mathcal{V}|$ steps. Consider one variable, $v \in \mathcal{V}$. Let A_1 be the set of vertices labeled by v, and A_2 be the set of vertices labeled by $\overline{v}$. Clearly $|A_1| = |A_2| = k$. Denote by C_i $(i = 1, 2)$ the set of vertices in A_i that are vertices of the tree $\mathcal{T}$, and put $U_i = A_i \setminus C_i$. We will assume that $|U_1| \leq |U_2|$, otherwise we change the role of A_1 and A_2 in what follows.

Let $N(U)$, for a set $U \subseteq A_1$, be the set of vertices in A_2 which are coupled with a vertex in U. Clearly $U_2 \cap N(U_1) = \emptyset$, because otherwise some coupling terminal is not connected to $\mathcal{T}$. Hence $N(U_1) \subseteq C_2$.

As our expander is $\frac{\beta+\gamma}{\beta-\gamma}$-good, it implies that either $|N(U_1)| \geq k + 1 - |U_1|$, or $|N(U_1)| \geq \frac{\beta+\gamma}{\beta-\gamma}|U_1|$.

We see that the first condition is not satisfied, as

$$k - |U_1| \geq k - |U_2| = |C_2| \geq |N(U_1)|.$$

Thus we can apply the second one to get

$$|C_2| \geq |N(U_1)| \geq \frac{\beta + \gamma}{\beta - \gamma}|U_1|. \tag{5}$$

Now we modify $\mathcal{T}$ to the new well-behaved Steiner tree $\mathcal{T}_{\text{new}}$ as follows: all vertices in A_1 and none in A_2 are in $\mathcal{T}_{\text{new}}$, and for any distinguished vertex u which is labeled by a literal distinct from v and $\overline{v}$,

$$u \in \mathcal{T}_{\text{new}} \Leftrightarrow u \in \mathcal{T}.$$

We also connect the coupling terminals accordingly.

Applying formula (3) for well-behaved Steiner trees we obtain

$$w(\mathcal{T}) - w(\mathcal{T}_{\text{new}}) = (N - N_{\text{new}})\beta + (M - M_{\text{new}})\gamma.$$

Clearly, $N - N_{\text{new}} = |C_2| - |U_1|$ and $M_{\text{new}} \le M + |C_2| + |U_1|$, hence

$$w(\mathcal{T}) - w(\mathcal{T}_{\text{new}}) \ge (|C_2| - |U_1|)\beta - (|C_2| + |U_1|)\gamma = |C_2|(\beta - \gamma) - |U_1|(\beta + \gamma),$$

which is non-negative, by (5). Thus $w(\mathcal{T}_{\text{new}}) \le w(\mathcal{T})$.

Now we apply a similar modification to $\mathcal{T}_{\text{new}}$ with another variable. It is easy to see that if we have done this for all variables, one after another, the result $\mathcal{T}^*$ is a standard tree for some assignment, with $w(\mathcal{T}^*) \le w(\mathcal{T})$. Consequently, $w(\mathcal{T}^*) = \text{OPT}$. □

Theorem 13. *For an integer $d \ge 3$ let $c(d)$ be the constant defined in (2). Further let an (α, β, γ)-gadget with $\beta > \gamma > 0$ and $\frac{\beta+\gamma}{\beta-\gamma} < c(d)$ be given. Then for any constant r, $1 < r < 1 + \frac{\gamma}{3d+2\alpha+3\beta}$, it is* NP*-hard to approximate the* Steiner Tree problem *within a factor r.*

Moreover, if the gadget above is (uniformly) quasi-bipartite, the same inapproximability results apply to the (uniformly) quasi-bipartite instances of the STP as well.

Proof. Let an integer $d \ge 3$, an (α, β, γ)-gadget and a number r with the above properties be fixed.

We can choose and keep fixed from now on an $\varepsilon \in \left(0, \frac{1}{4}\right)$ such that

$$r < 1 + \frac{(1-4\varepsilon)\gamma}{3d + 2\alpha + 3\beta + 2\varepsilon\gamma}.$$

Let $k(\varepsilon)$ be an integer such that for any integer $k \ge k(\varepsilon)$ the conclusion of Theorem 2 holds. Since $\frac{\beta+\gamma}{\beta-\gamma} < c(d)$, we can consider and keep fixed from now on one $\frac{\beta+\gamma}{\beta-\gamma}$-good d-regular expander graph D with k by k bipartition such that $k \ge k(\varepsilon)$. It will play the role of a constant in our (polynomial time, and approximation preserving) reduction from the gap problem $P_0(\varepsilon, k)$ (respectively, $P_1(\varepsilon, k)$) to the problem of approximating STP within r. (Strictly speaking, we do not construct this reduction; we only show that there exists one. But this clearly suffices for proving NP-hardness.) Hence (with everything above fixed, including k and D) we are ready to describe the reduction. Given an instance L of the problem $P_0(\varepsilon, k)$ (respectively, $P_1(\varepsilon, k)$) with n equations, whose optimum MAX of the maximal number of satisfiable equations is promised to be either at most $n(\frac{1}{2} + \varepsilon)$ or at least $n(1 - \varepsilon)$, the reduction described above produces the corresponding instance of the Steiner Tree problem. Since the assumptions of Lemma 12 are satisfied, the optimum OPT is achieved on a standard Steiner tree. Hence, using (3), the optimum OPT of the corresponding instance of the Steiner tree problem is

$$\text{OPT} = n\alpha + \frac{3}{2}nd + \frac{3}{2}n\beta + (n - \text{MAX})\gamma,$$

which has to be now either at least $n\alpha + \frac{3}{2}nd + \frac{3}{2}n\beta + n\left(\frac{1}{2} - \varepsilon\right)\gamma$, or at most $n\alpha + \frac{3}{2}nd + \frac{3}{2}n\beta + n\varepsilon\gamma$.

Hence even the partial decision subproblem of the STP, namely the problem to distinguish between these two cases, is NP-hard. Consequently, since

$$\frac{n\alpha + \frac{3}{2}nd + \frac{3}{2}n\beta + n\left(\frac{1}{2} - \varepsilon\right)\gamma}{n\alpha + \frac{3}{2}nd + \frac{3}{2}n\beta + n\varepsilon\gamma} = 1 + \frac{(1-4\varepsilon)\gamma}{2\alpha + 3d + 3\beta + 2\varepsilon\gamma} > r,$$

it is NP-hard to approximate the STP within r.

Moreover, it can be easily seen that if the gadget above is (uniformly) quasi-bipartite, our reduction produces (uniformly) quasi-bipartite instances of the STP, and the inapproximability results apply to those instances as well. □

Theorem 14. *Given an integer $d \ge 3$, let $q(d) = \min\left\{\frac{c(d)-1}{2c(d)}, \frac{1}{4}\right\}$, $r(d) = 1 + \frac{q(d)}{3(d+1-q(d))}$, where $c(d)$ is the constant defined in (2). Then for any constant r, $1 < r < r(d)$, it is* NP*-hard to approximate the optimal solution of the* Steiner Tree problem *within a factor r.*

In particular, since $c(6) > 1.76222$ implies $r(6) > 1.01063$, inapproximability within a factor 1.01063 ($> \frac{96}{95}$) follows for the STP, unless P = NP.

Proof. Let an integer $d \ge 3$ and a number r, $1 < r < r(d)$, be fixed. We can find $\gamma \le \frac{1}{4}$ with $\gamma < \frac{c(d)-1}{2c(d)}$ (i.e., $\frac{1}{1-2\gamma} < c(d)$) and such that $r < 1 + \frac{\gamma}{3(d+1-\gamma)}$, and apply Theorem 13 with the $(0, 1-\gamma, \gamma)$-gadget from Example 3 (with γ as above, $\alpha = 0$, and $\beta = 1 - \gamma$). □

Theorem 15. *Given an integer $d \ge 3$ and let $r(d) = 1 + \frac{c(d)-1}{6d\cdot c(d)+21c(d)-3}$, where $c(d)$ is the constant defined in (2). Then it is* NP*-hard to approximate the optimal solution of (uniformly) quasi-bipartite* Steiner Tree problem *within a factor r, for any r, $1 < r < r(d)$.*

In particular, since $c(7) > 1.94606$ implies $r(7) > 1.00791$, inapproximability within a factor 1.00791 ($> \frac{128}{127}$) follows for the (uniformly) quasi-bipartite Steiner Tree problem, *unless* P = NP.

Proof. Let an integer $d \geq 3$ and a number r, $1 < r < r(d)$, be fixed. We can find $\gamma < \frac{c(d)-1}{2c(d)}$ such that $r < 1 + \frac{\gamma}{3(d+3+\gamma)}$, and apply Theorem 13 with the uniformly quasi-bipartite $(3+3\gamma, 1-\gamma, \gamma)$-gadget from Example 4 (with γ as above, $\alpha = 3+3\gamma$, and $\beta = 1-\gamma$). □

Remark 16. The same inapproximability results as obtained in Theorem 14 can be proved for more special instances of the Steiner tree problem, for example, for the unweighted version of STP (when all edges have the same weight 1). To see that, we can clearly assume that γ in our gadget from Example 3 is rational, $\gamma = \frac{p}{q}\ (\leq \frac{1}{4})$ with integral p and q. Multiplying by q all weights in instances produced by our reduction, we obtain the equivalent problem on instances with weight of each edge $2p$, $q-2p$, or q. Now we replace each edge e with integral weight $w(e)$ by a path of $w(e)$ edges of weight 1 each, without changing the terminal set. Our inapproximability results translate to such instances in a straightforward way.

3. Expander graphs

Expander graphs play an important role in many constructions. They are useful in the design of sorting algorithms, and in constructions of various concentrators, superconcentrators, and connectors.

It is rather difficult to construct explicitly infinite families of (c, τ, d)-expanders with fixed (properly chosen) parameters $c > 1$, $d \geq 3$, and $\tau \in (0, 1)$. Such constructions were first given by Margulis [12], Gabber and Galil [7], and Lubotzky et al. [10]. However, in some applications also expanders, that we do not know how to construct efficiently, can be useful. For the purpose of this paper the *existence* of expanders with certain parameters is sufficient. For expanders, the existence of which is guaranteed by probabilistic methods, we can obtain better expansion parameters than for those constructed explicitly.

Random d-regular bipartite graphs

In what follows, d will be a fixed integer, $d \geq 3$. Consider an integer $k \geq d$, and two vertex sets, $V_1 = \{(1, j) : 1 \leq j \leq k\}$ and $V_2 = \{(2, j) : 1 \leq j \leq k\}$. The model of random labeled d-regular bipartite graphs with k by k bipartition (V_1, V_2) can be introduced as follows: For each vertex $(i, j) \in V_1 \cup V_2$ we consider a set $\mathcal{V}_{i,j}$ of d new vertices (the sets $\mathcal{V}_{i,j}$ being pairwise disjoint), and put $\mathcal{V}_i := \cup_{j=1}^{k} \mathcal{V}_{i,j}$ for $i = 1, 2$. Let $\mathcal{M}\ (=\mathcal{M}(k))$ be the set of all perfect matchings between $\mathcal{V}_1$ and $\mathcal{V}_2$. Clearly, $|\mathcal{M}| = (kd)!$. Consider $M \in \mathcal{M}$ randomly chosen, with the uniform probability distribution. Let G_M be the d-regular bipartite multigraph with bipartition (V_1, V_2), in which each $(1, r)$ and $(2, s)$ is joined by the same number of edges as $\mathcal{V}_{1,r}$ and $\mathcal{V}_{2,s}$ are matched by M. In other words, G_M is obtained from M by merging each set $\mathcal{V}_{i,j}$ into the vertex (i, j).

An important fact is that the portion of those M in $\mathcal{M}$ for which G_M is a simple graph is at least $\rho(d) > 0$ (for every sufficiently large k), and each such labeled simple graph (i.e., d-regular, bipartite, with bipartition (V_1, V_2)) corresponds to the same number of matchings $M \in \mathcal{M}$. Therefore the problem of proving that almost all d-regular bipartite graphs have some specified property reduces to proving such a result for almost all matchings in $\mathcal{M}(k)$, when $k \to \infty$.

Some authors alternatively use ordered d-tuples of perfect matchings between V_1 and V_2, instead of perfect matchings between $\mathcal{V}_1$ and $\mathcal{V}_2$, which leads essentially to the same computations.

The following theorem shows the existence of d-regular bipartite graphs with certain expansion parameters (see [5,2,15], and Theorem 6.6 in [6] for more details). In previous sections it was used to make our parametrized approximation lower bounds explicit.

Theorem 17. *Let $0 < \tau < \frac{1}{c} < 1$ be real numbers and d be an integer such that $d > \frac{H(\tau)+H(c\tau)}{H(\tau)-c\tau H(\frac{1}{c})}$, where $H(x) = -x \ln x - (1-x)\ln(1-x)$ for $x \in (0, 1)$. Then almost all random d-regular bipartite (multi)graphs are (c, τ, d)-expanders.*

Conclusion

The methods of this paper provide a new motivation for the study of bounds on expansion parameters of low degree graphs that provably exist. For our purposes we need not restrict ourselves to expanders that can be effectively constructed; the existence is enough. There is a substantial gap between the known upper and lower bounds for parameters of the best possible expanders. We believe that lower bounds on our expander constants $c(d)$ can be improved significantly. This would improve our inapproximability results. Another way to improve the results would be to provide the gadgets with better parameters.

Acknowledgement

The second author was partially supported by VEGA grant no. 1/3106/06.

References

[1] S. Arora, C. Lund, R. Motwani, M. Sudan, M. Szegedy, Proof verification an hardness of approximation problems, in: Proceedings of the 33rd Annual Symposium on Foundations of Computer Science, 1992, pp. 14–23.
[2] L.A. Bassalygo, Asymptotically optimal switching circuits, Problems in Information Transmission 17 (1981) 206–211.

[3] M. Bern, P. Plassmann, The Steiner problem with edge lengths 1 and 2, Information Processing Letters 32 (1989) 171–176.
[4] M. Chlebík, J. Chlebíková, Approximation hardness for small occurrence instances of NP-hard problems, in: Proceedings of the 5th Conference on Algorithms and Complexity, CIAC, in: LNCS, vol. 2653, 2003, pp. 152–164 (also ECCC Report TR02-73, 2002).
[5] F.R.K. Chung, On concentrators, superconcentrators, generalizers, and nonblocking networks, The Bell System Technical Journal 58 (1978) 1765–1777.
[6] F.R.K. Chung, Spectral Graph Theory, in: CBMS Regional Conference Series in Mathematics, American Mathematical Society, ISSN 0160-7642, ISBN 0-8218-0315-8, 1997.
[7] O. Gabber, Z. Galil, Explicit construction of linear-sized superconcentrators, Journal of Computer and System Sciences 22 (1981) 407–420.
[8] J. Håstad, Some optimal inapproximability results, Journal of ACM 48 (2001) 798–859.
[9] S. Hougardy, C. Gröpl, T. Nierhoff, H.J. Prömel, Approximation algorithms for the Steiner tree problem in graphs, in: X. Cheng, D.-Z. Du (Eds.), Steiner Trees in Industry, Kluwer Academic Publishers, 2001, pp. 235–279.
[10] A. Lubotzky, P. Phillips, P. Sarnak, Ramanujan graphs, Combinatorica 8 (1988) 261–277.
[11] R.M. Karp, Reducibility among combinatorial problems, in: Complexity of Computer Computations, in: Proc. Sympos. IBM Thomas J. Watson Res. Center, Yorktown Heights, NY, 1972, Plenum, New York, 1972, pp. 85–103.
[12] G.A. Margulis, Explicit construction of concentrators, Problems of Information Transmission 9 (1973) 325–332.
[13] C.H. Papadimitriou, S. Vempala, On the approximability of the traveling salesman problem, in: Proceedings of the 32nd ACM Symposium on the Theory of Computing, Portland, 2000.
[14] G. Robins, A. Zelikovsky, Improved Steiner tree approximation in graphs, in: Proceedings of the 11th Annual ACM-SIAM Symposium on Discrete Algorithms, 2000, pp. 770–779.
[15] P. Sarnak, Some applications of modular forms, in: Cambridge Tracts in Mathematics, vol. 99, Cambridge University Press, Cambridge, 1990.
[16] M. Thimm, On the approximability of the steiner tree problem, in: Proceedings of the 26th International Symposium (MFCS 2001 Mariánske Lázne, Czech Republic, 27–31 August, 2001), in: LNCS, vol. 2136, Springer, 2001, pp. 678–689; Theoretical Computer Science 295 (2003) 387–402 (journal version).