On the Approximability of the Traveling Salesman Problem

Christos H. Papadimitriou*
Computer Science Division,
U.C. Berkeley,
Berkeley, CA 94720.

Santosh Vempala†
Department of Mathematics,
Massachusetts Institute of Technology,
Cambridge, MA 02139.

Abstract

We show that the traveling salesman problem with triangle inequality cannot be approximated with a ratio better than $\frac{117}{116}$ when the edge lengths are allowed to be asymmetric and $\frac{220}{219}$ when the edge lengths are symmetric, unless P=NP. The best previous lower bounds were $\frac{2805}{2804}$ and $\frac{3813}{3812}$ respectively. The reduction is from Håstad's maximum satisfiability of linear equations modulo 2, and is nonconstructive.

1 Introduction

Despite a recent avalanche of better and better —occasionally optimal— approximability lower bounds for several NP-hard optimization problems, based on improved PCPs and reductions [2, 3, 8, 9, 11, 16], there has been very little progress on the traveling salesman problem (with triangle inequality, of course). The original reduction in [14] only proves MAXSNP-completeness of the special case of symmetric distances, each of which is either 1 or 2; no explicit lower bound is given, but it is clear that no ratio better than about 1.000001 can be obtained from that construction. This can be improved by more sophisticated methods to about $\frac{5381}{5380}$ [12], further improved to $\frac{3813}{3812}$ in [4]; this is presently the best known bound.

Even for the *asymmetric* traveling salesman problem the best known lower bound is $\frac{2805}{2804}$ [12]. For this latter problem, the best known *upper* bound is $\log n$ [7], so the miniscule lower bound is especially annoying.

*Supported in part by NSF ITR Grant CCR-0121555. Email:`christos@cs.berkeley.edu`
†Supported by NSF award CCR-0307536 and a Sloan foundation fellowship. Email:`vempala@math.mit.edu`

In this paper, we prove that the asymmetric traveling salesman problem cannot be approximated to a ratio smaller than $\frac{117}{116}$, and the symmetric traveling salesman problem cannot be approximated to a ratio smaller than $\frac{220}{219}$, thus improving the best known inapproximability bounds by more than an order of magnitude[1]. Our reduction starts from Håstad's maximum satisfiability of linear equations modulo 2 with three variables per equation [8], a very strong and useful lower bound. Our reduction is essentially *nonconstructive*, since it relies on a probabilistic construction of a graph with specialized expansion-like properties needed in our proof — the size of this graph, called a *pusher*, is finite, related to the number of occurrences of each variable in Håstad's proof, and thus the existence of a *deterministic* reduction is immediate. Deriving inapproximability thresholds (i.e., finding such reductions) is ideal ground for the probabilistic method since one cares about existence and not explicit construction. Another unusual feature of our reduction is that, in contrast to all previous related constructions, it has no "variable gadget" —that is, no graph in which the choice between two values for each variable is "centrally" decided— relying for consistency on a sophisticated method of checking the value of each literal against enough opposite literals; this should be a more widely useful method (see [6, 15] for two other applications of this idea, since the first appearance of the present proof in [13]).

2 The Gadgets

An instance of the asymmetric traveling salesman problem is a $n \times n$ matrix D of nonnegative distances satisfying the triangle inequality: $D(i,j) + D(j,k) \geq D(i,k)$ for all i, j, k (in the symmetric case, D is also symmetric). We wish to find among all permutations τ (called *tours*) of $\{1, \ldots, n\}$ the one that minimizes $\sum_{i=1}^{n} D_{\tau(i),\tau(i+1)}$ —here addition in subscripts is modulo n. In this paper we shall represent, for clarity of presentation, such instances in terms of an underlying directed graph with n nodes (undirected in the symmetric case) and nonnegative weights on the edges; the distances then are the shortest-path distances on this graph. Once an underlying graph is specified, we can extend our notion of a *tour* to include all closed walks in the graph that visit all nodes, some possibly many times. Any such walk can then be rendered as a permutation of the nodes, with no worse cost, by simply omitting previously visited nodes, and taking advantage of the definition of D as a shortest path matrix. In the sequel, a "tour" will refer to such a walk.

Our main result for the asymmetric traveling salesman problem is the following:

Theorem 2.1. *For every $\epsilon > 0$ it is NP-hard to approximate the asymmetric traveling salesman problem within ratio $\frac{117}{116} - \epsilon$.*

[1] in [13], we erroneously claimed that a similar construction yields better constants.

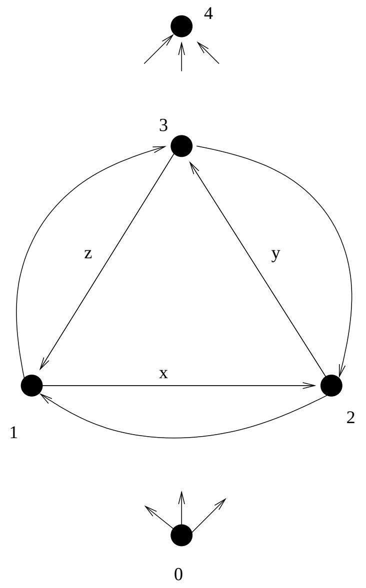


Figure 1: The equation gadget. All edges other than x, y, z have length one.

We start from the following important result, paraphrased from [8] to suit our purposes.

Theorem 2.2. *[8] For every $\epsilon > 0$ there is an integer k, depending on ϵ, such that it is NP-hard to tell whether a set of n linear equations modulo 2 with three variables per equation and with at most k occurrences of each variable has an assignment that satisfies $n(1-\epsilon)$ equations, or has no assignment that satisfies more than $n(\frac{1}{2}+\epsilon)$ equations.*

We shall assume that all equations are of the form $x + y + z = 0 \bmod 2$, where x, y, z are variables or negations, and that each variable appears the same number of times negated and unnegated (the former condition can be enforced by flipping some literals, and the latter by repeating each equation three more times with all possible pairs of literals negated).

Suppose then that we are given such a set of equations. We shall describe the instance of the asymmetric traveling salesman problem as a directed graph with edge weights. The graph, as usual, consists of several gadgets. The *equation gadget* is shown in Figure 1. There is one copy of this graph for each equation, with node 4 of one coinciding with node 0 of the next. The three edges labeled x, y, z correspond to the literals of the equation $x + y + z = 0 \bmod 2$. As we shall see, these three edges are in reality whole structures (see Figure 2); all other edges in the figure are real edges of cost one. The key property of this gadget is this:

Lemma 2.3. *For any subset S of the three labeled edges the following is true: There is a Hamilton path from 0 to 4 that traverses precisely S from among the three labeled edges x, y, z if and only if the cardinality of S is even. Otherwise, if the cardinality of S is odd, the shortest* *path from 0 to 4 that traverses all nodes, and precisely the edges in S among the labeled edges, has length 5.*

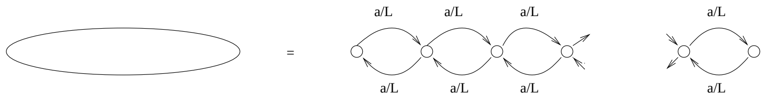
Proof. The first part (existence of Hamilton path) is evident: If S is empty then use (besides the edges from 0 and to 4) any two non-labeled edges connecting the nodes, and if $|S| = 2$ then take the two edges in S and no other edges. For the second part, if all three labeled edges are traversed then there is a cycle and thus a node repetition. If one labeled edge is used, notice that there is only one edge avoiding node repetition that can be used before the edge, and only one after (the edges from 0 and to 4); since this traversal would omit the third node, a node repetition is necessary. □

Thus, if we take a traversal of a labeled edge to mean that the corresponding literal is one, then a Hamilton path is tantamount to satisfaction.

The whole graph will consist of one equation gadget for each equation, connected in arbitrary order in tandem (with the 4 node of one identified with the 0 node of the next) and closing a cycle (the 4 node of the last is identified with the 0 node of the first). The intention is that tours will traverse these gadgets in order, traversing the true (according to a truth assignment) subset of the literals in each, wasting a weight of 1 whenever an equation is not satisfied. For any $0-1$ assignment to the literals, we define a *standard* tour as the one which visits the equation gadgets one by one, and within each gadget it uses a path that visits exactly the subset of edges that correspond to literals that are set to 1. If the equation is satisfied, the equation gadget is visited by a Hamilton path, otherwise with a repetition, as described in the lemma. The number of repeated nodes in the standard tour will thus capture the number of unsatisfied equations.

What we have described so far is the "skeleton" of the construction (and of the standard tour); the main action takes place within the labeled edges. Each labeled edge in each equation gadget is in fact a whole structure consisting of d undirected paths called *bridges* (shown as ellipses) connected by directed edges called *linking edges*. (d will soon be fixed to be 6; we are keeping several free parameters to demonstrate the generality of our construction and hopefully to help readers expose further improvements). The two blotted nodes in the figure are nodes of type 1,2, or 3 in the same equation gadget. The leftmost directed edge from a blotted node to a bridge, has length 1; all other linking edges have length b (Figure 2). We call this *the edge gadget*. Each bridge in the figure is in fact a bi-directed path with $L + 2$ edges, each of length a/L. Here L is a large integer; a and b will be eventually fixed to 4 and 2 respectively. A standard tour corresponding to an assignment is then exactly as described above, with labeled edges being replaced by the corresponding directed paths. So far, we have not ensured consistency among the occurrences of a variable, i.e., it is possible that the optimum tour of the graph is not a

An edge gadget consists of d bridges.



Each bridge is a bi−directed path of L+2 edges.

Figure 2: The edge gadget

standard tour — it might traverse one occurrence of a literal and avoid another — and hence there is no corresponding truth assignment.

To ensure consistency we need a further refinement of the construction (which will not affect a standard tour). Consider a variable x that appears k times negated and k times unnegated. The conclusion of the construction requires the *pairwise identification* of the bridges in the edge structures, so that each bridge in a positive edge gadget is identified with a bridge in a negative edge. The two bridges are identified so that each of the two endpoints of the resulting bridge has an entering edge and an exiting edge. Once the bridges have been identified, the neighborhood of each edge gadget is as shown in Figure 3. The "plan" for the identification is provided by a particular $k \times k$ d-regular bipartite multigraph $X = (V_1, V_2, E)$ with a special expander-like property, which we call a b-*pusher* (definition 1 below). For each variable x, we consider the b-pusher X, and identify the unnegated occurrences of x with V_1 and the negated occurrences with V_2, and we identify two bridges if the corresponding occurrences are connected by an edge in X. This concludes the description of the construction. (Even though it is not essential to our construction, it is easy to see, that the graph can have the orderly structure shown in Figure 3 with bridges divided into layers, and directed edges leading from one layer to the next; this follows from the fact that d-regular bipartite graphs are d-edge-colorable.)

Definition 1. *A d-regular bipartite graph $G = (V_1, V_2, E)$ is called a b-*pusher *if for any partition of V_1 into subsets U_1, S_1, T_1 and V_2 into U_2, S_2, T_2, such that there are no edges from vertices in U_1 to vertices in U_2, the number (T_1, T_2) of edges between vertices in T_1 and T_2 satisfies*

$$(b + \frac{1}{2})(T_1, T_2) \geq \min\{|U_1| + |T_2|, |U_2| + |T_1|\} - (b - \frac{1}{2})(|S_1| + |S_2|).$$

Theorem 5.1 shows the existence of such graphs for useful values of d and b.

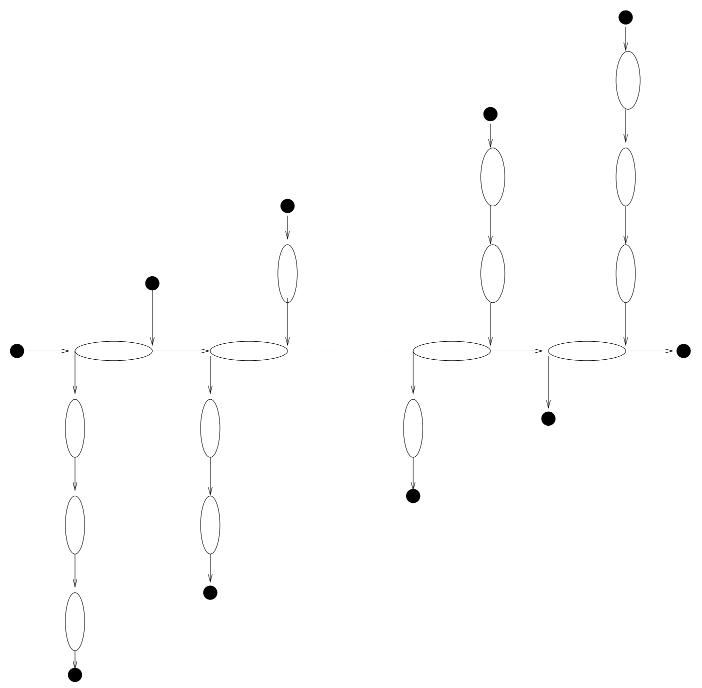


Figure 3: The neighborhood of an edge gadget

3 The Proof

Suppose that there is a truth assignment in the given set of equations that satisfies all but F of the equations. We have described tours as Eulerian walks of the underlying directed graph; the cost of a tour is then the sum of the weights of the edges in the walk, with the cost an edge added as many times as the edge is traversed. We claim that the cost of the standard tour is

$$\left[\frac{3n}{2} \cdot \left(d\frac{L+2}{L}a + db\right)\right] + [4n + F].$$

The first bracketed term is the cost of traversing all the $\frac{3dn}{2}$ bridges, each at cost $\frac{L+2}{L}a$ and one linking edge per bridge, with each of these edges costing b. (The first of the linking edges of an edge gadget, when traversed, is charged to the corresponding equation gadget.) The second term comes from the cost of traversing the n equation gadgets at 4 per satisfied equation, and 5 per unsatisfied equation.

The main step of the proof is to show that the optimum tour is standard (for some assignment, this is Lemma 3.2 below).

Let us fix a tour τ. Its cost can be split into two parts, corresponding to the bracketed terms in the cost of the standard tour above: The *bridge cost* of τ is the total cost of traversing the edges of length b and a/L in the bridges, whereas the *equation cost* is the cost of traversing the edges of length one on the equation gadgets (including the first edge of every bridge). If an edge

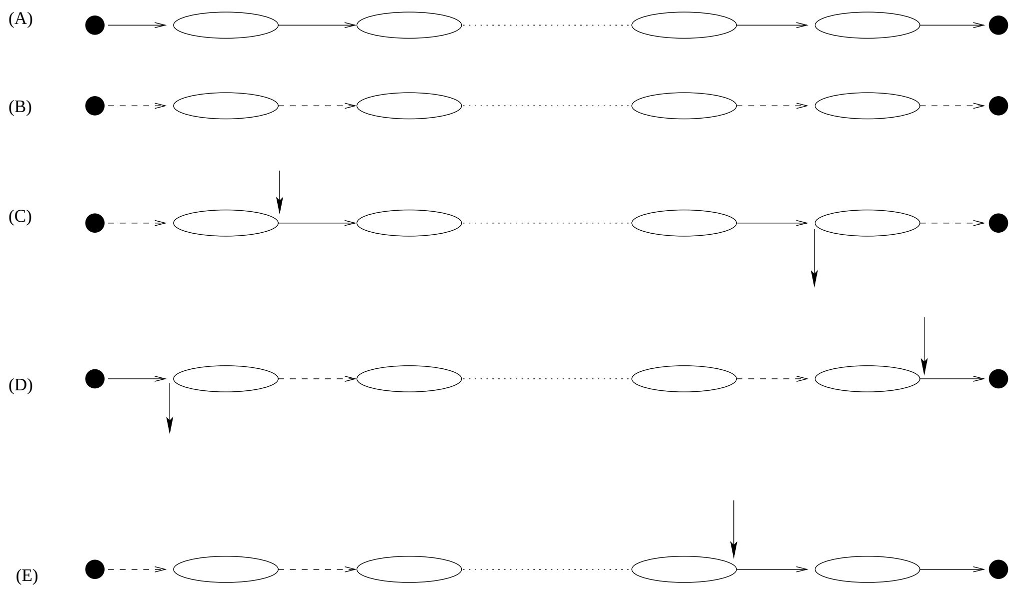


Figure 4: Traversals of an edge gadget

gadget has all of its linking edges traversed from left to right in τ, then we call the edge gadget *fully traversed* –and we think of the corresponding occurrence of a literal as true (Figure 4(A)). Otherwise, if none of these $d+1$ edges is traversed, we call the edge gadget *fully untraversed* —the occurrence is false (Figure. 4(B)). All other edge gadgets (and occurrences) are called *semitraversed*.

All semitraversed edge gadgets have this in common: There is at least one *reversal* somewhere in the edge gadget, that is, one linking edge is traversed and either the next or the previous one is not. We distinguish two types of semitraversed edge gadgets: (i) If the first and last linking edges are both untraversed, then the semitraversed edge gadget is of *type* U (e.g., Figure 4(C)). (ii) If the first and last linking edges are both traversed, then the semitraversed edge gadget is of *type* T (e.g., Figure 4(D)). (iii) Otherwise, it is of *type* S (e.g., Figure 4(E)). Note that types U and T have an even nonzero number of reversals, while type S has an odd number of reversals. The intuition for this classification is that vis-a-vis the equation gadget, type U and type T semi-traversed gadgets behave exactly like fully traversed and fully untraversed edge gadgets respectively, while type S gadgets might create more complex patterns by "entering" or "leaving" equation gadgets in the middle of the edge gadget.

A final species of waste is *double traversal*, which occurs when a particular bridge is traversed twice, once for the true value of its variable and once for the false.

Now, it is intuitively clear that reversals and double traversals are wasteful, in that the bridges or the linking edges must be traversed more intensely than in the standard tour in order

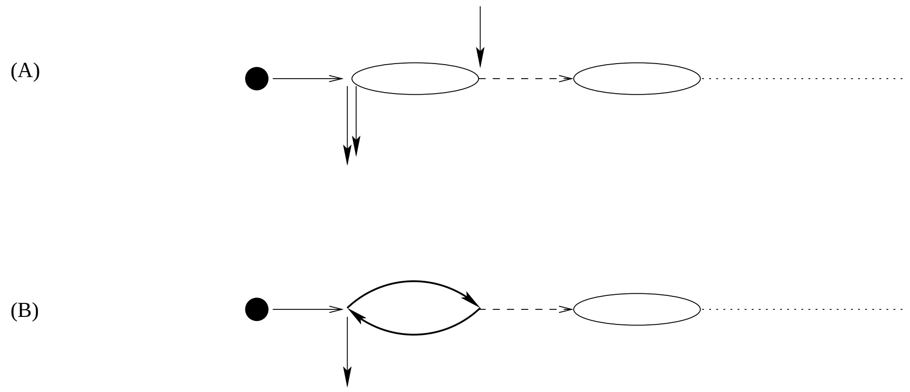


Figure 5: Cost of a reversal

to implement the reversal. To formalize this, we claim that there is a local cost of at least $\min(a/2, b)$ that we can assign to each reversal, and a local cost of $a+b$ to each double traversal:

Lemma 3.1. *The bridge cost of a tour is larger than that of the standard tour by at least $R \min(a/2, b) + D(a+b)$ where R is the number of reversals and D of double traversals.*

Proof. Consider a reversal, without loss of generality a bridge with an entering "horizontal" but no leaving horizontal edge (the other situation is analogous). Obviously, there is a vertical edge leaving the bridge. There are two cases: If there is a vertical edge entering the bridge (Figure 5(a)), then there is a locally assignable cost of b because of the second traversal of the vertical edge leaving the bridge. If there is no entering vertical edge (Figure 5(b)), then there is an extra cost of $(L+2)a/L$ in order for the tour to pick up the nodes in the bridge and come back to the left. However, only half of this cost can be assigned locally to this reversal —since the same bridge may also be the theater of a reversal in the vertical direction. Note that in the first case, when there is a vertical edge entering the bridge, there is no reversal happening in the vertical direction at this bridge; if a reversal happens earlier or later, it will incur a separate cost. Hence, every reversal can be assigned a local cost of at least $\min(a/2, b)$.

Double traversals are straightforward to account for: The edges of the bridge are traversed twice (extra cost at least a), and one of the linking edges leaving the bridge can be uniquely assigned to the double traversal. □

We can now prove the main result of this section:

Lemma 3.2. *There is an optimum tour that is a standard tour for some assignment.*

Proof. Let τ be any tour; we shall convert it, at no extra cost, to a standard tour.

Consider the bipartite graph with the positive occurrences of variable x on one side of the bipartition and the negative occurrences on the other side. Each vertex in this graph can be labeled as T, U, S^U, S^T or S^S depending on whether the corresponding edge gadget is traversed, untraversed or semitraversed and, in the latter case, on the type of the semitraversal. Let U_1 be the set of vertices labeled U on the left side of the bipartition. Similarly define $U_2, T_1, T_2, S_1^U, S_1^T, S_1^S, S_2^U, S_2^T$ and S_2^S and let u_1, t_1, etc. be the corresponding cardinalities of these sets. Without loss of generality assume that $u_1 + t_2 \leq u_2 + t_1$. Consider a modified tour where the occurrences on the U_1 side of the bipartition are all traversed and those on the U_2 are all untraversed, resulting in a tour where the variable x is traversed consistently. Do this for all variables, and then traverse the equation gadgets as in the standard tour for this assignment.

We will now show that the cost of the new tour is at most that of the old tour. There are two kinds of changes in the cost. First, the standard tour recovered certain wastefulness of the non-standard one in traversing the edge gadgets (the one captured by Lemma 3.1). Second, the standard tour may have incurred, in the worst case, some extra costs in traversing the equation gadgets (intuitively, by turning satisfied equations to unsatisfied ones). We shall argue, on a variable-by-variable basis, that the former changes compensate for the latter.

Consider a variable x and its sets U_1 etc. There are two kinds of wastefulness in the edge gadgets of this variable: First, since type U and type T semitraversals have at least two reversals per edge gadget and type S have at least one reversal, there are at least $2[s_1^U + s_1^T + s_2^U + s_2^T + \frac{1}{2}(s_1^S + s_2^S)]$ reversals in τ.

Second, suppose that in the bipartite graph X providing the connection plan there is an edge between a node in T_1 and a node in T_2. This means that the bridge corresponding to this edge is doubly traversed. Setting at this point $a = 2b$, the total waste related to x is, by Lemma 3.1

$$bR + 3bD \geq b[2s_1^U + 2s_1^T + 2s_2^U + 2s_2^T + s_1^S + s_2^S] + 3b(T_1, T_2), \tag{1}$$

where by (T_1, T_2) we denote the number of edges between T_1 and T_2 in X.

This saved waste will have to compensate for the increased cost of traversing the equation gadgets. Each edge gadget corresponding to an occurrence in the sets U_1, T_2 is traversed in a different manner in the modified tour, and therefore may increase the cost of traversing an equation from 4 to 5, i.e., the equation is unsatisfied in the final tour. Similarly for occurrences in the sets S_1^U, S_2^T. We claim that, besides this, at most one equation is lost for *every two semitraversed occurrences of type S*.

Consider a semitraversed edge gadget of x of type S. Call it f, and suppose w.l.o.g. that it goes between nodes labeled 1 and 2 in its equation gadget (Figure 1), and that its first linking edge is untraversed while its last is traversed. If the equation gadget already has local cost 5 in

τ (i.e., the cost of the outgoing edges from nodes labeled $0, 1, 2, 3$ plus any other incoming edges into node 4), then there is nothing to prove, since there is no increase in the cost. So assume that the local cost of the equation in the tour τ is 4. This implies that the nodes labeled $0, 1, 2, 3$ in the gadget have indegree and outdegree equal to one. Further we can assume that the other edge gadgets for this equation are of type U_2, T_1, S_1^T or S_2^U. Otherwise the cost of the equation has already been accounted for. We will now argue that there must be another semitraversed edge gadget of type S for the same equation gadget. This would imply our claim. Suppose not. Then the outgoing edge from 1 has to go to node 4 (since the first linking edge from on the edge gadget from 1 to 2 is assumed to be untraversed). The edge gadget between nodes labeled 2 and 3, and 3,1 must both be of type T_1 or S_1^T since the outgoing edges from $2, 3$ cannot go to 4. But then all three nodes $1, 2, 3$ have indegree 1 already and the outgoing edge from 0 has to go to one of them, which would make one of them have indegree two, and hence outdegree two. This contradicts the assumption that the local cost of the equation is only 4.

We conclude that the total cost of the extra equations lost because of the change in the tour is at most the sum, over all variables x, of

$$u_1 + t_2 + s_1^U + s_2^T + \frac{1}{2}(s_1^S + s_2^S). \tag{2}$$

It remains to prove that the RHS of (1) dominates (2). This is equivalent to showing that

$$3b(T_1, T_2) \geq u_1 + t_2 - 2b(s_1^T + s_2^U) - b(s_1^U + s_2^T) - (b - \frac{1}{2})(s_1^S + s_2^S).$$

Let $s_1 = s_1^U + s_1^T + s_1^S$ and $s_2 = s_2^U + s_2^T + s_2^S$. Clearly it would suffice to show that

$$3b(T_1, T_2) \geq u_1 + t_2 - (b - \frac{1}{2})(s_1 + s_2), \tag{3}$$

(Intuitively, this is saying that the worst case is when all the semitraversed edge gadgets are of type S).

Theorem 5.1 shows that almost all large enough d-regular graphs are b-pushers for $d = 6$ and $b = 2$. The defining property of 2-pushers implies (3). Thus the lemma follows. $\square$

Proof of Theorem 2.1. The proof of our main result is now a matter of accounting. From Håstad's theorem, we know that it is NP-hard to tell whether a set of n linear equations modulo 2 has an assignment that satisfies $n(1-\epsilon)$ linear equations or has no assignment that satisfies more than $n(\frac{1}{2}+\epsilon)$ equations. After applying our reductions, this corresponds to standard tours of length $(\frac{3}{2}d(a\frac{L+2}{L}+b)+4+\epsilon)n = (58+\frac{72}{L}+\epsilon)n$ and $(\frac{3}{2}d(a\frac{L+2}{L}+b)+\frac{9}{2}+\epsilon)n = (58.5+\frac{72}{L}+\epsilon)n$ respectively. By choosing L to be large enough, it follows that it is NP-hard to approximate the asymmetric traveling salesman problem to within $\frac{117}{116} - \epsilon$, for any $\epsilon > 0$. $\square$

4 The Symmetric TSP

We use the same basic ideas with a new set of gadgets suitable for the symmetric traveling salesman problem. Our main result is the following.

Theorem 4.1. *For every $\epsilon > 0$ it is NP-hard to approximate the symmetric traveling salesman problem within ratio $\frac{220}{219} - \epsilon$.*

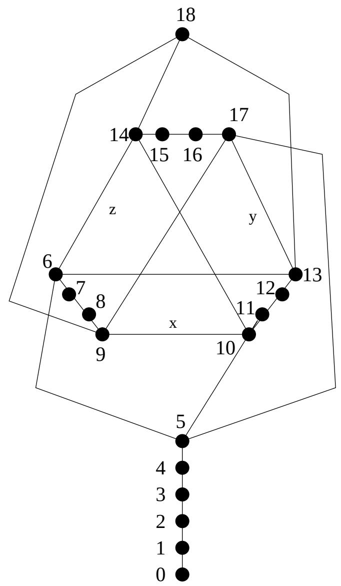
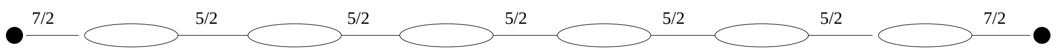


Figure 6: Equation gadget for the symmetric TSP

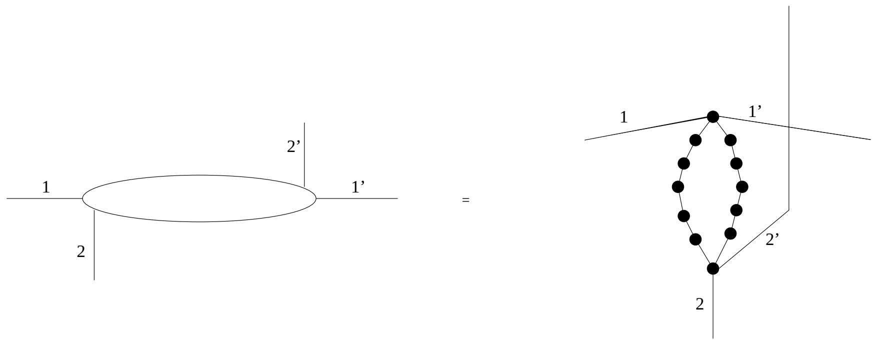
The equation gadget is shown in Figure 6; it is the asymmetric gadget, with each of the nodes labeled $1, 2, 3$ in Figure 1 replaced by an undirected path consisting of 4 nodes, one more than in Karp's original reduction from directed to undirected Hamilton cycle [10]. Node 0 is replaced by 6 nodes. The analog of Lemma 2.3 is this:

Lemma 4.2. *For any subset S of the three labeled edges the following is true: There is a Hamilton path from 0 to 18 that traverses precisely S from among the three labeled edges if and only if the cardinality of S is even. Otherwise, if the cardinality of S is odd, the shortest path from 0 to 18 that traverses the labeled edges in S and no other labeled edges, and visits all nodes, has length 20.*

Proof. Because of the four nodes replacing each node in the directed gadget, failing to traverse one such quadruple in the intended way incurs an extra cost of two — and hence in this case there is nothing to prove. Therefore, all quadruples are traversed as intended, and the argument reverts to the asymmetric case. □



An edge gadget consists of 6 bridges.



A bridge consists of 2 paths with L edges each; each edge has length 4/L.

Figure 7: Edge gadget for the symmetric TSP

In our construction we have an equation gadget for each equation, with all edge lengths equal to $\frac{1}{2}$; this way the difference between satisfied and unsatisfied equation gadgets is kept to 1.

The edge gadget is shown in Figure 7; the ovals in the figure are bridges, $2L$-node gadgets identical to the one shown, with each edge having length $4/L$. (Note that each oval now denotes a pair of paths *perpendicular* to the oval, in the sense that the linking edges 1 and $1'$ are incident upon the same node; likewise for 2 and $2'$.) The linking edges are of length $\frac{5}{2}$, except for the first and the last one, which have length $\frac{7}{2}$. The plan for identification of the bridges is again via the $k \times k$, 6-regular bipartite graph X.

There are two “standard” ways to traverse each bridge. One is by entering at 1 and exiting at $1'$ and the other is by entering at 2 and exiting at $2'$ (see Figure 7). The cost of traversing a bridge in one of the two standard ways is 8.

As before we define a standard tour as one that traverses all edge gadgets corresponding to true literals in the standard way, and visits the equations equation gadgets one by one, incurring an extra cost of 1 if the corresponding equation is unsatisfied. The cost of a standard tour that corresponds to a satisfying assignment with F unsatisfied equations is

$$\left[\frac{3n}{2}(6 \cdot 8 + 5 \cdot 5/2 + 2 \times (7/2 - 1/4))\right] + [9n + F] = \frac{219}{2}n + F.$$

The first bracketed term is again the bridge cost, and the second term is the equation cost, 9 per satisfied equation and 10 per unsatisfied equation. In the bridge cost, we add the cost of traversing the 6 bridges, the 5 inner linking edges, plus the two outer edges with cost 7/2; however, of that cost 1/4 is allocated to the equation gadget, so that the two combined make the edge gadget behave like an edge of length 1/2 (in the asymmetric case we instead had a first edge of length 1).

Nonstandard traversals by the pairs (1,2), (1,2′), (1′,2), (1′, 2′) of edges incur a minimum additional cost of 4 (we ignore the $O(1/L)$ terms that will be absorbed, as in the previous section, in the ϵ). Double traversals, in which 4 or more edges from among 1, 1′, 2, 2′ are traversed, cost an extra of at least 5/2 because of the extra traversals of these edges.

Notice that in the symmetric case, double traversals incur no extra cost within the bridge. Also, in allocating "local" costs to the bridges, we shall be thinking that the cost of extra traversals of a linking edge as being split equally between the two bridges it links. If the first or the last linking edge of an edge gadget is traversed more than once, the extra cost is allocated in its entirety to the adjacent bridge.

The only remaining case of non-standard traversal of a bridge is the one in which the tour enters and leaves a bridge from the same linking edge, say the edge 1. There are two cases. (a) The linking edge leads to another bridge: then that other bridge is bound to suffer an additional cost of at least 5/2 for the extra traversals of linking edges this will cause. (b) The doubly traversed linking edge leads to an equation gadget (and thus the bridge is the first or last in the edge gadget): the standard cost of traversing the linking edges once on each side (charged to the bridge) is $(\frac{7}{2} - \frac{1}{4}) + \frac{5}{4} = \frac{9}{2}$, which is the cost of the first linking edge ($\frac{7}{2}$) minus the portion $\frac{1}{4}$ which is charged to the equation gadget, plus half the cost of the other linking edge. The cost of the double traversal is $2 \times \frac{7}{2}$, all of which is charged to the bridge. Thus, the extra cost is $2 \times \frac{7}{2} - \frac{9}{2} = \frac{5}{2}$. Such a semitraversed edge (picked up by a linking edge that is connected to the equation gadget) is said to be of type P.

It follows that for each reversal we can assign to the bridges of the semitraversed edge gadget a local cost of at least 2, while semitraversals of type P cost 5/2. As we mentioned above, double traversals cost an extra 5/2:

Lemma 4.3. *The bridge cost of a tour is larger than that of the standard tour by at least $2R + \frac{5}{2}D + \frac{5}{2}P$ where R is the number of reversals in semitraversals other than those of type P, D is the number of double traversals, and P is the number of semitraversals of type P.*

Proof of Theorem 4.1. We show, as in the asymmetric case, that there is an optimum standard tour. The argument follows closely that for the asymmetric case. The expression for the local recovered waste of a tour at a variable gadget becomes now

$$4s_1^T + 4s_2^T + 4s_1^U + 4s_2^U + 2s_1^S + 2s_2^S + \frac{5}{2}(s_1^P + s_2^P)] + \frac{5}{2}(T_1, T_2), \tag{4}$$

where the superscript P denotes semitraversed edges of that type. The equation for the loss on the equation side is

$$u_1 + t_2 + s_1^U + s_2^T + s_1^P + s_2^P + \frac{1}{2}(s_1^S + s_2^S), \qquad (5).$$

Again (since the coefficient of s_1^P, s_2^P in (4) is $5/2$) the worse case is when we only have type-S semitraversals. Since this now reduces to the same inequality as in the asymmetric case, we can apply Theorem 5.1 to establish that (5) is dominated by (4). It follows that it is NP-hard to approximate the symmetric traveling salesman problem to a ratio smaller than $\frac{220}{219}$. $\square$

5 The Probabilistic Construction

In this section we prove the existence of bipartite multigraphs with the required property.

Theorem 5.1. *For k sufficiently large and any $b \geq 2$, almost every 6-regular bipartite graph on $2k$ vertices is a b-pusher.*

The proof relies on the following lemma.

Lemma 5.2. *For k sufficiently large, almost every 6-regular bipartite graph on $2k$ vertices has the following property: for any subset of vertices U contained entirely in V_1 or V_2, with $|U| = \alpha k$ and $|N(U)| = \beta k$,*

1. $0 \leq \alpha \leq .25 \implies \beta \geq 2\alpha.$
2. $.25 \leq \alpha \leq 0.5 \implies \beta \geq \alpha + .25$
3. $\alpha \geq 0.5 \implies \beta > .5\alpha + .5$

Proof of the Lemma. The probability that some subset U contained in V_1 or V_2 of size αk has at most βk neighbors is at most

$$P = \binom{k}{\alpha k}\binom{k}{\beta k}\frac{\binom{6\beta k}{6\alpha k}}{\binom{6k}{6\alpha k}}$$

Let $\tilde{P}$ the quantity obtained after using the approximation $\binom{r}{\gamma r} \simeq (\frac{1}{\gamma^\gamma(1-\gamma)^{1-\gamma}})^r$ in the above expression. (By $\simeq$ we mean that the logarithms of the two sides are within a $1 + o(1)$ factor). Then,

$$\frac{\log \tilde{P}}{k} = 5\alpha\log(\alpha) + 5(1-\alpha)\log(1-\alpha) - \beta\log(\beta) - (1-\beta)\log(1-\beta) - 6\alpha\log(\frac{\alpha}{\beta}) - 6(\beta-\alpha)\log(1-\frac{\alpha}{\beta})$$

If we substitute $\beta = 2\alpha$ in the above equation, we can check that the expression is non-positive for $0 \leq a \leq .25$ (the second derivative with respect to α is positive, and the endpoints can be checked). Similarly for $\beta = \alpha + .25$, and $\beta = .5\alpha + .5$ for the corresponding bounds. $\square$

Proof of Theorem 5.1. Let U_1, S_1, T_1 be a partition of V_1 and U_2, S_2, T_2 be a partition of V_2 and let $u_1, u_2, s_1, s_2, t_1, t_2$ denote their sizes. W.l.o.g. assume that $u_1 + t_2 \leq u_2 + t_1$. Let $T = (T_1, T_2)$ denote the number of edges between T_1 and T_2. We will show that

$$(b + \frac{1}{2})T \geq u_1 + t_2 - (b - \frac{1}{2})(s_1 + s_2). \tag{1}$$

Let $H = (b + \frac{1}{2})T - [u_1 + t_2 - (b - \frac{1}{2})(s_1 + s_2)]$. The following claim will be useful in the proof.

Claim 1. *There is a setting of u_i, s_i, t_i that minimizes H for which $T = 0$.*

Suppose $T > 0$. Then, there is an edge between w in T_1 and z in T_2. If we instead label z as being in S_2 or U_2, then T decreases and H does not increase.

Thus, we can assume that $T = 0$ at the minimum value of H. Our goal is to show that $H \geq 0$ for $b \geq 2$. Let $f(U_1), f(T_2)$ denote the number of neighbors of U_1, T_2 respectively. We consider the following cases.

Case A. $u_1, t_2 \leq 0.25$.

Then, using part (1) of Lemma 5.2, $f(U_1) \geq 2u_1, f(T_2) \geq 2t_2$ and so

$$\begin{aligned} H &\geq (b - \frac{1}{2})(f(U_1) - t_2 + f(T_2) - u_1) - (u_1 + t_2) \\ &\geq (b - \frac{3}{2})(u_1 + t_2) \geq 0 \quad \text{for } b \geq \frac{3}{2}. \end{aligned}$$

Case B. $u_1, t_2 \geq 0.25$. Using Lemma 5.2, this implies that $u_1, t_2 \leq 0.5$. Suppose $u_1 > 0.5$, then $f(U_1) > 0.75$ and so $u_2 < 0.25$; on the other hand $f(T_2) \geq 0.5$ and so $t_1 \leq 0.5$. This gives $u_2 + t_1 < 0.75 \leq u_1 + t_2$, a contradiction. Thus part (2) of the lemma is applicable in this case. Further,

$$u_1 + t_2 \leq u_2 + t_1 \leq 2 - f(U_1) - f(T_2) \leq 1.5 - u_1 - t_2$$

and so we get $u_1 + t_2 \leq 0.75$, and hence $s_1 + s_2 \leq 0.5$. Using this,

$$\begin{aligned} H &\geq (b - \frac{1}{2})(0.5) - (u_1 + t_2) \\ &\geq \frac{b}{2} - 1 \geq 0 \quad \text{for } b \geq 2. \end{aligned}$$

Case C. $u_1 \geq 0.25, t_2 \leq 0.25$. The remaining case ($u_1 \leq 0.25, t_2 \geq 0.25$) is identical to this one.

Here we can use part (1) of Lemma 5.2 to get that $f(T_2) \geq 2t_2$. There are two subcases.

C1. $u_1 \le 0.5$. By part (2) of Lemma 5.2, $f(U_1) \ge u_1 + 0.25$. Therefore,

$$\begin{aligned} H &\ge (b-\frac{1}{2})(f(U_1) - t_2 + \max\{0, f(T_2) - u_1\}) - (u_1 + t_2) \\ &\ge (b-\frac{1}{2})(u_1 + 0.25 - t_2) + (b-\frac{1}{2})\max\{0, 2t_2 - u_1\} - (u_1 + t_2) \end{aligned}$$

Now if $u_1 \ge 2t_2$, then

$$\begin{aligned} H &\ge 0.25(b-\frac{1}{2}) + (b-\frac{3}{2})u_1 - (b+\frac{1}{2})t_2 \\ &\ge 0.25(b-\frac{1}{2}) - (\frac{7}{2} - b)t_2 \\ &\ge \min\{\frac{b}{2} - 1, 0.25(b-\frac{1}{2})\} \end{aligned}$$

(based on whether $b \le 7/2$ or $b > 7/2$). If $u_1 < 2t_2$

$$\begin{aligned} H &\ge 0.25(b-\frac{1}{2}) + (b-\frac{3}{2})t_2 - u_1 \\ &\ge 0.25(b-\frac{1}{2}) - (\frac{7}{4} - \frac{b}{2})u_1 \\ &\ge \min\{\frac{b}{2} - 1, 0.25(b-\frac{1}{2})\}. \end{aligned}$$

In both subcases, $H \ge 0$ for $b \ge 2$.

C2. $u_1 > 0.5$. By part (3) of Lemma 5.2 $f(U_1) \le u_1/2 + 1/2$. Also,

$$u_1 + t_2 \le u_2 + t_1 \le 2 - u_1 - f(U_1) \le \frac{3}{2} - \frac{3u_1}{2} \le \frac{3}{4}.$$

Therefore,

$$\begin{aligned} H &\ge (b-\frac{1}{2})(f(U_1) - t_2) - (u_1 + t_2) \\ &\ge (b-\frac{1}{2})(\frac{u_1}{2} + \frac{1}{2} - t_2) - (u_1 + t_2) \\ &\ge \frac{b}{2} - \frac{1}{4} - (\frac{5}{4} - \frac{b}{2})u_1 - (b+\frac{1}{2})t_2. \end{aligned}$$

If $b \le 5/2$, then

$$\begin{aligned} H &\ge \frac{b}{2} - \frac{1}{4} - (\frac{5}{4} - \frac{b}{2})(\frac{3}{4} - t_2) - (b+\frac{1}{2})t_2 \\ &\ge \frac{7b}{8} - \frac{19}{16} - (\frac{3b}{2} - \frac{3}{4})t_2 \ge \frac{b}{2} - 1. \end{aligned}$$

If $b > 5/2$, then

$$H \ge \frac{b}{2} - \frac{1}{4} - (b+\frac{1}{2})t_2 \ge \frac{b}{4} - \frac{3}{8}.$$

Again, $H \ge 0$ for $b \ge 2$.

This completes the proof. □

6 Discussion

The methods and bounds of this paper provide a new motivation for the study of random graphs. In particular, better bounds for the random constructions would directly imply improved lower bounds for the approximability of the traveling salesman problem. The parameters we use here ($d = 6$ and $b = 2$) are the best we found for all values of d between 4 and 7, and we are convinced that the value of the $b \cdot d$ product, 12, is very close to its optimum value.

The immense intricacies of proving a rather modest lower bound for the asymmetric TSP, suggest that a constant approximation ratio, at least for the directed graph TSP, should be possible.

Acknowledgment. We are grateful to Alex Fabrikant for writing a matlab program to check our computations for the probabilistic construction, and to Muli Safra for interesting discussions. We would like to thank two anonymous referees for their useful criticism and for catching an error in the earlier version.

References

[1] N. Alon, "On the edge-expansion of graphs," Combinatorics, Probability and Computing, 6, 1997, 145-152.

[2] S. Arora, C. Lund, R. Motwani, M. Sudan, and M. Szegedy, "Proof verification and hardness of approximation problems," JACM 45(3), 1998, 501-555.

[3] P. Berman and M. Karpinski, "On some tighter inapproximability results," Proc. of ICALP 1999, 200-209.

[4] H.-J. Bckenhauer, S. Seibert, "Improved Lower Bounds on the Approximability of the Traveling Salesman Problem," RAIRO Theoretical Informatics and Applications 34, 2000, 213-255.

[5] B. Bollobás, "The isoperimetric number of random regular graphs," Europ. J. of Combinatorics 9, 241-244, 1988.

[6] L. Engebretsen and M. Karpinski "Approximation Hardness of TSP with Bounded Metrics," Proc. of ICALP 2001, 201-212.

[7] A. Frieze, G. Galbiati, and F. Maffioli, "On the worst-case performance of some algorithms for the asymmetric traveling salesman problem," Networks 12, 1982, 23-39.

[8] J. Håstad, "Some optimal inapproximability results," JACM 48(4), 2001, 798-859.

[9] J. Håstad, “Clique is hard to approximate within $n^{1-\epsilon}$,” Acta Mathematicа 182, 1999, 105-142.

[10] R. M. Karp, ” Reducibility Among Combinatorial Problems”, Complexity of Computer Computations, Plenum Press, 1972, 85-103.

[11] S. Khanna, M. Sudan, L. Trevisan and D.P. Williamson, “The Approximability of Constraint Satisfaction Problems,” SIAM J. Comput 30(6), 2000, 1863-1920.

[12] L. Engebretsen, “An Explicit Lower Bound for TSP with Distances One and Two,” Algorithmica, 35(4), 2003, 301-319.

[13] C. H. Papadimitriou and S. Vempala “On the approximability of the traveling salesman problem,” Proc. of STOC 2000, 126-133.

[14] C. H. Papadimitriou and M. Yannakakis, “The traveling salesman problem with distances one and two,” Math of OR 18(1), 1993, 1-11.

[15] M. Thimm, “On the Approximability of the Steiner Tree Problem,” MFCS 2001, 678-689.

[16] L. Trevisan, G. Sorkin, M. Sudan, D. Williamson, “Gadgets, approximation and linear programming,” SIAM J. Comput. 29(6), 2000, 2074-2097.