Improved Approximation Lower Bounds for TSP with Distances One and Two*

Ryo HIRADE and Toshiya ITOH

Global Scientific Information and Computing Center, Tokyo Institute of Technology,
2-12-1 O-okayama, Meguro-ku, Tokyo 152-8550, Japan

Received May 14, 2001; final version accepted November 2, 2001

The metric traveling salesman problem $\triangle$ -TSP is the traveling salesman problem in which the distances among cities satisfy the triangle inequality. In this paper, we consider the metric traveling salesman problem $\triangle(1, 2)$ -TSP with distances one and two and $\triangle(1, 2, 3)$ -TSP with distances one, two, and three as the special cases of $\triangle$ -TSP. Since $\triangle(1, 2)$ -TSP is $\mathcal{NP}$ -complete, it is $\mathcal{NP}$ -hard to find an optimal solution for $\triangle(1, 2)$ -TSP. So in polynomial time, we wish to find an approximate solution for $\triangle(1, 2)$ -TSP. However $\triangle(1, 2)$ -TSP is $\mathcal{APX}$ -complete, there is a nontrivial approximation lower bound for $\triangle(1, 2)$ -TSP. For any $\varepsilon > 0$, Engebretsen showed that it is $\mathcal{NP}$ -hard to approximate the symmetric $\triangle(1, 2)$ -TSP within $5381/5380 - \varepsilon$; the asymmetric $\triangle(1, 2)$ -TSP within $2805/2804 - \varepsilon$, and Böchenhauer, et al. showed that it is $\mathcal{NP}$ -hard to approximate the symmetric $\triangle(1, 2, 3)$ -TSP within $3813/3812 - \varepsilon$. In this paper, we improve those lower bounds and show that for any $\varepsilon > 0$, it is $\mathcal{NP}$ -hard to approximate the symmetric $\triangle(1, 2)$ -TSP within $1027/1026 - \varepsilon$ (Corollary 4.5); the asymmetric $\triangle(1, 2)$ -TSP within $535/534 - \varepsilon$ (Corollary 4.7); the symmetric $\triangle(1, 2, 3)$ -TSP within $817/816 - \varepsilon$ (Theorem 5.2); the asymmetric $\triangle(1, 2, 3)$ -TSP within $364/363 - \varepsilon$ (Theorem 5.3). We also show that for any $\varepsilon > 0$, it is $\mathcal{NP}$ -hard to approximate SHORTEST-SUPERSTRING within $1279/1278 - \varepsilon$ (Corollary 6.3).

KEYWORDS: TSP, E3-LIN, Approximation Algorithms, $\mathcal{APX}$ -complete, SHORTEST-SUPERSTRING

1 Introduction

1.1 Background

Let $C = \{c_1, c_2, \dots, c_n\}$ be the set of cities and $d: C \times C \rightarrow \mathbf{R}^+$ be the distance function between each pair of cities in C , i.e., $d(c_i, c_j)$ is the distance of any pair of cities $c_i, c_j \in C$ and $d(c, c) = 0$ for any $c \in C$, where $\mathbf{R}^+$ is the nonnegative reals. The function d is *symmetric* if $d(c_i, c_j) = d(c_j, c_i)$ for any pair of cities $c_i, c_j \in C$; *asymmetric* otherwise. Let $\triangle$ -TSP = $\{(C, d)\}$ be the *metric traveling salesman problem* for which the function d satisfies the triangle inequality and we use $\triangle(1, 2)$ -TSP = $\{(C, d)\}$ (resp. $\triangle(1, 2, 3)$ -TSP = $\{(C, d)\}$) to denote the metric traveling salesman problem with *distances one and two* (resp. *distances one, two, and three*) for which d satisfies the triangle inequality and $d(c_i, c_j) \in \{1, 2\}$ (resp. $d(c_i, c_j) \in \{1, 2, 3\}$) for any pair of distinct cities $c_i, c_j \in C$. Note that $\triangle(1, 2)$ -TSP arises in various applications such as an mapping parallel programs to parallel architectures [17] and code optimization [3]. It is obvious that $\triangle(1, 2)$ -TSP and $\triangle(1, 2, 3)$ -TSP are the special cases of $\triangle$ -TSP. The goal is to find an optimal solution for $\triangle(1, 2)$ -TSP, i.e., a minimum weight tour that visits every city exactly once. For any $(C, d) \in \triangle(1, 2)$ -TSP, we use $opt(C, d)$ to denote the weight of the optimal tour on (C, d) .

It is $\mathcal{NP}$ -hard to compute an optimal solution for $\triangle(1, 2)$ -TSP [14], which implies that there is little hope to efficiently compute an optimal solution for $\triangle(1, 2)$ -TSP (and for $\triangle(1, 2, 3)$ -TSP and $\triangle$ -TSP). So it is interesting to find an approximate solution for $\triangle$ -TSP, i.e., given $(C, d) \in \triangle$ -TSP, find in polynomial time a tour on (C, d) with weight as close to $opt(C, d)$ as possible.

1.2 Previous Works

We say that $\triangle$ -TSP = $\{(C, d)\}$ is symmetric (resp. asymmetric) if the function d is symmetric (resp. asymmetric). As for $\triangle$ -TSP, Christofides [8] showed that the symmetric $\triangle$ -TSP is approximable within $3/2$ and Frieze, Galbiati, and Maffioli [12] showed that the asymmetric $\triangle$ -TSP is approximable within $O(\log n)$. As for $\triangle$ -TSP, Papadimitriou and Yannakakis [16] showed that the symmetric $\triangle(1, 2)$ -TSP is approximable within $7/6$ and Vishwanathan [20] proved that the asymmetric $\triangle(1, 2)$ -TSP is approximable within $17/12$. On the other hand, Papadimitriou and Yannakakis [16] showed that $\triangle(1, 2)$ -TSP is $\mathcal{APX}$ -complete, which implies that any of $\triangle(1, 2)$ -TSP, $\triangle(1, 2, 3)$ -TSP, and $\triangle$ -TSP do not have a polynomial time approximation scheme unless $\mathcal{P} = \mathcal{NP}$ [1]. So there exists a nontrivial approximation lower bound for $\triangle(1, 2)$ -TSP, $\triangle(1, 2, 3)$ -TSP, and $\triangle$ -TSP unless $\mathcal{P} = \mathcal{NP}$.

Let $X = \{x_1, x_2, \dots, x_n\}$ be the set of Boolean variables and $E[X] = \{e_1, e_2, \dots, e_m\}$ be the system of linear equations on X . For an integer $k \geq 1$, we say that $E[X] \in \text{Ek-LIN}$ if $E[X]$ is the system of linear equa-

* A preliminary version of this work appeared in the IEICE Technical Report, COMP99-84, pp. 41–48 (March 2000).

tions over $GF(2)$ such that e_j is $x_{j_1} + x_{j_2} + \dots + x_{j_k} \equiv c_j \pmod{2}$ for each $1 \leq j \leq m$, i.e., each $e_j \in E[X]$ has exactly k variables $x_{j_i} \in X$. For integers $k \geq 1$ and $l \geq 2$, we also say that $E[X] \in \text{Ol-Ek-LIN}$ if $E[X] \in \text{Ek-LIN}$ and $x_i \in X$ occurs in $E[X]$ exactly l times for each $1 \leq i \leq n$. The goal is to find an assignment $a \in \{0, 1\}^n$ to the Boolean variables X under which $e_j \in E[X]$ holds as many as possible. Håstad [13] showed that for any $\varepsilon > 0$, it is $\mathcal{NP}$ -hard to approximate E2-LIN within $12/11 - \varepsilon$ and E3-LIN within $2 - \varepsilon$. For the special case of E2-LIN, Berman and Karpinski [6] showed that for any $\varepsilon > 0$, it is $\mathcal{NP}$ -hard to approximate O3-E2-LIN within $332/331 - \varepsilon$.

For $\Delta(1, 2)$ -TSP, Engebretsen [10] reduced O3-E2-LIN to $\Delta(1, 2)$ -TSP *deterministically* by giving an *equation gadget* for each $e_j \in E[X]$ of two variables and a *variable cluster* for each $x_i \in X$ and showed that for any $\varepsilon > 0$, it is $\mathcal{NP}$ -hard to approximate the symmetric $\Delta(1, 2)$ -TSP within $5381/5380 - \varepsilon$ and the asymmetric $\Delta(1, 2)$ -TSP within $2805/2804 - \varepsilon$. For $\Delta(1, 2, 3)$ -TSP, Böchenhauer, et al. [4] reduced O3-E2-LIN to $\Delta(1, 2, 3)$ -TSP *deterministically* (in a way similar to that by Engebretsen [10]) and showed that for any $\varepsilon > 0$, it is $\mathcal{NP}$ -hard to approximate the symmetric $\Delta(1, 2, 3)$ -TSP within $3813/3812 - \varepsilon$. These are the first explicit approximation lower bounds for the symmetric and asymmetric $\Delta(1, 2)$ -TSP and $\Delta(1, 2, 3)$ -TSP which are also for Δ -TSP, however, we still have a gap to the approximation upper bounds for $\Delta(1, 2)$ -TSP or Δ -TSP. For Δ -TSP, Papadimitriou and Vempala [15] reduced E3-LIN to Δ -TSP *probabilistically* by constructing an equation gadget for each $e_j \in E[X]$ of three variables and no variable cluster and showed that for any $\varepsilon > 0$, it is $\mathcal{NP}$ -hard to approximate the symmetric Δ -TSP within $129/128 - \varepsilon$ and the asymmetric Δ -TSP within $41/40 - \varepsilon$. Note that the lower bounds given by Papadimitriou and Vempala [15] improve the lower bounds for Δ -TSP directly derived from those by Engebretsen [10] or Böchenhauer, et al. [4], however, they do not necessarily imply the lower bounds for either $\Delta(1, 2)$ -TSP or $\Delta(1, 2, 3)$ -TSP. In fact, the instance reduced from E3-LIN by Papadimitriou and Vempala [15] is in Δ -TSP but not in $\Delta(1, 2)$ -TSP or $\Delta(1, 2, 3)$ -TSP.

1.3 Our Methods

To derive the explicit approximation lower bound for $\Delta(1, 2)$ -TSP, Engebretsen [10] reduced O3-E2-LIN to $\Delta(1, 2)$ -TSP as follows: Let $E[X] \in \text{O3-E2-LIN}$ and construct an equation gadget $2EG_j$ for each $e_j \in E[X]$ of two variables and a variable cluster VC_i for each $x_i \in X$. Note that $2EG_j$ consists of 12 (resp. 5) vertices for the symmetric (resp. asymmetric) case and that VC_i consists of 6 (resp. 5) vertices for the symmetric (resp. asymmetric) case. Connect $2EG_j$'s and VC_i 's together in the appropriate way and let $(C, d) \in \Delta(1, 2)$ -TSP. An assignment $a \in \{0, 1\}^n$ to X naturally specifies a subtour T_j^a on $2EG_j$. The equation gadget $2EG_j$ has the property that there exists a subtour T_j^a on $2EG_j$ visiting all cities in $2EG_j$ by traversing only edges of weight one if $e_j \in E[X]$ holds under the assignment $a \in \{0, 1\}^n$; a subtour T_j^a on $2EG_j$ visiting all cities in $2EG_j$ by traversing edges of weight one and a single edge of weight two otherwise. Thus there exist e equations in $E[X]$ that do not hold under an assignment $a \in \{0, 1\}^n$ iff there exists a tour on $(C, d) \in \Delta(1, 2)$ -TSP of weight $\|C\| + e$.

We say that $E[X] \in \text{O3-E2/3-LIN}$ if there exists a partition of $E[X] = E_2[X] \cup E_3[X]$ such that $E_2[X] \in \text{E2-LIN}$ and $E_3[X] \in \text{E3-LIN}$ and each $x_i \in X$ occurs in $E[X]$ exactly three times. The goal is to find an assignment $a \in \{0, 1\}^n$ to X under which $e_j \in E[X]$ holds as many as possible. Berman and Karpinski [7] showed that for any $\varepsilon > 0$, it is $\mathcal{NP}$ -hard to approximate O3-E2/3-LIN within $62/61 - \varepsilon$. To reduce O3-E2/3-LIN to $\Delta(1, 2)$ -TSP, we need to construct a new equation gadget for $e_j \in E_3[X]$ of three variables. In this paper, we shall construct an equation gadget $3EG_j$ for $e_j \in E_3[X]$ with the same property as $2EG_j$ by Engebretsen [10]. In fact, our equation gadget $3EG_j$ consists of 27 (resp. 12) vertices for the symmetric (resp. asymmetric) case. It should be noted that the equation gadget $3EG_j$ is completely different from that by Papadimitriou and Vempala [15] and can be easily modified to derive the approximation lower bounds for $\Delta(1, 2, 3)$ -TSP (as the way be

Table 1: Summary of the Approximation Lower Bounds

	Previous	Ours
Symmetric $\Delta(1, 2)$ -TSP	$\frac{5381}{5380} \approx 1.00019$ [10]	$\frac{1027}{1026} \approx 1.00097$ (Corollary 4.5)
Asymmetric $\Delta(1, 2)$ -TSP	$\frac{2805}{2804} \approx 1.00036$ [10]	$\frac{535}{534} \approx 1.00019$ (Corollary 4.7)
Symmetric $\Delta(1, 2, 3)$ -TSP	$\frac{3813}{3812} \approx 1.00026$ [4]	$\frac{817}{816} \approx 1.00012$ (Theorem 5.2)
Asymmetric $\Delta(1, 2, 3)$ -TSP	—	$\frac{364}{363} \approx 1.00027$ (Theorem 5.3)
SHORTEST-SUPERSTRING	$\frac{6721}{6720} \approx 1.00015$ [5, 10]	$\frac{1279}{1278} \approx 1.00078$ (Corollary 6.3)

Böchenhauer, et al. [4]).

1.4 Main Results

In this paper, we improve the approximation lower bounds for the symmetric and asymmetric $\triangle$ -TSP and $\triangle(1, 2, 3)$ -TSP. Table 1 summarizes the results obtained in this paper. We note that SHORTEST-SUPERSTRING is the minimization problem. By the reduction from the asymmetric $\triangle(1, 2)$ -TSP of bounded degree, Blum, et al. [5] showed that SHORTEST-SUPERSTRING is $\mathcal{APX}$ -complete. Since our reduction from O3-E2/3-LIN (or from O3-E2-LIN [10]) to $\triangle(1, 2)$ -TSP constructs the asymmetric $\triangle(1, 2)$ -TSP of bounded degree, the approximation lower bound above for SHORTEST-SUPERSTRING immediately follows from that for the asymmetric $\triangle(1, 2)$ -TSP (further details on SHORTEST-SUPERSTRING will be given in Section 6).

2 Preliminaries

To describe an instance (C, d) of $\triangle(1, 2)$ -TSP, we construct a graph $G = (C, E_d)$ by specifying the edges of weight one, i.e., for any pair of cities $c_i, c_j \in C$, $(c_i, c_j) \in E_d$ if $d(c_i, c_j) = 1$; $(c_i, c_j) \notin E_d$ otherwise. Note that the family of graphs $G = (C, E_d)$ completely describes the class of $\triangle(1, 2)$ -TSP. So we regard $G = (C, E_d)$ as an instance of $\triangle(1, 2)$ -TSP. To describe an instance (C, d) of $\triangle(1, 2, 3)$ -TSP, we also construct a graph $G = (C, E_d)$ by specifying the edges of weight one, i.e., for any pair of cities $c_i, c_j \in C$, $(c_i, c_j) \in E_d$ if $d(c_i, c_j) = 1$; $(c_i, c_j) \notin E_d$ otherwise. For any $c_i, c_j \in C$, let $d_0(c_i, c_j)$ be the distance of c_i and c_j along the edges of weight one in G . For the weight $w(c_i, c_j)$ of remaining edges (c_i, c_j) that do not explicitly appear in E_d , define $w(c_i, c_j) = \min \{3, d_0(c_i, c_j)\}$ for $(C, E_d) \in \triangle(1, 2, 3)$ -TSP. Note that the family of graphs $G = (C, E_d)$ describes only the subclass of $\triangle(1, 2, 3)$ -TSP, however, this is enough for our purpose (because we deal with the approximation lower bound for $\triangle(1, 2, 3)$ -TSP). Thus we regard $G = (C, E_d)$ as an instance of $\triangle(1, 2, 3)$ -TSP.

For any tour T on $G = (C, E_d)$, the tour T visits every vertex in C exactly once and can be viewed as the collection of edges in E_d . For an instance $G = (C, E_d)$ of $\triangle(1, 2)$ -TSP or $\triangle(1, 2, 3)$ -TSP and a tour T on G , let G^T be a subgraph of G with edges of G traversed by T , i.e., $G^T = (C, E_d \cap T)$. For any instance $G = (C, E_d)$ of $\triangle(1, 2)$ -TSP or $\triangle(1, 2, 3)$ -TSP and any tour T on G , we say that $c \in C$ is an *endpoint* with respect to T and G if the degree of c in G^T is *one*. We say that $c \in C$ is a *double endpoint* with respect to T and G if the degree of c in G^T is *zero* and we count it as two endpoints.

Let $X = \{x_1, x_2, \dots, x_n\}$ be the set of Boolean variables and $E[X] = \{e_1, e_2, \dots, e_m\}$ be the system of linear equations over $GF(2)$ on the set X of Boolean variables.

Definition 2.1 [13]: For any integer $k \geq 1$, we say that $E[X] \in E_k\text{-LIN}$ if $E[X]$ is the system of linear equations over $GF(2)$ such that e_j is $x_{j_1} + x_{j_2} + \dots + x_{j_k} \equiv c_j \pmod{2}$ for each $1 \leq j \leq m$.

In the sequel, we omit “(mod 2)” and simply write $x_{j_1} + x_{j_2} + \dots + x_{j_k} = 0$ and $x_{j_1} + x_{j_2} + \dots + x_{j_k} = 1$.

Definition 2.2 [6]: For any integers $k \geq 1$ and $l \geq 2$, we say that $E[X] \in Ol\text{-}E_k\text{-LIN}$ if $E[X] \in E_k\text{-LIN}$ and $x_i \in X$ occurs in $E[X]$ exactly l times for each $1 \leq i \leq n$.

Definition 2.3 [7]: We say that $E[X] \in O3\text{-}E2/3\text{-LIN}$ if there exists a partition $E[X] = E_2[X] \cup E_3[X]$ such that $E_2[X] \in E2\text{-LIN}$ and $E_3[X] \in E3\text{-LIN}$, and each $x_i \in X$ occurs in $E[X]$ exactly three times.

3 The Reduction to $\triangle(1, 2)$ -TSP

We reduce $E[X] = E_2[X] \cup E_3[X] \in O3\text{-}E2/3\text{-LIN}$ to the symmetric or asymmetric $\triangle(1, 2)$ -TSP. For each $e_j \in E_2[X]$, we construct the symmetric (resp. asymmetric) equation gadget $2EG_j$ in Figure 1 (resp. Figure 2), and for each $e_j \in E_3[X]$, we also construct the symmetric (resp. asymmetric) equation gadget $3EG_j$ in Figure 3 (resp. Figure 4). Note that each occurrence of a variable $x \in X$ corresponds to a vertex A_x in the equation gadgets. Since each variable $x \in X$ occurs exactly three times in $E[X]$, we have three vertices A_x corresponding to the variable x . For each variable $x \in X$, we link these three vertices A_x together by a variable cluster VC_x in Figure 5, which will force those vertices A_x to represent the same value for all three occurrences of the variable $x \in X$.

To construct an instance $G = (C, E_d)$ of the symmetric $\triangle(1, 2)$ -TSP from an instance $E_{2/3}[X]$ of $O3\text{-}E2/3\text{-LIN}$ (see Figure 6), we connect the equation gadgets $2EG_j$ and $2EG_{j+1}$ by letting $\tau_j = \sigma_{j+1}$, the equation gadgets $3EG_j$ and $3EG_{j+1}$ by letting $t_j = s_{j+1}$, and the variables clusters VC_i and VC_{i+1} by letting $\beta_i = \alpha_{i+1}$. Thus the vertices σ_j, s_j , and α_i have degree two. To construct an instance $G = (C, E_d)$ of the asymmetric $\triangle(1, 2)$ -TSP from an instance $E_{2/3}[X]$ of $O3\text{-}E2/3\text{-LIN}$ (see Figure 7), we connect the equation gadgets $2EG_j$ and $2EG_{j+1}$ by the edge $\tau_j \sigma_{j+1}$, the equation gadgets $3EG_j$ and $3EG_{j+1}$ by the edge $t_j s_{j+1}$, and the variables clusters VC_i and VC_{i+1} by the edge $\beta_i \alpha_{i+1}$. Thus the vertices σ_j, s_j , and α_i have indegree one, and the vertices τ_j, t_j , and β_i have outdegree one.

3.1 The Equation Gadget for E2-LIN

Let $E_2[X] \in E2\text{-LIN}$. It follows from Definition 2.1 that for each $1 \leq j \leq m_2$, $e_j \in E_2[X]$ is either of the form $x + y = 0$ or of the form $x + y = 1$. The following figures describe the symmetric and asymmetric equation gadgets $2EG_j$ for $e_j \in E_2[X]$ given by Engebretsen [10]. In Figures 1 and 2, the vertices A_x and A_y cor-

Figure 1: The Symmetric Equation Gadget $2EG_j = (C_j^2, E_j^2)$ for $e_j \in E_2[X]$. Figure 2: The Asymmetric Equation Gadget $2EG_j = (C_j^2, E_j^2)$ for $e_j \in E_2[X]$.

respond to the variables x and y , respectively, and are called Λ -vertices. The edges connected to Λ -vertices from below are called Λ -edges and the pair of Λ -edges connected to Λ_x (resp. Λ_y) is called the Λ -pair of Λ_x (resp. Λ_y). For any tour T on $2EG_j$, we say that a Λ -pair is *traversed* if both of the Λ -edges are traversed by T ; *untraversed* if none of the Λ -edges are traversed by T ; *semitraversed* otherwise. Note that the traversed Λ -pair of Λ_x (resp. Λ_y) corresponds to $x = 1$ (resp. $y = 1$) and the untraversed Λ -pair of Λ_x (resp. Λ_y) corresponds to $x = 0$ (resp. $y = 0$).

Lemma 3.1 [10, Lemma 1]: *Let T be a tour on the equation gadget $2EG_j$ in Figure 1 (resp. Figure 2) such that there are no semitraversed Λ -pairs. Then the following holds:*

- (a) *If there are zero or two traversed Λ -pairs in Figure 1-a (resp. Figure 2-a), it is possible to modify the tour T such that there are no endpoints in $2EG_j$ without increasing its weight and changing the way of traversing the Λ -pairs. Otherwise, it is possible to modify the tour T such that there are two endpoints in $2EG_j$ without increasing its weight and changing the way of traversing the Λ -pairs and it is impossible to construct a tour with less than two endpoints in $2EG_j$.*
- (b) *If there is exactly one traversed Λ -pairs in Figure 1-b (resp. Figure 2-b), it is possible to modify the tour T such that there are no endpoints in $2EG_j$ without increasing its weight and changing the way of traversing the Λ -pairs. Otherwise, it is possible to modify the tour T such that there are two endpoints in $2EG_j$ without increasing its weight and changing the way of traversing the Λ -pairs and it is impossible to construct a tour with less than two endpoints in $2EG_j$.*

Lemma 3.2 [10, Lemmas 2 & 3]: *Let T be a tour on the equation gadget $2EG_j$ in Figure 1 (resp. Figure 2) such that there are $k \in \{1, 2\}$ semitraversed Λ -pairs. Then it is possible to modify the tour T such that there are k endpoints in $2EG_j$ without increasing its weight and changing the way of traversing the Λ -pairs and it is impossible to construct a tour with less than k endpoints in $2EG_j$.*

3.2 The Equation Gadget for E3-LIN

Let $E_3[X] \in \text{E3-LIN}$. It follows from Definition 2.1 that for each $1 \leq j \leq m_3$, $e_j \in E_3[X]$ is either of the form $x + y + z = 0$ or of the form $x + y + z = 1$. In the following figures, we present the symmetric and asymmetric equation gadgets $3EG_j$ for $e_j \in E_3[X]$. In Figures 3 and 4, the vertices Λ_x , Λ_y , and Λ_z correspond to the variables x , y , and z , respectively, and are called Λ -vertices. For any tour T on $3EG_j$, let $\text{Config}(T; 3EG_j) \in \{t, u, s\}^3$ be the configurations of Λ -pairs of Λ_x , Λ_y , and Λ_z , where t , u , and s are traversed, untraversed, and semitraversed, respectively.

It is analogous to the equation gadget $2EG_j$ that in the equation gadget $3EG_j$, the traversed Λ -pairs of Λ_x , Λ_y , and Λ_z correspond to $x = 1$, $y = 1$, $z = 1$, respectively, and the untraversed Λ -pairs Λ_x , Λ_y , and Λ_z correspond to $x = 0$, $y = 0$, and $z = 0$, respectively.

Proposition 3.3: *For any tour T on the equation gadget $3EG_j$, the number of semitraversed Λ -pairs in $3EG_j$ with respect to T is odd iff the number of endpoints in $3EG_j$ with respect to T is odd.*

In the following, we analyze the behavior of the symmetric and asymmetric equation gadget $3EG_j$ with respect

Figure 3: The Symmetric Equation Gadget $3EG_j = (C_j^3, E_j^3)$ for $e_j \in E_3[X]$. Figure 4: The Asymmetric Equation Gadget $3EG_j = (C_j^3, E_j^3)$ for $e_j \in E_3[X]$.

to the number of untraversed Λ -pairs. In fact, we show that if the equation gadget $3EG_j$ has no semitraversed Λ -pairs, then the number of endpoints in $3EG_j$ is determined by the number of traversed Λ -pairs (Lemma 3.4) and if the equation gadget $3EG_j$ has $1 \leq k \leq 3$ semitraversed Λ -pairs, then the number of endpoints in $3EG_j$ is k regardless of the number of traversed Λ -pairs (Lemma 3.5 for $k = 1$, Lemma 3.6 for $k = 2$, and Lemma 3.7 for $k = 3$).

Lemma 3.4: *Let T be a tour on the equation gadget $3EG_j$ in Figure 3 (resp. Figure 4) such that there are no semitraversed Λ -pairs. Then the following holds:*

- (a) *If there are zero or two traversed Λ -pairs in Figure 3-a (resp. Figure 4-a), it is possible to modify the tour T such that there are no endpoints in $3EG_j$ without increasing its weight and changing the way of traversing the Λ -pairs. Otherwise, it is possible to modify the tour T such that there are two endpoints in $3EG_j$ without increasing its weight and changing the way of traversing the Λ -pairs and it is impossible to construct a tour with less than two endpoints in $3EG_j$.*
- (b) *If there are one or three traversed Λ -pairs in Figure 3-b (resp. Figure 4-b), it is possible to modify the tour T such that there are no endpoints in $3EG_j$ without increasing its weight and changing the way of traversing the Λ -pairs. Otherwise, it is possible to modify the tour T such that there are two endpoints in $3EG_j$ without increasing its weight and changing the way of traversing the Λ -pairs and it is impossible to construct a tour with less than two endpoints in $3EG_j$.*

Proof: We show the lemma for the symmetric case (the asymmetric case can be shown analogously).

Since no semitraversed Λ -pairs exist, we have the following cases: (Case 3.4.1) $\text{Config}(T; 3EG_j) = (u, u, u)$; (Case 3.4.2) $\text{Config}(T; 3EG_j) = (u, t, t)$; (Case 3.4.3) $\text{Config}(T; 3EG_j) = (t, u, t)$; (Case 3.4.4) $\text{Config}(T; 3EG_j) = (t, t, u)$; (Case 3.4.5) $\text{Config}(T; 3EG_j) = (t, u, u)$; (Case 3.4.6) $\text{Config}(T; 3EG_j) = (u, t, u)$; (Case 3.4.7) $\text{Config}(T; 3EG_j) = (u, u, t)$; (Case 3.4.8) $\text{Config}(T; 3EG_j) = (t, t, t)$.

For the statement (a), we have the partial tours on $3EG_j$ with no endpoints for Cases 3.4.1 to 3.4.4.

Case 3.4.1: $s_j x_0 \Lambda_x x_1 x_2 x_3 x_4 x_5 x_6 x_7 v y_0 \Lambda_y y_1 y_2 y_3 y_4 y_5 y_6 y_7 w z_0 \Lambda_z z_1 z_2 z_3 z_4 z_5 z_6 z_7 t_j$;

Case 3.4.2: $s_j x_0 \Lambda_x x_1 x_2 x_3 y_4 y_5 y_6 y_1 y_2 y_3 z_4 z_5 z_6 z_1 z_2 z_3 x_4 x_5 x_6 x_7 v y_0 y_7 w z_0 z_7 t_j$;

Case 3.4.3: $s_j x_0 x_7 v y_0 \Lambda_y y_1 y_2 y_3 z_4 z_5 z_6 z_1 z_2 z_3 x_4 x_5 x_6 x_1 x_2 x_3 y_4 y_5 y_6 y_7 w z_0 z_7 t_j$;

Case 3.4.4: $s_j x_0 x_7 v y_0 y_7 w z_0 \underline{A_z z_1 z_2 z_3 x_4 x_5 x_6 x_1 x_2 x_3 y_4 y_5 y_6 y_1 y_2 y_3 z_4 z_5 z_6 z_7 t_j}$.

For the statement (a), we have the partial tours on $3EG_j$ with two endpoints for Cases 3.4.5 to 3.4.8, where the endpoints are specified by underlined vertices.

Case 3.4.5: $s_j \underline{x_0} \quad \underline{x_1 x_2 x_3 x_4 x_5 x_6 x_7 v y_0 A_y y_1 y_2 y_3 y_4 y_5 y_6 y_7 w z_0 A_z z_1 z_2 z_3 z_4 z_5 z_6 z_7 t_j}$;

Case 3.4.6: $s_j x_0 \underline{A_x x_1 x_2 x_3 x_4 x_5 x_6 x_7 v y_0} \quad \underline{y_1 y_2 y_3 y_4 y_5 y_6 y_7 w z_0 A_z z_1 z_2 z_3 z_4 z_5 z_6 z_7 t_j}$;

Case 3.4.7: $s_j x_0 \underline{A_x x_1 x_2 x_3 x_4 x_5 x_6 x_7 v y_0 A_y y_1 y_2 y_3 y_4 y_5 y_6 y_7 w z_0} \quad \underline{z_1 z_2 z_3 z_4 z_5 z_6 z_7 t_j}$;

Case 3.4.8: $s_j \underline{x_0} \quad \underline{x_1 x_2 x_3 y_4 y_5 y_6 y_1 y_2 y_3 z_4 z_5 z_6 z_1 z_2 z_3 x_4 x_5 x_6 x_7 v y_0 y_7 w z_0 z_7 t_j}$.

To complete the proof of the statement (a), we need to show that it is impossible to construct the partial tour on $3EG_j$ with less than two endpoints for Cases 3.4.5 to 3.4.8. From Proposition 3.3 and the assumption that there are no semitraversed Λ -pairs in $3EG_j$, it follows that there are even number of endpoints in $3EG_j$ with respect to T . Assume that there are no endpoints in $3EG_j$.

For Case 3.4.5, any (partial) tour T on $3EG_j$ must traverse the following edges $s_j x_0, x_7 v y_0 A_y y_1 y_2 y_3, y_7 w z_0 A_z z_1 z_2 z_3, z_7 t_j, x_1 x_2 x_3, x_4 x_5 x_6, y_4 y_5 y_6, z_4 z_5 z_6$ in Figure 3-a to avoid making the vertices $s_j, x_2, x_5, v, A_y, y_2, y_5, w, A_z, z_2, z_5, t_j$ endpoints. To avoid making the vertices y_6, y_7, z_6, z_7 endpoints, the tour T must traverse the edges $y_6 y_7$ and $z_6 z_7$. So any (partial) tour T on $3EG_j$ with no endpoints must traverse the edges $s_j x_0, x_1 x_2 x_3, x_4 x_5 x_6, x_7 v y_0 A_y y_1 y_2 y_3, y_4 y_5 y_6 y_7 w z_0 A_z z_1 z_2 z_3, z_4 z_5 z_6 z_7 t_j$. If the tour T traverses the edge $x_6 x_7$, then the vertex x_0 is an endpoint. If the tour T traverses the edge $x_1 x_6$, then it makes a cycle $y_4 y_5 y_6 y_7 w z_0 A_z z_1 z_2 z_3 x_4 x_5 x_6 x_1 x_2 x_3 y_4$ in $3EG_j$. Thus it follows from Proposition 3.3 that it is impossible to construct partial tour on $3EG_j$ with less than two endpoints for Case 3.4.5. In a way similar to Case 3.4.5, we can also show that it is the case for Cases 3.4.6 to 3.4.8.

For the statement (b), we have the partial tours on $3EG_j$ with no endpoints for Cases 3.4.5 to 3.4.8.

Case 3.4.5: $s_j x_0 x_1 x_2 x_3 y_4 y_5 y_6 A_y y_1 y_2 y_3 z_4 z_5 z_6 A_z z_1 z_2 z_3 x_4 x_5 x_6 x_7 v y_0 y_7 w z_0 z_7 t_j$;

Case 3.4.6: $s_j x_0 x_7 v y_0 y_1 y_2 y_3 z_4 z_5 z_6 A_z z_1 z_2 z_3 x_4 x_5 x_6 A_x x_1 x_2 x_3 y_4 y_5 y_6 y_7 w z_0 z_7 t_j$;

Case 3.4.7: $s_j x_0 x_7 v y_0 y_7 w z_0 z_1 z_2 z_3 x_4 x_5 x_6 A_x x_1 x_2 x_3 y_4 y_5 y_6 A_y y_1 y_2 y_3 z_4 z_5 z_6 z_7 t_j$;

Case 3.4.8: $s_j x_0 x_1 x_2 x_3 x_4 x_5 x_6 x_7 v y_0 y_1 y_2 y_3 y_4 y_5 y_6 y_7 w z_0 z_1 z_2 z_3 z_4 z_5 z_6 z_7 t_j$.

For the statement (b), we have the partial tours on $3EG_j$ with two endpoints for Cases 3.4.1 to 3.4.4, where the endpoints are specified by underlined vertices.

Case 3.4.1: $s_j x_0 x_1 A_x x_6 x_5 x_4 z_3 z_2 z_1 A_z z_6 z_5 z_4 y_3 y_2 y_1 A_y y_6 y_5 y_4 x_3 x_2 \quad \underline{x_7 v y_0 y_7 w z_0 z_7 t_j}$;

Case 3.4.2: $s_j \underline{x_0} \quad \underline{x_5 x_4 x_3 x_2 x_1 A_x x_6 x_7 v y_0 y_1 y_2 y_3 y_4 y_5 y_6 y_7 w z_0 z_1 z_2 z_3 z_4 z_5 z_6 z_7 t_j}$;

Case 3.4.3: $s_j x_0 x_1 x_2 x_3 x_4 x_5 x_6 x_7 v y_0 \quad \underline{y_5 y_4 y_3 y_2 y_1 A_y y_6 y_7 w z_0 z_1 z_2 z_3 z_4 z_5 z_6 z_7 t_j}$;

Case 3.4.4: $s_j x_0 x_1 x_2 x_3 x_4 x_5 x_6 x_7 v y_0 y_1 y_2 y_3 y_4 y_5 y_6 y_7 w z_0 \quad \underline{z_5 z_4 z_3 z_2 z_1 A_z z_6 z_7 t_j}$.

The rest of the proof of the statement (b) is similar to that of the statement (a). ■

Lemma 3.5: *Let T be a tour on the equation gadget $3EG_j$ in Figure 3 (resp. Figure 4) such that there is one semitraversed Λ -pair. Then it is possible to modify the tour T such that there is one endpoint in $3EG_j$ without increasing its weight and changing the way of traversing the Λ -pairs and it is impossible to construct a tour with less than one endpoint in $3EG_j$.*

Proof: We show the lemma for the symmetric case (the asymmetric case can be shown analogously).

Since there is one semitraversed Λ -pair, we have the following cases: (Case 3.5.1) $\text{Config}(T; 3EG_j) = (s, u, u)$; (Case 3.5.2) $\text{Config}(T; 3EG_j) = (s, u, t)$; (Case 3.5.3) $\text{Config}(T; 3EG_j) = (s, t, u)$; (Case 3.5.4) $\text{Config}(T; 3EG_j) = (s, t, t)$; (Case 3.5.5) $\text{Config}(T; 3EG_j) = (u, s, u)$; (Case 3.5.6) $\text{Config}(T; 3EG_j) = (u, s, t)$; (Case 3.5.7) $\text{Config}(T; 3EG_j) = (t, s, u)$; (Case 3.5.8) $\text{Config}(T; 3EG_j) = (t, s, t)$; (Case 3.5.9) $\text{Config}(T; 3EG_j) = (u, u, s)$; (Case 3.5.10) $\text{Config}(T; 3EG_j) = (u, t, s)$; (Case 3.5.11) $\text{Config}(T; 3EG_j) = (t, u, s)$; (Case 3.5.12) $\text{Config}(T; 3EG_j) = (t, t, s)$.

It follows from Proposition 3.3 that for both equation gadgets $3EG_j$ in Figures 3-a and 3-b, it is impossible to construct a tour with less than one endpoint in $3EG_j$. It is easy to verify that for Cases 3.5.1 to 3.5.12, there exist tours (with one endpoint) on $3EG_j$ in Figures 3-a and 3-b. In fact, we have the (partial) tours on $3EG_j$ with one endpoint in Figure 3-a, e.g.,

Case 3.5.9: $s_j x_0 A_x x_1 x_2 x_3 x_4 x_5 x_6 x_7 v y_0 A_y y_1 y_2 y_3 y_4 y_5 y_6 y_7 w z_0 A_z \quad \underline{z_1 z_2 z_3 z_4 z_5 z_6 z_7 t_j}$;

Case 3.5.11: $s_j x_0 x_7 v y_0 A_y y_1 y_2 y_3 z_4 z_5 z_6 z_1 z_2 z_3 x_4 x_5 x_6 x_1 x_2 x_3 y_4 y_5 y_6 y_7 w z_0 A_z \quad \underline{z_7 t_j}$;

Case 3.5.12: $s_j x_0 x_7 v y_0 y_7 w z_0 A_z \quad \underline{z_1 z_2 z_3 x_4 x_5 x_6 x_1 x_2 x_3 y_4 y_5 y_6 y_1 y_2 y_3 z_4 z_5 z_6 z_7 t_j}$,

and we have the (partial) tours on $3EG_j$ with one endpoint in Figure 3-b, e.g.,

Case 3.5.9: $s_j x_0 x_7 v y_0 y_7 w z_0 z_1 z_2 z_3 x_4 x_5 x_6 A_x x_1 x_2 x_3 y_4 y_5 y_6 A_y y_1 y_2 y_3 z_4 z_5 z_6 A_z z_7 t_j$;

Case 3.5.11: $s_j x_0 x_1 x_2 x_3 y_4 y_5 y_6 A_y y_1 y_2 y_3 z_4 z_5 z_6 A_z z_1 z_2 z_3 x_4 x_5 x_6 x_7 v y_0 y_7 w z_0 z_7 t_j$;

Case 3.5.12: $s_j x_0 x_1 x_2 x_3 x_4 x_5 x_6 x_7 v y_0 y_1 y_2 y_3 y_4 y_5 y_6 y_7 w z_0 z_1 z_2 z_3 z_4 z_5 z_6 z_7 t_j$.

For the asymmetric case, we say that a Λ -pair is *semitraversed* if one of incoming and outgoing Λ -edges is traversed by a tour. So the asymmetric case can be shown analogously. ■

Lemma 3.6: *Let T be a tour on the equation gadget $3EG_j$ in Figure 3 (resp. Figure 4) such that there are two semitraversed Λ -pairs. Then it is possible to modify the tour T such that there are two endpoints in $3EG_j$ without increasing its weight and changing the way of traversing the Λ -pairs and it is impossible to construct a tour with less than two endpoints in $3EG_j$.*

Proof: We show the lemma for the symmetric case (the asymmetric case can be shown analogously).

We have the following cases: (Case 3.6.1) $\text{Config}(T; 3EG_j) = (s, s, u)$; (Case 3.6.2) $\text{Config}(T; 3EG_j) = (s, s, t)$; (Case 3.6.3) $\text{Config}(T; 3EG_j) = (s, u, s)$; (Case 3.6.4) $\text{Config}(T; 3EG_j) = (s, t, s)$; (Case 3.6.5) $\text{Config}(T; 3EG_j) = (u, s, s)$; (Case 3.6.6) $\text{Config}(T; 3EG_j) = (t, s, s)$.

It is easy to verify that for Cases 3.6.1 to 3.6.6, there exist tours (with two endpoints) on $3EG_j$ in Figures 3-a and 3-b. In fact, we have the tours on $3EG_j$ with two endpoints in Figure 3-a, e.g.,

Case 3.6.5: $s_j x_0 A_x x_1 x_2 x_3 x_4 x_5 x_6 x_7 v y_0 A_y y_1 y_2 y_3 y_4 y_5 y_6 y_7 w z_0 A_z z_1 z_2 z_3 z_4 z_5 z_6 z_7 t_j$;

Case 3.6.6: $s_j x_0 x_7 v y_0 A_y y_1 y_2 y_3 z_4 z_5 z_6 z_1 z_2 z_3 x_4 x_5 x_6 x_1 x_2 x_3 y_4 y_5 y_6 y_7 w z_0 A_z z_7 t_j$,

and we have the tours on $3EG_j$ with two endpoints in Figure 3-b, e.g.,

Case 3.6.5: $s_j x_0 x_7 v y_0 y_1 y_2 y_3 z_4 z_5 z_6 A_z z_1 z_2 z_3 x_4 x_5 x_6 A_x x_1 x_2 x_3 y_4 y_5 y_6 A_y y_7 w z_0 z_7 t_j$;

Case 3.6.6: $s_j x_0 x_1 x_2 x_3 x_4 x_5 x_6 x_7 v y_0 y_1 y_2 y_3 y_4 y_5 y_6 A_y y_7 w z_0 z_1 z_2 z_3 z_4 z_5 z_6 A_z z_7 t_j$.

To complete the proof, we need to show that it is impossible to construct the partial tour on $3EG_j$ with less than two endpoints for Cases 3.6.1 to 3.6.6. From Proposition 3.3 and the assumption that there are two semitraversed Λ -pairs in $3EG_j$, it follows that there are even number of endpoints in $3EG_j$ with respect to T . Assume that there are no endpoints in $3EG_j$ of Figure 3-a.

For Case 3.6.1, any partial tour T on $3EG_j$ must traverse the edges $s_j x_0$, $x_7 v y_0$, $y_7 w z_0 A_z z_1 z_2 z_3$, $z_7 t_j$, $x_1 x_2 x_3$, $x_4 x_5 x_6$, $y_1 y_2 y_3$, $y_4 y_5 y_6$, $z_4 z_5 z_6$ in Figure 3-a to avoid making the vertices s_j , v , w , A_z , x_2 , x_5 , y_2 , y_5 , z_2 , z_5 , t_j endpoints. To avoid making the vertex A_x an endpoint, the tour T must traverse the edges $A_z x_0$ or $A_x x_1$. If the edge $A_z x_0$ is traversed by T , then it makes the vertex x_1 or x_7 an endpoint. If the edge $A_x x_1$ is traversed by T , then it makes the vertex x_0 or x_7 an endpoint. So it follows from Proposition 3.3 that it is impossible to construct partial tour on $3EG_j$ with less than two endpoints for Case 3.6.1. In a way similar to Case 3.6.1, we can also show that it is the case for Cases 3.6.2 to 3.6.6. As for $3EG_j$ in Figure 3-b, we can show that it is impossible to construct partial tour on $3EG_j$ in with less than two endpoints for Cases 3.6.1 to 3.6.6 by the discussion similar to the above. ■

Lemma 3.7: *Let T be a tour on the equation gadget $3EG_j$ in Figure 3 (resp. Figure 4) such that there are three semitraversed Λ -pairs. Then it is possible to modify the tour T such that there are three endpoints in $3EG_j$ without increasing its weight and changing the way of traversing the Λ -pairs and it is impossible to construct a tour with less than three endpoints in $3EG_j$.*

Proof: We show the lemma for the symmetric case (the asymmetric case can be shown analogously).

It is easy to see that there exist tours (with three endpoints) on $3EG_j$ in Figures 3-a and 3-b, i.e.,

Figure 3-a: $s_j x_0 A_x x_1 x_2 x_3 x_4 x_5 x_6 x_7 v y_0 A_y y_1 y_2 y_3 y_4 y_5 y_6 y_7 w z_0 A_z z_1 z_2 z_3 z_4 z_5 z_6 z_7 t_j$;

Figure 3-b: $s_j x_0 x_1 x_2 x_3 x_4 x_5 x_6 A_x x_7 v y_0 y_1 y_2 y_3 y_4 y_5 y_6 A_y y_7 w z_0 z_1 z_2 z_3 z_4 z_5 z_6 A_z z_7 t_j$.

To complete the proof, we need to show that it is impossible to construct the partial tour T on $3EG_j$ in Figures 3-1 and 3-b with less than three endpoints. From Proposition 3.3 and the assumption that there are three semitraversed Λ -pairs in $3EG_j$, it follows that there are odd number of endpoints in $3EG_j$ with respect to T . Assume that there is one endpoint in $3EG_j$ of Figure 3-a.

We have the following cases: (Case 3.7.1) The endpoint is one of the vertices A_x , A_y , A_z ; (Case 3.7.2) The endpoint is one of the vertices $x_1, \dots, x_6$, $y_1, \dots, y_6$, $z_1, \dots, z_6$; (Case 3.7.3) The endpoint is one of the vertices s_j , x_0 , x_7 , v , y_0 , y_7 , w , z_0 , z_7 , t_j .

For Case 3.7.1, we have the following subcases: (Case 3.7.1-x) The endpoint is the vertex A_x ; (Case 3.7.1-y) The endpoint is the vertex A_y ; (Case 3.7.1-z) The endpoint is the vertex A_z . For Case 3.7.1-x, there exists a tour T on $3EG_j$ with one endpoint at A_x such that $\text{Config}(T; 3EG_j) = (s, s, s)$. So replacing the semitraversed Λ -pair of A_x by the traversed one, we have a tour T' on $3EG_j$ with no endpoints such that $\text{Config}(T'; 3EG_j) = (t, s, s)$, which contradicts Lemma 3.6 for Case 3.6.6. Thus Case 3.7.1-x never holds. In a way similar to this, Cases

3.7.1- y and 3.7.1- z contradict to Lemma 3.6 for Cases 3.6.4 and 3.6.2, respectively, and thus Cases 3.7.1- y and 3.7.1- z never hold.

For Case 3.7.2, we have the following subcases: (Case 3.7.2- x) The endpoint is one of the vertices $x_1, \dots, x_6$; (Case 3.7.2- y) The endpoint is one of the vertices $y_1, \dots, y_6$; (Case 3.7.2- z) The endpoint is one of the vertices $z_1, \dots, z_6$. For Case 3.7.2- x , any tour T on $3EG_j$ must traverse the edges $s_jx_0, x_7vy_0, y_7wz_0, z_7t_j, y_4y_5y_6, z_4z_5z_6$ to avoid making the vertices s_j, v, w, t_j, y_5, z_5 endpoints. Since the vertex A_y is not an endpoint, the tour T must traverse A_yy_0 or A_yy_1 . If the tour T traverses the edge A_yy_0 , then it makes y_1 or y_7 an endpoint. If the tour T traverses the edge A_yy_1 , then it makes y_0 or y_6 an endpoint. Thus Case 3.7.2- x never holds. In a way similar to this, Cases 3.7.2- y and 3.7.2- z never hold.

For Case 3.7.3, we can move the endpoint out of the set $W = \{s_j, x_0, x_7, v, y_0, y_7, w, z_0, z_7, t_j\}$ without increasing the number of endpoints. Let T be a tour on $3EG_j$ with one endpoint at one of the vertices in W . If $s_j \in W$ is the endpoint, then the tour T must traverse $A_xx_0x_7$. We modify the tour T by making A_x an endpoint and traversing the edge $s_jx_0x_7$. Then we have a tour T' on $3EG_j$ with one endpoint at A_x , which contradicts the result for Case 3.7.1- x . Obviously, the same discussion made for the vertex $s_j \in W$ can be applied to the vertices $v, w, t_j \in W$. If $x_0 \in W$ is the endpoint, then the tour T traverses x_0x_7 or traverses s_jx_0 and A_xx_1 . If the edge x_0x_7 is traversed by T , then it makes s_j an endpoint, which contradicts the assumption that there is one endpoint in $3EG_j$ of Figure 3-a. If the edges s_jx_0 and A_xx_1 are traversed by T , then we modify the tour T by making x_1 an endpoint and traversing the edge $s_jx_0A_x$. So we have a tour T' on $3EG_j$ with one endpoint at x_1 , which contradicts the result for Case 3.7.2- x . Note that the same discussion made for the vertex $x_0 \in W$ can be applied to the vertices $y_0, z_0 \in W$. If $x_7 \in W$ is the endpoint, then the tour T traverses $s_jx_0x_7$ or traverses x_7vy_0 and $s_jx_0A_x$. If the edge $s_jx_0x_7$ is traversed by T , then it makes v an endpoint, which contradicts the assumption that there is one endpoint in $3EG_j$. If the edges x_7vy_0 and $s_jx_0A_x$ are traversed by T , then we modify T by making x_1 an endpoint and traversing the edge $x_4x_5x_6x_7vy_0$. So we have a tour T' on $3EG_j$ with one endpoint at x_1 , which contradicts the result for Case 3.7.2- x . The same discussion made for the vertex $x_7 \in W$ can be applied to the vertices $y_7, z_7 \in W$. Thus Case 3.7.3 never holds.

Hence it follows from Proposition 3.3 that it is impossible to construct tours on $3EG_j$ in Figure 3-a with less than three endpoints. By the discussion similar to the above, we can show that it is impossible to construct tours on $3EG_j$ in Figure 3-b with less than three endpoints. ■

3.3 The Variable Cluster and the $\Delta(1, 2)$ -TSP Instance

Let $E_{2/3}[X] \in \text{O3-E2/3-LIN}$. It follows from Definition 2.3 that $x_i \in X$ occurs in $E_{2/3}[X]$ exactly three times for each $1 \leq i \leq n$. The following are the symmetric and asymmetric variable clusters VC_i for $x_i \in X$ [10]. In Figures 5-a and 5-b, the upper three vertices in VC_i correspond to the variable $x_i \in X$.

To construct an instance $G = (C, E_d)$ of $\Delta(1, 2)$ -TSP from an instance $E_{2/3}[X]$ of O3-E2/3-LIN, we connect the equation gadgets $2EG_j$, $3EG_j$, and the variables clusters VC_i as follows:

Constructing G of the Symmetric $\Delta(1, 2)$ -TSP: (1) For each $1 \leq j < m_2$, connect $2EG_j$ and $2EG_{j+1}$ to form a partial chain C_{2EG} by letting $\tau_j = \sigma_{j+1}$; (2) For each $1 \leq j < m_3$, connect $3EG_j$ and $3EG_{j+1}$ to form a partial chain C_{3EG} by letting $t_j = s_{j+1}$; (3) For each $1 \leq i < n$, connect VC_i and VC_{i+1} to form a partial chain C_{VC} by letting $\beta_i = \alpha_{i+1}$; (4) Connect partial chains C_{2EG} , C_{3EG} , and C_{VC} to construct an instance $G = (C, E_d)$ of the symmetric $\Delta(1, 2)$ -TSP by letting $\tau_{m_2} = s_1$, $t_{m_3} = \alpha_1$, and $\beta_n = \sigma_1$ and by joining the upper three vertices in VC_i ($1 \leq i \leq n$) to the corresponding A -vertex A_{x_i} in $2EG_j$ and $3EG_j$ (the resulting instance $G = (C, E_d)$ of the symmetric $\Delta(1, 2)$ -TSP is described in Figure 6).

Constructing G of the Asymmetric $\Delta(1, 2)$ -TSP: (1) For each $1 \leq j < m_2$, connect $2EG_j$ to $2EG_{j+1}$ by the edge $\tau_j\sigma_{j+1}$ to form a partial chain C_{2EG} ; (2) For each $1 \leq j < m_3$, connect $3EG_j$ to $3EG_{j+1}$ by the edge t_js_{j+1} to form a partial chain C_{3EG} ; (3) For each $1 \leq i < n$, connect VC_i to VC_{i+1} by the edge $\beta_i\alpha_{i+1}$ to form a partial chain C_{VC} ; (4) Connect partial chains C_{2EG} , C_{3EG} , and C_{VC} to construct an instance $G = (C, E_d)$ of the asymmetric $\Delta(1, 2)$ -TSP by the edges $\tau_{m_2}s_1$, $t_{m_3}\alpha_1$, and $\beta_n\sigma_1$ and by joining the upper three vertices in VC_i ($1 \leq i \leq n$) to the corresponding A -vertex A_{x_i} in $2EG_j$ and $3EG_j$ (the resulting instance $G = (C, E_d)$ of the asymmetric $\Delta(1, 2)$ -TSP is described in Figure 7).

Lemma 3.8 [10, Lemmas 4 & 5]: *For any tour T on $G \in \Delta(1, 2)$ -TSP with semitraversed A -pairs, it is possible to modify T such that there are no semitraversed A -pairs without increasing its weight.*

Figure 5: The Variable Cluster $VC_i = (C_i, E_i)$ for $x_i \in X$.

Figure 6: The Instance $G = (C, E_d)$ of the Symmetric $\Delta(1, 2)$ -TSP. Figure 7: The Instance $G = (C, E_d)$ of the Asymmetric $\Delta(1, 2)$ -TSP.

4 An Approximation Lower Bound for $\Delta(1, 2)$ -TSP

For any $G = (C, d) \in \Delta(1, 2)$ -TSP and any tour T on G , let $w(T)$ be the weight of the tour T on G and $2e(T)$ be the total number of endpoints in C with respect to T and G . Note that the total number of endpoints is always even. We also note that any tour T on $G = (C, d)$ can be viewed as the collection of disjoint paths on G^T and we call these paths *partial tours* on G^T .

Proposition 4.1 [10]: *For any $G = (C, d) \in \Delta(1, 2)$ -TSP and any tour T on G , the following equality holds: $w(T) = \|C\| + e(T)$, where $\|A\|$ denotes the cardinality of the set A .*

Let $G = (C, E_d) \in \Delta(1, 2)$ -TSP that is reduced from $E_{2/3}[X] \in \text{O3-E2/3-LIN}$. So in a way similar to Theorem 1 by Engebretsen [10], the following theorem follows from Lemmas 3.1 to 3.8.

Theorem 4.2: *For any tour T with $2e(T)$ endpoints in $G \in \Delta(1, 2)$ -TSP, there exists an assignment $a \in \{0, 1\}^n$ to X under which at most $e(T)$ equations in $E_{2/3}[X]$ do not hold.*

For any $E_{2/3}[X] \in \text{O3-E2/3-LIN}$, let $\text{opt}(E_{2/3}[X])$ be the maximum number of equations in $E_{2/3}[X]$ that simultaneously hold, and $E_{2/3}[X] = E_2[X] \cup E_3[X]$, where $E_2[X] \in \text{E2-LIN}$ and $E_3[X] \in \text{E3-LIN}$.

Theorem 4.3 [7]: *For any $\varepsilon > 0$, it is $\mathcal{NP}$ -hard for $E_{2/3}[X] \in \text{O3-E2/3-LIN}$ with $\|E_2[X]\| = 60m$ and $\|E_3[X]\| = 2m$ to decide whether $\text{opt}(E_{2/3}[X]) \geq (62 - \varepsilon)m$ or $\text{opt}(E_{2/3}[X]) \leq (61 + \varepsilon)m$.*

4.1 An Approximation Lower Bound for the Symmetric $\Delta(1, 2)$ -TSP

Recall the equation gadget $2EG_j$ in Figure 1, the equation gadget $3EG_j$ in Figure 3, and the variable cluster VC_i in Figure 5-a. The theorem below follows from Proposition 4.1 and Theorems 4.2 and 4.3.

Theorem 4.4: *For any $\varepsilon > 0$, it is $\mathcal{NP}$ -hard to decide whether an instance $G = (C, E_d)$ of the symmetric $\Delta(1, 2)$ -TSP with $1026m$ vertices has $\text{opt}(G) \geq (1027 - \varepsilon)m$ or $\text{opt}(G) \leq (1026 + \varepsilon)m$.*

Proof: Let $G = (C, E_d)$ be an instance of the symmetric $\Delta(1, 2)$ -TSP that is reduced from an instance $E_{2/3}[X] = E_2[X] \cup E_3[X]$ of O3-E2/3-LIN , where $\|E_2\| = 60m$ and $\|E_3\| = 2m$. Since each variable $x \in X$ occurs in $E_{2/3}[X]$ exactly three times, the total number of variables is $42m = (60m \times 2 + 2m \times 3)/3$. So from $2EG_j$ in Figure 1, $3EG_j$ in Figure 3, VC_i in Figure 5-a, and the construction for an instance G of the symmetric $\Delta(1, 2)$ -TSP, it follows that $\|C\| = 1026m = 60m \times 12 + 2m \times 27 + 42m \times 6$. Thus the theorem follows from Proposition 4.1 and Theorems 4.2 and 4.3. ■

Corollary 4.5: *For any $\varepsilon > 0$, it is $\mathcal{NP}$ -hard to approximate the symmetric $\Delta(1, 2)$ -TSP within a factor $1027/1026 - \varepsilon$ (≈ 1.00097) to the optimum.*

4.2 An Approximation Lower Bound for the Asymmetric $\Delta(1, 2)$ -TSP

Recall the equation gadget $2EG_j$ in Figure 2, the equation gadget $3EG_j$ in Figure 4, and the variable cluster VC_i in Figure 5-b. The theorem below follows from Proposition 4.1 and Theorems 4.2 and 4.3.

Theorem 4.6: *For any $\varepsilon > 0$, it is $\mathcal{NP}$ -hard to decide whether an instance $G = (C, E_d)$ of the asymmetric $\Delta(1, 2)$ -TSP with $534m$ vertices has $\text{opt}(G) \geq (535 - \varepsilon)m$ or $\text{opt}(G) \leq (534 + \varepsilon)m$.*

Proof: Let $G = (C, E_d)$ be an instance of the asymmetric $\Delta(1, 2)$ -TSP that is reduced from an instance $E_{2/3}[X]$

$= E_2[X] \cup E_3[X]$ of O3-E2/3-LIN, where $\|E_2\| = 60m$ and $\|E_3\| = 2m$. Since each variable $x \in X$ occurs in $E_2[X]$ exactly three times, the number of variables is $42m = (60m \times 2 + 2m \times 3)/3$. So from $2EG_j$ in Figure 2, $3EG_j$ in Figure 4, VC_i in Figure 5-b, and the construction for an instance G of the asymmetric $\triangle(1, 2)$ -TSP, it follows that $\|C\| = 534m = 60m \times 5 + 2m \times 12 + 42m \times 5$. Thus the theorem follows from Proposition 4.1 and Theorems 4.2 and 4.3. ■

Corollary 4.7: *For any $\varepsilon > 0$, it is $\mathcal{NP}$ -hard to approximate the asymmetric $\triangle(1, 2)$ -TSP within a factor $535/534 - \varepsilon$ (≈ 1.0019) to the optimum.*

5. An Approximation Lower Bound for $\triangle(1, 2)$ -TSP

To reduce an instance $E[X] = E_2[X] \cup E_3[X]$ of O3-E2/3-LIN to an instance $G = (C, E_d)$ of $\triangle(1, 2, 3)$ -TSP, we need to modify the equation gadget $2EG_j$ (in Figures 1 and 2), the equation gadget $3EG_j$ (in Figures 3 and 4), and the variable cluster VC_i (in Figure 5) to $2EG'_j$, $3EG'_j$, and VC'_i , respectively. The modified equation gadgets $2EG'_j$, $3EG'_j$ and the modified variable cluster VC'_i are given in Figures 8–12, respectively.

The way of connecting the equation gadgets $2EG'_j$, $3EG'_j$ and the variable cluster VC'_i for the symmetric and asymmetric $\triangle(1, 2, 3)$ -TSP is the same with that for the asymmetric $\triangle(1, 2)$ -TSP in Figure 7, i.e., we connect them by the edges $\tau_j\sigma_{j+1}$, t_js_{j+1} , $\beta_i\alpha_{i+1}$, $\tau_{m_2}s_1$, $t_{m_3}\alpha_1$, and $\beta_n\sigma_1$. For the instances $G = (C, E_d)$ of the symmetric and asymmetric $\triangle(1, 2, 3)$ -TSP, the weight of the edges (c_i, c_j) that do not explicitly appear in E_d is given by $w(c_i, c_j) = \min\{3, d_0(c_i, c_j)\}$, where $d_0(c_i, c_j)$ is the distance of c_i and c_j along the edges of weight one in G . Thus the following proposition holds.

Figure 8: The Symmetric Equation Gadget $2EG'_j = (C_j^2, E_j^2)$ for $e_j \in E_2[X]$.

Figure 9: The Asymmetric Equation Gadget $2EG'_j = (C_j^2, E_j^2)$ for $e_j \in E_2[X]$.

Figure 10: The Symmetric Equation Gadget $3EG'_j = (C_j^3, E_j^3)$ for $e_j \in E_3[X]$.

Figure 11: The Asymmetric Equation Gadget $3EG'_j = (C'_j, E'_j)$ for $e_j \in E_3[X]$. Figure 12: The Variable Cluster $VC'_i = (C_i, E_i)$ for $x_i \in X$. Figure 13: The Instance $G = (C, E_d)$ of the Symmetric $\Delta(1, 2, 3)$ -TSP. Figure 14: The Instance $G = (C, E_d)$ of the Asymmetric $\Delta(1, 2, 3)$ -TSP.

Proposition 5.1 [4]: For any $G = (C, d) \in \Delta(1, 2, 3)$ -TSP given in Figures 13 and 14 and any tour T on G , the following equality holds: $w(T) = \|C\| + 2e(T)$.

The theorems below follow from Proposition 5.1, Theorems 4.2 and 4.3, and the construction for an instance $G = (C, E_d)$ of the symmetric and asymmetric $\Delta(1, 2, 3)$ -TSP, respectively.

Theorem 5.2: For any $\varepsilon > 0$, it is $\mathcal{NP}$ -hard to approximate the symmetric $\Delta(1, 2, 3)$ -TSP within a factor $817/816 - \varepsilon$ (≈ 1.0012) to the optimum.

Theorem 5.3: For any $\varepsilon > 0$, it is $\mathcal{NP}$ -hard to approximate the asymmetric $\Delta(1, 2, 3)$ -TSP within a factor $364/363 - \varepsilon$ (≈ 1.0027) to the optimum.

6 An Approximation Lower Bound for SHORTEST-SUPERSTRING

Let Σ be a finite alphabet and we use $a \# b$ to denote the concatenation of strings $a, b \in \Sigma^*$. We say that $s \in \Sigma^*$ is a superstring of $t \in \Sigma^*$ if there exist $u, v \in \Sigma^*$ such that $s = u \# t \# v$, i.e., t is a substring of s . For $S =$

$\{s_1, s_2, \dots, s_n\}$ such that $s_i \in \Sigma^*$ for each $1 \leq i \leq n$, we say that $s \in \Sigma^*$ is a superstring for S if s is a superstring for each s_i . Then SHORTEST-SUPERSTRING is the minimization problem that given $S = \{s_1, s_2, \dots, s_n\}$, find the shortest superstring $s \in \Sigma^*$ for S . Blum, et al. [5, Theorem 11] showed that SHORTEST-SUPERSTRING is approximable within 3 (it is improved to $5/2$ by Sweedyk [19] and the related results on the approximation upper bounds for SHORTEST-SUPERSTRING can be found in the literatures [18, 9, 2]). Blum, et al. [5, Theorem 18] also showed that SHORTEST-SUPERSTRING is $\mathcal{APX}$ -complete by the reduction from the asymmetric $\Delta(1, 2)$ -TSP.

Let $G = (C, E_d)$ be an instance of the asymmetric $\Delta(1, 2)$ -TSP, and for each vertex $v \in C$, let $N(v)$ be the set of all vertices $u \in C$ adjacent to the vertex v . We say that the vertex $v \in C$ is *dominating* if $\text{outdegree}(v) = 1$ and $\text{indegree}(N(v)) = 1$ and the vertex $v \in C$ is *separately dominating* if v is dominating and $\text{outdegree}(N(v)) \geq 2$. We also say that $G = (C, E_d)$ is *separately dominating* if any vertex $v \in C$ of $\text{outdegree}(v) = 1$ is separately dominating.

Theorem 6.1: *For some integer $B \geq 2$, let $G = (C, E_d)$ be a separately dominating instance of the asymmetric $\Delta(1, 2)$ -TSP such that $1 \leq \text{outdegree}(v) \leq B$ for all vertices $v \in C$ and there exist l separately dominating vertices $v \in C$. Then for any integer $k > 0$, there exists a polynomial time reduction from an instance $G = (C, E_d)$ of $\Delta(1, 2)$ -TSP to an instance S of SHORTEST-SUPERSTRING such that $\text{opt}(G) = \|C\| + e$ iff $\text{opt}(S) = (k + 3)\|C\| + (k + 2)\|E_d\| - (2k + 5)l + (k + 1)e$.*

Proof: We can show the theorem in a way similar to Blum, et al. [5, Theorem 18]. For some integer $B \geq 2$, let $G = (C, E_d)$ be a separately dominating instance of the asymmetric $\Delta(1, 2)$ -TSP such that $1 \leq \text{outdegree}(v) \leq B$ for all vertices $v \in C$ and there are l separately dominating vertices $v \in C$. Let $\text{SD}(G) = \{v \in C : v \text{ is separately dominating}\}$ and $\text{NSD}(G) = \{u \in N(v) : v \in \text{SD}(G)\}$. It is obvious that $\text{SD}(G) \cap \text{NSD}(G) = \emptyset$. By the definition of $N(v)$, we can regard N as a function $N : C \rightarrow 2^C$. So from the definition of the separately dominating vertices, it follows that for each vertex $v \in \text{SD}(G)$, there exists the *unique* vertex $u \in \text{NSD}(G)$ such that $\{u\} = N(v)$. This implies that the function $N : \text{SD}(G) \rightarrow \text{NSD}(G)$ is a bijection and thus we have that $\|\text{SD}(G)\| = \|\text{NSD}(G)\| = l$.

For each vertex $v \in C - \{\text{SD}(G) \cup \text{NSD}(G)\}$, enumerate vertices $w \in N(v)$ in the cyclic manner, i.e., $N(v) = \{w_0, w_1, \dots, w_{d_v-1}\}$. It is obvious that $\text{outdegree}(v) = d_v \geq 2$ for any vertex $v \in C - \{\text{SD}(G) \cup \text{NSD}(G)\}$ from the definitions of $\text{SD}(G)$ and $\text{NSD}(G)$. For each vertex $v \in \text{SD}(G)$, enumerate vertices $w \in N(u) = N(N(v))$ in the cyclic manner for the unique vertex $u \in \text{NSD}(G)$ such that $\{u\} = N(v)$, i.e., $N(u) = \{w_0, w_1, \dots, w_{d_u-1}\} = N(N(v)) = \{w_0, w_1, \dots, w_{d_{N(v)}-1}\}$. Since $v \in \text{SD}(G)$, it is obvious that $d_{N(v)} = d_u = \text{outdegree}(u) \geq 2$ for any vertex $v \in \text{NSD}(G)$.

Let $B > 0$ be a fixed integer. Then for a separately dominating instance $G = (C, E_d)$ of the asymmetric $\Delta(1, 2)$ -TSP such that $1 \leq \text{outdegree}(v) \leq B$ for all vertices $v \in C$, we reduce an instance $G = (C, E_d)$ of the asymmetric $\Delta(1, 2)$ -TSP to an instance S of SHORTEST-SUPERSTRING as follows:

- (0) Let $S = \emptyset$;
- (1) For each vertex $v \in C - \{\text{SD}(G) \cup \text{NSD}(G)\}$,
 - (1-1) Enumerate vertices $w \in N(v)$ in the cyclic manner, i.e., $N(v) = \{w_0, w_1, \dots, w_{d_v-1}\}$,
 - (1-2) For the vertex v and each vertex $w_i \in N(v)$, construct $d_v + 1$ strings as follows:

$$\begin{aligned}
 v &\rightarrow s_v = \#^k v \# v'; \\
 w_0 &\rightarrow s_{w_0} = v' \#^k w_0 v' \#^k w_1; \\
 w_1 &\rightarrow s_{w_1} = v' \#^k w_1 v' \#^k w_2; \\
 &\vdots \\
 w_{d_v-2} &\rightarrow s_{w_{d_v-2}} = v' \#^k w_{d_v-2} v' \#^k w_{d_v-1}; \\
 w_{d_v-1} &\rightarrow s_{w_{d_v-1}} = v' \#^k w_{d_v-1} v' \#^k w_0.
 \end{aligned}$$

- (1-3) Let $S = S \cup \{s_v, s_{w_0}, s_{w_1}, \dots, s_{w_{d_v-1}}, s_{w_{d_v-2}}\}$.
- (2) For each vertex $v \in \text{SD}(G)$, let $u \in \text{NSD}(G)$ be the unique vertex such that $\{u\} = N(v)$.
 - (2-1) enumerate vertices $w \in N(u)$ in the cyclic manner, i.e., $N(u) = \{w_0, w_1, \dots, w_{d_u-1}\} = N(N(v)) = \{w_0, w_1, \dots, w_{d_{N(v)}-1}\}$,
 - (2-2) For the vertex v and each vertex $w_i \in N(u)$, construct $d_u + 1$ strings as follows:

$$\begin{aligned}
 v &\rightarrow s_v = \#^k v \# u'; \\
 w_0 &\rightarrow s_{w_0} = u' \#^k w_0 u' \#^k w_1; \\
 w_1 &\rightarrow s_{w_1} = u' \#^k w_1 u' \#^k w_2; \\
 &\vdots \\
 w_{d_u-2} &\rightarrow s_{w_{d_u-2}} = u' \#^k w_{d_u-2} u' \#^k w_{d_u-1}; \\
 w_{d_u-1} &\rightarrow s_{w_{d_u-1}} = u' \#^k w_{d_u-1} u' \#^k w_0.
 \end{aligned}$$

(2-3) Let $S = S \cup \{s_v, s_{w_0}, s_{w_1}, \dots, s_{w_{d_v-1}}, s_{w_{d_v-2}}\}$.

Note that for a separately dominating instance $G = (C, E_d)$ of the asymmetric $\triangle(1, 2)$ -TSP such that $1 \leq \text{out-degree}(v) \leq B$ for all vertices $v \in C$, the optimal tour on G always traverses edges $vw_i \in E_d$ for each $v \in \text{SD}(G)$ and each $w_i \in N(v)$. Recall that there are $l = \|\text{SD}(G)\| = \|\text{NSD}(G)\|$ separately dominating vertices $v \in C$ and that the function $N: \text{SD}(G) \rightarrow \text{NSD}(G)$ is a bijection. Thus in a way similar to the proof of Theorem 18 by Blum, et al. [5], it follows that $\text{opt}(G) = \|C\| + e$ iff

$$\begin{aligned}
\text{opt}(S) &= \left\{ \sum_{v \in C - \{\text{SD}(G) \cup \text{NSD}(G)\}} [(k+3) + (k+2)d_v] \right\} + \left\{ \sum_{v \in \text{SD}(G)} [(k+3) + (k+2)d_{N(v)}] \right\} + (k+1)e \\
&= \left\{ \sum_{v \in C - \{\text{SD}(G) \cup \text{NSD}(G)\}} [(k+3) + (k+2)d_v] \right\} + \left\{ \sum_{u \in \text{NSD}(G)} [(k+3) + (k+2)d_u] \right\} + (k+1)e \\
&= \left\{ \sum_{v \in C - \text{SD}(G)} [(k+3) + (k+2)d_v] \right\} + (k+1)e \\
&= \left\{ \sum_{v \in C - \text{SD}(G)} (k+3) \right\} + (k+2) \left\{ \sum_{v \in C} d_v \right\} - (k+2) \left\{ \sum_{v \in \text{SD}(G)} d_v \right\} + (k+1)e \\
&= (k+3)\{ \|C\| - \|\text{SD}(G)\| \} + (k+2)\|E_d\| - (k+2) \left\{ \sum_{v \in \text{SD}(G)} 1 \right\} + (k+1)e \\
&= (k+3)\|C\| + (k+2)\|E_d\| - (k+3)\|\text{SD}(G)\| - (k+2)\|\text{SD}(G)\| + (k+1)e \\
&= (k+3)\|C\| + (k+2)\|E_d\| - (2k+5)\|\text{SD}(G)\| + (k+1)e \\
&= (k+3)\|C\| + (k+2)\|E_d\| - (2k+5)l + (k+1)e,
\end{aligned}$$

where the third equality follows from the fact that $\text{SD}(G) \cap \text{NSD}(G) = \emptyset$ and the last equality follows from the fact that $\|\text{SD}(G)\| = \|\text{NSD}(G)\| = l$. ■

From the inapproximability result for the asymmetric $\triangle(1, 2)$ -TSP by Engebretsen [10] and Theorem 6.1, we can show that for any $\varepsilon > 0$, it is $\mathcal{NP}$ -hard to approximate SHORTEST-SUPERSTRING within a factor $6721/6720 - \varepsilon$ (≈ 1.00015) to the optimum, however, we can improve it by Theorems 4.6 and 6.1.

Theorem 6.2: For any $\varepsilon > 0$ and any integer $k > 0$, it is $\mathcal{NP}$ -hard to decide whether an instance S of SHORTEST-SUPERSTRING has $\text{opt}(S) \geq \{(1279 - \varepsilon)k + (2983 - \varepsilon)\}m = \{1279 - \varepsilon + (2983 - \varepsilon)/k\}km$ or $\text{opt}(S) \leq \{(1278 + \varepsilon)k + (2982 + \varepsilon)\}m = \{1278 + \varepsilon + (2982 + \varepsilon)/k\}km$.

Proof: We note that the instance $G = (C, E_d)$ of the asymmetric $\triangle(1, 2)$ -TSP in Figure 7 is separately dominating and that $1 \leq \text{out-degree}(v) \leq 2$ for all vertices $v \in C$. There are $108m = 1 \times 60m + 3 \times 2m + 1 \times 42m$ separately dominating vertices $v \in C$. In fact, the vertex τ_j of the equation gadget $2EG_j$ in Figure 2 is separately dominating for each $1 \leq j \leq m_2 = 60m$, the vertices v_j, w_j, t_j of the equation gadget $3EG_j$ in Figure 4 are separately dominating for each $1 \leq j \leq m_3 = 2m$, and the vertex β_i of the variable cluster VC_i in Figure 5-b is separately dominating for each $1 \leq i \leq n = 42m$. So from Figures 2, 4, 5-b, and 7, we have that $\|C\| = 5 \times 60m + 12 \times 2m + 5 \times 42m = 534m$ and $\|E_d\| = 11 \times 60m + 24 \times 2m + 6 \times 42m = 960m$. Thus the theorem follows from Theorems 4.6 and 6.1. ■

Corollary 6.3: For any $\varepsilon > 0$, it is $\mathcal{NP}$ -hard to approximate SHORTEST-SUPERSTRING within a factor $1279/1278 - \varepsilon$ (≈ 1.00078) to the optimum.

7 Concluding Remarks

In this paper, we have shown that for any $\varepsilon > 0$, it is $\mathcal{NP}$ -hard to approximate the symmetric $\triangle(1, 2)$ -TSP within a factor $1027/1026 - \varepsilon$ to the optimum (Corollary 4.5); it is $\mathcal{NP}$ -hard to approximate the asymmetric $\triangle(1, 2)$ -TSP within a factor $535/534 - \varepsilon$ to the optimum (Corollary 4.7); it is $\mathcal{NP}$ -hard to approximate SHORTEST-SUPERSTRING within a factor $1279/1278 - \varepsilon$ to the optimum (Corollary 6.3), however, we still have a gap to the known upper bounds.

Recently, Engebretsen and Karpinski [11] slightly improved our results by reducing O3-E2/3-LIN to $\triangle(1, 2)$ -TSP and showed that for any $\varepsilon > 0$, it is $\mathcal{NP}$ -hard to approximate the symmetric $\triangle(1, 2)$ -TSP within a factor $743/742 - \varepsilon$ to the optimum; it is $\mathcal{NP}$ -hard to approximate the asymmetric $\triangle(1, 2)$ -TSP within a factor $321/320 - \varepsilon$ to the optimum. In our and Engebretsen-Karpinski's reductions, it is crucial that for $E_{2/3}[X] \in \text{O3-E2/3-LIN}$, each variable $x \in X$ occurs in $E_{2/3}[X]$ exactly three times. To improve the approximation lower bound for $\triangle(1, 2)$ -TSP, we may have the following possible ways:

- (1) One of the most straightforward ways is to improve the approximation lower bound for O3-E2-LIN or O3-E2-e-LIN.
- (2) At present, we do not know any explicit approximation lower bound for O3-E3-LIN. If we could show a larger approximation lower bound for O3-E3-LIN, then it would lead to improving the approximation lower bounds for $\triangle(1, 2)$ -TSP and SHORTEST-SUPERSTRING.

- (3) As we have mentioned above, it is crucial in our and Engebretsen-Karpinski's reductions that for $E_{2/3}[X] \in \text{O3-E2/3-LIN}$, each variable $x \in X$ occurs in $E_{2/3}[X]$ exactly three times. If we could reduce E2-LIN or E3-LIN to $\Delta(1, 2)$ -TSP, then it would also lead to improving the approximation lower bounds for $\Delta(1, 2)$ -TSP and SHORTEST-SUPERSTRING.

Acknowledgments

The authors would like to thank anonymous referees for several valuable comments on the earlier version of this paper.

REFERENCES

- [1] Arora, A., Lund, C., Motwani, R., Sudan, M., and Szegedy, M., "Proof Verification and Hardness of Approximation Problems," *Proc. of the 33rd Annual IEEE Symposium on Foundations of Computer Science*, pp. 14–23 (1992).
- [2] Armen, C. and Stein, C., "A $2\frac{2}{3}$ -Superstring Approximation Algorithm," *Discrete Applied Mathematics*, Vol. 88, pp. 29–57 (1998).
- [3] Boesch, F. T. and Gimpel, J. F., "Covering the Points of a Digraph with Point-Disjoint Paths and Its Application to Code Optimization," *Jour. Assoc. Comput. Mach.*, Vol. 24, pp. 192–198 (1977).
- [4] Böchenhauer, H. J., Hromkovič, J., Klasing, R., Seibert, S., and Unger, W., "An Improved Lower Bound on the Approximability of Metric TSP and Approximation Algorithms for the TSP with Sharpened Triangle Inequality," *Proc. of the 17th Symposium on Theoretical Aspects of Computer Science*, Lecture Notes in Computer Science 1770, pp. 382–394 (2000).
- [5] Blum, A., Jiang, T., Li, M., Tromp, J., and Yannakakis, M., "Linear Approximation of Shortest Superstrings," *Proc. of the 23rd Annual ACM Symposium on Theory of Computing*, pp. 328–336 (1991).
- [6] Berman, P. and Karpinski, M., "On Some Tighter Inapproximability Results," Technical Report 29, Electronic Colloquium on Computational Complexity (June 1998).
- [7] Berman, P. and Karpinski, M., "On Some Tighter Inapproximability Results, Further Improvements," Technical Report 65, Electronic Colloquium on Computational Complexity (November 1998).
- [8] Christofides, N., "Worst-Case Analysis of a New Heuristic for the Traveling Salesman Problem," Technical Report 388, GSIA, Carnegie-Mellon University (1976).
- [9] Czumaj, A., Gąsieniec, L., Piotrów, M. and Rytter, W., "Parallel and Sequential Approximation of Shortest Superstrings," *Proc. of the 4th Scandinavian Workshop on Algorithm Theory*, Lecture Notes in Computer Science 824, pp. 95–106 (1994).
- [10] Engebretsen, L., "An Explicit Lower Bound for TSP with Distances One and Two," *Proc. of the 16th Annual Symposium on Theoretical Aspect of Computer Science*, Lecture Notes in Computer Science 1563, pp. 373–382 (1999).
- [11] Engebretsen, L. and Karpinski, M., "Approximation Hardness of TSP with Bounded Metrics," Technical Report 89, Electronic Colloquium on Computational Complexity (December 2000).
- [12] Frieze, A., Galbiati, G., and Maffioli, F., "On the Worst-Case Performance of Some Algorithms for the Asymmetric Traveling Salesman Problem," *Networks*, Vol. 12, pp. 23–39 (1982).
- [13] Håstad, J., "Some Optimal Inapproximability Results," *Proc. of the 29th Annual ACM Symposium on Theory of Computing*, pp. 1–10 (1997).
- [14] Karp, D. S., "Reducibility among Combinatorial Problems," *Complexity of Computer Computations*, pp. 85–103, Plenum (1972).
- [15] Papadimitriou, C. H. and Vempala, S., "On the Approximability of the Traveling Salesman Problem," *Proc. of the 32nd Annual ACM Symposium on Theory of Computing*, pp. 126–133 (2000).
- [16] Papadimitriou, C. H. and Yannakakis, M., "The Traveling Salesman Problem with Distances One and Two," *Math. Oper. Res.*, Vol. 18, pp. 1–11 (1993).
- [17] Pinter, S. and Wolfstahl, Y., "One Mapping Process to Processors," *Internat. Jour. Parallel Programming*, Vol. 16, pp. 1–15 (1987).
- [18] Teng, S. and Yao, F., "Approximating Shortest Superstrings," *Proc. of the 34th Annual IEEE Symposium on Foundations of Computer Science*, pp. 158–165 (1993).
- [19] Sweedyk, Z., "A $2\frac{1}{2}$ -Approximation Algorithm for Shortest Superstring," *SIAM J. Comput.*, Vol. 29, pp. 954–986 (2000).
- [20] Vishwanathan, S., "An Approximation Algorithm for the Asymmetric Traveling Salesman Problem with Distances One and Two," *Inform. Process. Lett.*, Vol. 44, pp. 297–302 (1992).