TSP instances that require exponentially many subtour elimination constraints

Eleonora Vercesi[1,2] and Austin Buchanan[3]

[1] Università della Svizzera Italiana (USI)
[2] Dalle Molle Institute for Artificial Intelligence USI-SUPSI
[3] Oklahoma State University

Abstract. We construct pathological TSP instances that require exponentially many subtour elimination constraints (SECs) in the sense that, if a smaller collection of SECs is imposed, then the Dantzig-Fulkerson-Johnson model returns a disconnected solution costing less than an optimal tour. For asymmetric instances on n vertices, we prove a growth rate of $\Omega(\varphi^n)$, where $\varphi \approx 1.618$ is the golden ratio. A node-splitting reduction gives metric instances with $\Omega(\varphi^{n/2})$ growth. We also give instances in the plane under the Euclidean, Rectilinear, and Maximum metrics that require exponentially many SECs. Concorde struggles on them, with solve times growing exponentially despite tiny integrality gaps.

1 Introduction

The classic TSP formulation of Dantzig, Fulkerson, and Johnson (DFJ) [4] can be written as follows, where $G = (V, E)$ is a complete undirected graph on n vertices and c_e is the cost to traverse edge $e \in E$. Traditionally, at least on paper, the model imposes exponentially many subtour elimination constraints (SECs) (1c), one for each set S in the set family $\mathcal{S}_{\text{all}} \coloneqq \{S \subset V \mid 2 \leq |S| \leq n-2\}$. However, for notational convenience, we parameterize the DFJ model by $\mathcal{S} \subseteq \mathcal{S}_{\text{all}}$.

$$\text{DFJ}(G, \mathcal{S}) = \min \sum_{e \in E} c_e x_e \tag{1a}$$

$$\text{s.t.} \sum_{e \in \delta(v)} x_e = 2 \qquad \forall v \in V \tag{1b}$$

$$\sum_{e \in \delta(S)} x_e \geq 2 \qquad \forall S \in \mathcal{S} \tag{1c}$$

$$x_e \in \{0, 1\} \qquad \forall e \in E. \tag{1d}$$

Here, $\delta(S)$ denotes the subset of edges with one endpoint in S. In a slight abuse of notation, for singleton sets $S = \{v\}$, we simply write $\delta(v)$ rather than $\delta(\{v\})$.

In practice, a tiny fraction of the exponentially many SECs often suffices to prove optimality [11,12]. Indeed, Vercesi and Buchanan [15] show that the classic `dantzig42` instance of Dantzig, Fulkerson, and Johnson requires only four SECs. That is, there exists a set family $\mathcal{S} = \{S_1, S_2, S_3, S_4\}$ for which the DFJ model

over just those four SECs (solved as an integer program) has objective value equal to the length of an optimal tour, denoted $\mathrm{TSP}(G)$. Moreover, there is no smaller set family with this property. These notions are formalized in Definition 1.

Definition 1. *A set family $\mathcal{S} \subseteq \mathcal{S}_{\text{all}}$ is sufficient if $\mathrm{DFJ}(G, \mathcal{S}) = \mathrm{TSP}(G)$, and the number of required SECs is the size of a smallest sufficient set family, i.e.,*

$$\begin{aligned} R(G) \coloneqq \min\ & |\mathcal{S}| \\ \text{s.t. } & \mathrm{DFJ}(G, \mathcal{S}) = \mathrm{TSP}(G) \\ & \mathcal{S} \subseteq \mathcal{S}_{\text{all}}. \end{aligned}$$

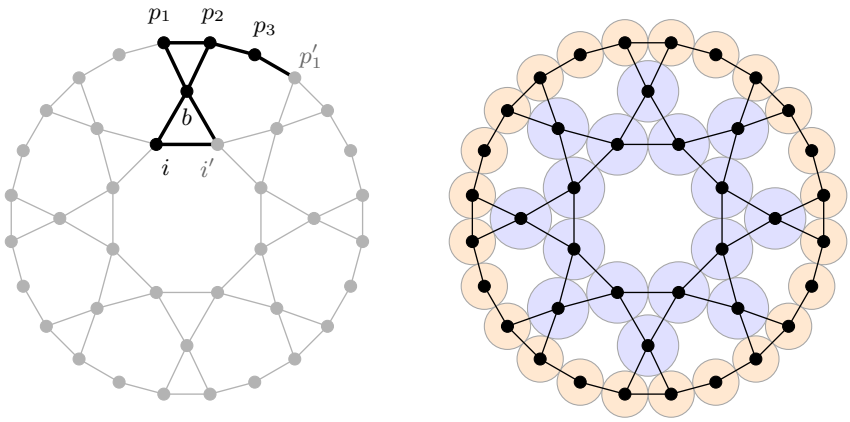
By this definition, an instance G *requires* $R(G)$ subtour constraints in the sense that all set families $\mathcal{S}$ of cardinality less than $R(G)$ yield cost strictly less than TSP optimal. Vercesi and Buchanan [15] computed $R(G)$ for TSPLIB instances, finding that it is often small. Indeed, across all TSPLIB instances with fewer than 100 vertices, 64% of them have $R(G) \leq 5$, and 91% have $R(G) \leq 12$. Meanwhile, HardTSPLIB instances [16], which are hard for the Concorde TSP solver, require significantly more, with a particular 30-vertex instance requiring 235 SECs. Similarly, some of the most challenging metric TSP instances from the literature, the Rectilinear 3-dimensional instances of Zhong [17], empirically seem to require a quadratic number of SECs, $\Theta(n^2)$. Still, it remained an open question whether any class of instances requires exponentially many SECs, $2^{\Omega(n)}$.

Our contributions. In this paper, we answer this question in the affirmative: there *do* exist pathological TSP instances that require exponentially many SECs. Section 2 gives our simplest construction, the *Ball Bearing* construction, which requires $2^{n/5}$ SECs and can be embedded in the plane, yielding TSP instances under the Euclidean, Rectilinear, and Maximum metrics that also require $2^{n/5}$ SECs. Section 3 gives a more complex *Fibonacci* construction that increases the exponential growth to $\Omega(\varphi^n)$ for asymmetric TSP and $\Omega(\varphi^{n/2})$ for metric TSP, where $\varphi \approx 1.618$ is the golden ratio. Section 4 applies the Concorde solver [1,2] to the constructions. Empirically, solve times grow exponentially despite tiny integrality gaps. This stands in contrast with HardTSPLIB and Rectilinear 3-dimensional instances, which stymie Concorde with their large integrality gaps.

2 Ball Bearing Construction

This section constructs pathological TSP instances from Ball Bearing graphs, denoted $\mathrm{BB}(k)$. As depicted in Fig. 1a, they are obtained by arranging $k \geq 3$ gadgets in a ring. Each gadget has five vertices, with an inner vertex i, ball vertex b, and three outer (or peripheral) vertices p_1, p_2, p_3, yielding $n = 5k$ total vertices. Within a gadget $g = \{i, b, p_1, p_2, p_3\}$, there are five edges $\{ib, bp_1, bp_2, p_1p_2, p_2p_3\}$, and the gadget is connected to the next gadget $g' = \{i', b', p_1', p_2', p_3'\}$ in the ring by three edges $\{ii', bi', p_3p_1'\}$. In total, $\mathrm{BB}(k)$ has $8k$ edges.

The following lemma refers to 2-factors. In a graph $G = (V, E)$, an edge subset $F \subseteq E$ is called a 2-factor if $|F \cap \delta(v)| = 2$ for all $v \in V$.



(a) The graph BB(8). (b) A coin embedding for Euclidean metric.

Fig. 1: The Ball Bearing construction.

Lemma 1. *In every 2-factor of* BB(k)*, every ball vertex b neighbors both vertices from $\{i, i'\}$ or both vertices from $\{p_1, p_2\}$; it cannot neighbor one from each.*

Proof. Since every peripheral vertex (of type) p_3 has degree two in BB(k), its two incident edges belong to every 2-factor F. Accordingly, one edge is forced incident to p_1 (by prior gadget) and one to p_2. If the edge p_1p_2 between them belongs to F, then p_1 and p_2 already have degree two, forcing b to pick bi and bi'. Otherwise, if edge p_1p_2 does *not* belong to F, then p_1 and p_2 each have one remaining edge to select (bp_1 and bp_2), so b neighbors both p_1 and p_2 in F. □

Lemma 2. *The graphs* BB(k) *are themselves non-Hamiltonian but admit $2^{n/5}$ vertex bipartitions $\{S, \bar{S}\}$ for which S and $\bar{S}$ induce Hamiltonian subgraphs.*

Proof. Let F be a 2-factor of BB(k). Consider an inner vertex i and peripheral vertex p_3. By Lemma 1, the cycle $C \subseteq F$ that contains p_3 cannot cross into the inner ring, so i belongs to a different cycle, thus F is not a Hamiltonian cycle.

Denote by B the set of ball vertices and I the inner vertices. Every vertex subset S with $I \subseteq S \subseteq I \cup B$ and its complement $\bar{S} = V(\text{BB}(k)) \setminus S$ induce a Hamiltonian subgraph (by snaking along the inner ring and outer ring, respectively, picking up ball vertices from the given side). There are $2^{|B|} = 2^{n/5}$ such sets. □

By giving BB(k)'s edges weight one, and missing edges weight two, we arrive at a 1,2-TSP instance that requires $2^k = 2^{n/5}$ SECs. To prove this, we recall a general notion (Definition 2) and result (Proposition 1) that we use throughout.

Definition 2. *A subset $S \subset V$ is cheap if* $\text{TSP}(G[S]) + \text{TSP}(G[\bar{S}]) < \text{TSP}(G)$.

If a vertex subset $S \subset V$ is cheap, then every sufficient set family $\mathcal{S}$ contains either S or its complement $\bar{S}$ (and the associated SECs are equivalent under

the degree-two constraints (1b)). This follows, for example, from Vercesi and Buchanan [15]'s characterization of sufficient set families (see their Theorem 4 and their §4.2.1). The idea here is that, if $\mathcal{S}$ contains neither S nor its complement $\bar{S}$, then the union of a TSP tour of $G[S]$ and a TSP tour over $G[\bar{S}]$ will satisfy the constraints of DFJ$(G, \mathcal{S})$, as these are the only two constraints from the full DFJ model that would "cut off" such a solution. (This statement applies to the symmetric version of the DFJ model given in the introduction, as well as to its asymmetric variant.) This implies the following proposition.

Proposition 1. *If $\mathcal{C}$ is a collection of cheap sets of G, no two of which partition V, then G requires at least $|\mathcal{C}|$ SECs.*

Theorem 1. *The 1,2-TSP instance derived from* BB(k) *requires at least $2^{n/5}$ SECs.*

Proof. By Lemma 2, BB(k) is not Hamiltonian, so the associated 1,2-TSP instance G has cost $\text{TSP}(G) \geq n + 1$. Meanwhile, also by Lemma 2, BB(k) admits $2^{n/5}$ vertex bipartitions $\{S, \bar{S}\}$ for which S and $\bar{S}$ induce Hamiltonian subgraphs, so the 1,2-weighting has $\text{TSP}(G[S]) + \text{TSP}(G[\bar{S}]) = n$. Since G admits $2^{n/5}$ cheap sets, the result follows by Proposition 1. □

Denote by SEP(G) the optimal value of the linear program (LP) (1) in which the integrality constraint (1d) is replaced by $x_e \in [0, 1]$ and all SECs are imposed (i.e., $\mathcal{S} = \mathcal{S}_{\text{all}}$). This LP is the so-called **S**ubtour **E**limination **P**roblem. The *integrality gap* of a TSP instance is defined as the ratio between TSP(G) and SEP(G); this number is always between 1 and $\frac{3}{2} - 10^{-36}$ for metric instances [7,8,9]. In the literature, computationally hard-to-solve instances are associated with large integrality gaps [5,16,17]. However, the 1,2-TSP instances derived from BB(k) have tiny integrality gaps, as noted in Proposition 2 and proved in Appendix A.

Proposition 2. *The 1,2-TSP instances G derived from* BB(k) *have*

$$\frac{\text{TSP}(G)}{\text{SEP}(G)} = \frac{n+1}{n} \to 1 \text{ as } n \to \infty.$$

2.1 Planar Embeddings

The Ball Bearing graphs are planar, and it is known by the Koebe-Andreev-Thurston theorem [10, Ch.5] that every planar graph admits a *coin embedding*, where vertices are represented by non-overlapping circular disks in the plane, with two disks being tangent precisely when their vertices are connected by an edge. Fig. 1b gives an example. This immediately gives the following corollary.

Corollary 1. *The Euclidean TSP instance derived from a coin embedding of* BB(k) *requires at least $2^{n/5}$ SECs.*

Proof. Let r_v be the radius of vertex v in the coin embedding. Then, the 2-factors from Lemma 2 all have cost $f \coloneqq \sum_{v \in V} 2r_v$, since they enter and exit each coin once, never traversing a gap. These $2^{n/5}$ sets are cheap since any TSP tour must use at least one edge from outside BB(k) (due to non-Hamiltonicity, Lemma 2), forcing an optimal tour to traverse a gap between non-touching coins, while *still* having to traverse each coin's radius twice, forcing $\mathrm{TSP}(G) > f$. □

The particular coin embedding depicted in Fig. 1b "works" for all $k \geq 3$, where the inner and ball coins can be taken as unit disks. Their positions, as well as the radii of the outer coins, follow from this choice. For the details, refer to the accompanying code, which includes generators for all proposed embeddings.

When moving to the Maximum and Rectilinear metrics, we would like a general-purpose hammer like the Koebe-Andreev-Thurston theorem saying that all planar graphs admit embeddings as non-overlapping square coins or diamond coins, respectively. Such a theorem does not exist. Nevertheless, we construct suitable embeddings for the Ball Bearing graphs.

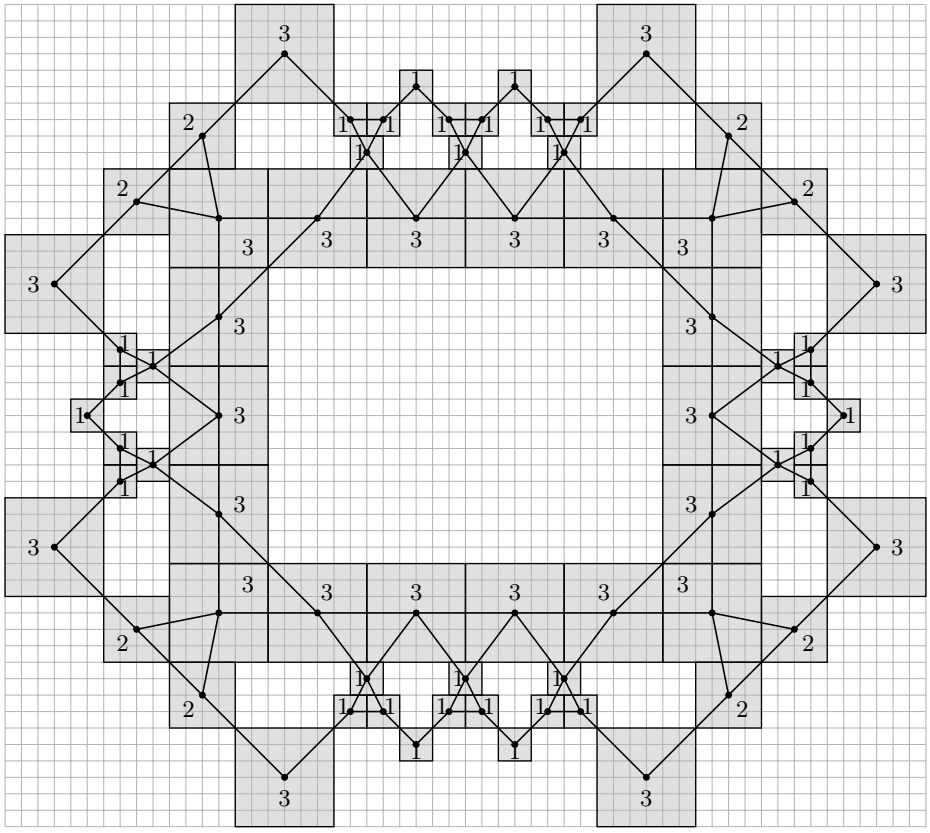


Fig. 2: A square coin embedding of BB(14) for Maximum metric.

Fig. 2 shows an embedding of BB(14) as square coins for the Maximum metric. The numbers in the figure indicate the radii, and each square is centered at integer-valued (x, y)-coordinates. So, the distances between any two vertices (x_1, y_1) and (x_2, y_2) under the Maximum metric, i.e., $\max\{|x_2 - x_1|, |y_2 - y_1|\}$, will also be integer-valued. Although the figure only shows the case of $k = 14$ gadgets, the same idea applies for any even $k \geq 8$ (by adding or subtracting pairs of gadgets on the straightaways). Fig. 8 of Appendix B depicts the $k = 8$ case.

By rotating the construction 45 degrees, we obtain a suitable embedding for the Rectilinear metric, see Fig. 9 of Appendix B for BB(10). Algebraically, this means mapping the (x, y)-coordinates of each square to $(\frac{x+y}{2}, \frac{x-y}{2})$ for the corresponding diamond. These coordinates will be half-integral. If desired, multiply by two to achieve integer-valued coordinates and distances under the Rectilinear metric, i.e., $|x_2 - x_1| + |y_2 - y_1|$. Theorem 2 summarizes these observations.

Theorem 2. *For every even $k \geq 8$, there exists a TSP instance in the plane on $n = 5k$ vertices, under the Maximum and Rectilinear metrics, that requires at least $2^{n/5}$ SECs. Moreover, the (x, y)-coordinates of each vertex are small integers (specifically, linear in n); the distances between vertices are too.*

Proof. Analogous to the proof of Corollary 1. □

3 Fibonacci Construction

The previous section gave a relatively simple construction based on the Ball Bearing graphs, which allowed us to easily obtain TSP instances in the plane under the Euclidean, Rectilinear, and Maximum metrics.

Here, we give a more complex construction. We start with a *directed* graph, the Fibonacci-Sum digraph FS(k), which, like before, is non-Hamiltonian, but admits exponentially many vertex bipartitions $\{S, \bar{S}\}$ that induce Hamiltonian subgraphs. By 1,2-weighting, this gives an ATSP instance that requires on the order of $\varphi^n \approx 1.618^n$ SECs. Then, by applying a node-splitting reduction, similar to that of Jonker and Volgenant [6], which doubles the number of nodes, we obtain a symmetric (indeed metric) instance that requires on the order of $\varphi^{n/2}$ SECs. This gives substantially faster growth in SECs, $\varphi^n \approx 2^{0.6942n}$ and $\varphi^{n/2} \approx 2^{0.3471n}$, than the Ball Bearing construction, $2^{n/5} = 2^{0.2n}$.

As depicted in Fig. 3, the Fibonacci-Sum digraph FS(k) is obtained by first arranging an even number $n = k+2$ of vertices in a directed path $(s, 1, 2, \ldots, k, t)$. For every $i, j \in [k]$ with $i < j$, we add edge (i, j) if i and j have the same parity (i.e., odd or even). So far, this gives a directed acyclic graph with $k + 1$ *path* edges and $2\binom{k/2}{2}$ *parity* edges. Last, we add k *back* edges, shown in dashes, which connect every odd $i \in [k]$ back to s, and t back to every even $j \in [k]$. The second drawing in Fig. 3 emphasizes the digraph's odd/even nature.

Before giving our main theorem, we review two key facts about Fibonacci numbers. Let F_k be the k-th Fibonacci number, with $F_1 = F_2 = 1$ and $F_j = F_{j-1} + F_{j-2}$ for $j \geq 3$. It is known that $F_k \geq \varphi^{k-2}$, see Cormen et al. [3, Lemma 19.3]. An interesting fact about the Fibonacci numbers is that F_k equals the

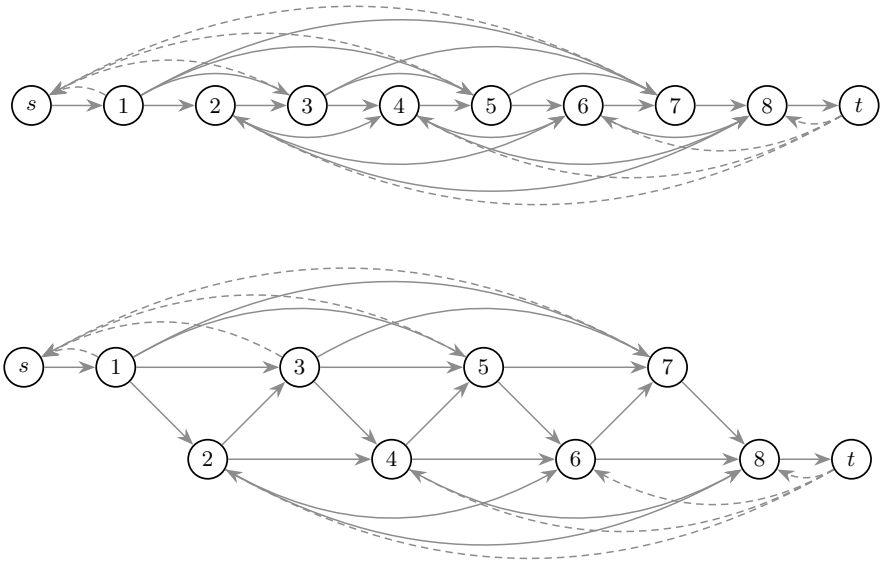


Fig. 3: Two drawings of the Fibonacci-Sum digraph FS(8).

number of *compositions* of k into odd parts, see Stanley [13, Exercise 1.35(d)]. For instance, Table 1 gives the $F_8 = 21$ compositions of 8 into odd parts.

These two facts about Fibonacci numbers are documented in Lemma 3.

Lemma 3 ([13,3]). *Let k be a positive integer. The number of compositions of k into odd parts is precisely F_k, and $F_k \geq \varphi^{k-2}$.*

We now state and prove our main theorem.

Theorem 3. *For even $k \geq 2$, FS(k) is itself non-Hamiltonian but admits F_k many vertex bipartitions $\{S, \bar{S}\}$ with $s \in S$ for which S and $\bar{S}$ induce Hamiltonian subgraphs, so the associated 1,2-ATSP instance requires at least $F_k \geq \varphi^{n-4}$ SECs.*

Proof. First, we show that each composition of k into odd parts gives a vertex bipartition $\{S, \bar{S}\}$, as shown in Table 1. Let $a_1 + a_2 + \cdots + a_p$ be a composition of k into odd parts. Since k is even, but each part a_i is odd, the number of parts p must be even. We partition $[k] = \{1, 2, \ldots, k\}$ into sets $(R_1, R_2, \ldots, R_p)$, where the *run* $R_1 = \{1, 2, \ldots, a_1\}$ contains the first a_1 vertices, $R_2 = \{a_1+1, a_1+2, \ldots, a_1+a_2\}$ contains the next a_2 vertices, etc. Generally, each run i

$$R_i = \{b_i, b_i + 1, \ldots, f_i\}$$

begins at $b_i = \sum_{j=1}^{i-1} a_j + 1$ and finishes at $f_i = \sum_{j=1}^{i} a_j$. Observe that the first run R_1 begins and finishes at odd numbers, b_1 and f_1, while R_2 begins and

Table 1: The $F_8 = 21$ compositions of $k = 8$ into odd parts.

#	composition	runs	S	$\bar{S}$
1	$1+7$	$1 \mid 2345678$	$\{s, 1\}$	$\{2, 3, 4, 5, 6, 7, 8, t\}$
2	$3+5$	$123 \mid 45678$	$\{s, 1, 2, 3\}$	$\{4, 5, 6, 7, 8, t\}$
3	$5+3$	$12345 \mid 678$	$\{s, 1, 2, 3, 4, 5\}$	$\{6, 7, 8, t\}$
4	$7+1$	$1234567 \mid 8$	$\{s, 1, 2, 3, 4, 5, 6, 7\}$	$\{8, t\}$
5	$1+1+1+5$	$1 \mid 2 \mid 3 \mid 45678$	$\{s, 1, 3\}$	$\{2, 4, 5, 6, 7, 8, t\}$
6	$1+1+3+3$	$1 \mid 2 \mid 345 \mid 678$	$\{s, 1, 3, 4, 5\}$	$\{2, 6, 7, 8, t\}$
7	$1+1+5+1$	$1 \mid 2 \mid 34567 \mid 8$	$\{s, 1, 3, 4, 5, 6, 7\}$	$\{2, 8, t\}$
8	$1+3+1+3$	$1 \mid 234 \mid 5 \mid 678$	$\{s, 1, 5\}$	$\{2, 3, 4, 6, 7, 8, t\}$
9	$1+3+3+1$	$1 \mid 234 \mid 567 \mid 8$	$\{s, 1, 5, 6, 7\}$	$\{2, 3, 4, 8, t\}$
10	$1+5+1+1$	$1 \mid 23456 \mid 7 \mid 8$	$\{s, 1, 7\}$	$\{2, 3, 4, 5, 6, 8, t\}$
11	$3+1+1+3$	$123 \mid 4 \mid 5 \mid 678$	$\{s, 1, 2, 3, 5\}$	$\{4, 6, 7, 8, t\}$
12	$3+1+3+1$	$123 \mid 4 \mid 567 \mid 8$	$\{s, 1, 2, 3, 5, 6, 7\}$	$\{4, 8, t\}$
13	$3+3+1+1$	$123 \mid 456 \mid 7 \mid 8$	$\{s, 1, 2, 3, 7\}$	$\{4, 5, 6, 8, t\}$
14	$5+1+1+1$	$12345 \mid 6 \mid 7 \mid 8$	$\{s, 1, 2, 3, 4, 5, 7\}$	$\{6, 8, t\}$
15	$1+1+1+1+1+1+3$	$1 \mid 2 \mid 3 \mid 4 \mid 5 \mid 678$	$\{s, 1, 3, 5\}$	$\{2, 4, 6, 7, 8, t\}$
16	$1+1+1+1+1+3+1$	$1 \mid 2 \mid 3 \mid 4 \mid 567 \mid 8$	$\{s, 1, 3, 5, 6, 7\}$	$\{2, 4, 8, t\}$
17	$1+1+1+1+3+1+1$	$1 \mid 2 \mid 3 \mid 456 \mid 7 \mid 8$	$\{s, 1, 3, 7\}$	$\{2, 4, 5, 6, 8, t\}$
18	$1+1+1+3+1+1+1$	$1 \mid 2 \mid 345 \mid 6 \mid 7 \mid 8$	$\{s, 1, 3, 4, 5, 7\}$	$\{2, 6, 8, t\}$
19	$1+1+3+1+1+1+1$	$1 \mid 234 \mid 5 \mid 6 \mid 7 \mid 8$	$\{s, 1, 5, 7\}$	$\{2, 3, 4, 6, 8, t\}$
20	$3+1+1+1+1+1+1$	$123 \mid 4 \mid 5 \mid 6 \mid 7 \mid 8$	$\{s, 1, 2, 3, 5, 7\}$	$\{4, 6, 8, t\}$
21	$1+1+1+1+1+1+1+1$	$1 \mid 2 \mid 3 \mid 4 \mid 5 \mid 6 \mid 7 \mid 8$	$\{s, 1, 3, 5, 7\}$	$\{2, 4, 6, 8, t\}$

finishes at even numbers, b_2 and f_2, etc. This holds because each a_i is odd. Let

$$S = \{s\} \cup R_1 \cup R_3 \cup \cdots \cup R_{p-1} \qquad \bar{S} = \{t\} \cup R_2 \cup R_4 \cup \cdots \cup R_p.$$

We argue that S and $\bar{S}$ induce Hamiltonian subgraphs, indeed following the precise sequence given above. That is, starting from s, traverse run R_1 in order, then run R_3 in order, and so on, and finally return to s. Similarly, for $\bar{S}$, start from t, traverse run R_2 in order, then run R_4 in order, and so on, and return back to t. The edges of these directed cycles exist in FS(k). Namely, within each run R_i, they follow the path edges. When jumping from one run R_i to the next R_{i+2}, the finishing vertex f_i and the beginning vertex b_{i+2} have an odd number of vertices between them (namely, a_{i+1}), and so f_i and b_{i+2} have the same parity, and so (f_i, b_{i+2}) is a parity edge. The first and last edges of S's tour are $(s, 1 = b_1)$ and (f_{p-1}, s), with the former being a path edge, and the latter being a back edge. For $\bar{S}$'s tour, (t, b_2) is a back edge, and $(f_p = k, t)$ is a path edge. We have thus shown that the F_k many compositions of k into odd parts yield (distinct) suitable partitions of FS(k)'s vertices, and $F_k \geq \varphi^{k-2} = \varphi^{n-4}$ by Lemma 3.

Finally, we show that FS(k) is non-Hamiltonian. For contradiction purposes, suppose that it *does* admit a directed Hamiltonian cycle C. Denote by ρ the number of path edges in C. We count ρ in two ways, finding it to be both odd and even, which will yield our contradiction. Write C as $(s, h_1, h_2, \ldots, h_\alpha, t, q_1, q_2, \ldots, q_\beta, s)$. We have $\alpha, \beta \geq 1$, since there are no edges between s and t. The cycle has precisely

two back edges, (t, q_1) and (q_β, s); all others are path edges or parity edges. Recall that path edges take the form $(i, i+1)$, while parity edges take the form (i, j) with $i < j$ and $j - i$ even. Observe that the path edges cross between the odd numbers (i.e., $1, 3, \ldots, k-1$) and the even numbers (i.e., $2, 4, \ldots, k$), while the parity edges and back edges do not. Accordingly, the subpath $P_1 = (s, h_1, \ldots, h_\alpha, t)$ of C, whose interior starts on the odd row and ends on the even row, must have an odd number of path edges, ρ_1. Similarly, subpath $P_2 = (t, q_1, \ldots, q_\beta, s)$ has an odd number of path edges, ρ_2. Since ρ_1 and ρ_2 are odd, their sum ρ is even.

Next, we show that ρ is odd. Classify the vertices $\{1, 2, \ldots, k\}$ as either type 1 or type 2, depending on whether they lie on subpath P_1 or P_2. Observe that edges $(s, 1)$ and (k, t) are forced into C because these are the only edges exiting s and entering t, respectively. Accordingly, these edges belong to subpath P_1 and so vertices 1 and k have type 1. Along the path $(s, 1, 2, \ldots, k, t)$ count how many pairs of consecutive vertices have different types. Denote by c this number of "cut" edges, each of which points from type 1 to type 2, or from type 2 to type 1. This number c must be even, since we start and end at type 1. Also observe that a given path edge $(i, i+1)$ belongs to C if and only if its vertices have the same type. So, of the $k+1$ path edges, c are cut, and the remaining $\rho = (k+1) - c$ belong to C. Since $k+1$ is odd and c is even, their difference ρ is odd.

In conclusion, if FS(k) admits a Hamiltonian cycle C, then the number of path edges ρ in C is odd and even, a contradiction. So, FS(k) is non-Hamiltonian. □

By applying a node-splitting reduction, we arrive at Corollary 2. The reduction is similar to that of Jonker and Volgenant [6] but exploits properties of FS(k) to keep the edge costs in $\{1, 2, 3\}$ and yields a symmetric (indeed *metric*) instance.

Corollary 2. *For even $k \geq 2$, there exists a metric TSP instance on $n = 2(k+2)$ vertices with edge costs in $\{1, 2, 3\}$ that requires at least $F_k \geq \varphi^{n/2-4}$ SECs.*

Proof. We apply a node-splitting reduction to FS(k). Namely, each vertex v of FS(k) is replaced by two vertices v and v' with an edge between them of cost one. Every directed edge (u, v) in FS(k) is replaced by undirected edge $\{u', v\}$ of cost two. All other (missing) edges are given cost three. Fig. 4 illustrates this TSP instance $G = (V, E)$, omitting the cost-3 edges for clarity. The only possible obstruction to being metric in a 1,2,3-weighted instance is a 1-1-3 triangle, but no two cost-1 edges share an endpoint, so the instance is metric.

Recall that FS(k) has F_k bipartitions $\{S, \bar{S}\}$ with both S and $\bar{S}$ inducing Hamiltonian subgraphs, with the resulting 1,2-ATSP instance $D = (N, A)$ having

$$\mathrm{TSP}(D[S]) + \mathrm{TSP}(D[\bar{S}]) = |N| < \mathrm{TSP}(D). \tag{2}$$

These same bipartitions $\{S, \bar{S}\}$, but in expanded form $\{T, \bar{T}\}$, where $T = \{v, v' \mid v \in S\}$ and $\bar{T} = \{v, v' \mid v \in \bar{S}\}$, satisfy

$$\mathrm{TSP}(G[T]) + \mathrm{TSP}(G[\bar{T}]) = |V| + |N| = 3|N| < \mathrm{TSP}(G).$$

The first equality holds because the directed cycles over S and $\bar{S}$, when expanded (using the cost-1 edges $\{v, v'\}$), yield cycles over T and $\bar{T}$ of cost $|V| + |N| = 3|N|$.

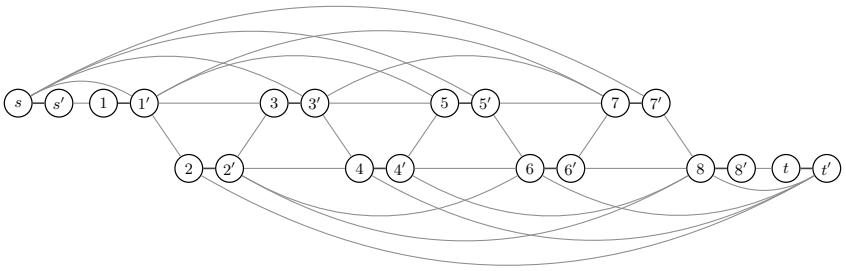


Fig. 4: Applying node-splitting to FS(8). Edges have three different costs: 1 for edges of the form $\{v, v'\}$, 2 for other drawn edges, and 3 for non-drawn edges.

The inequality holds because $\text{TSP}(G) \geq 3|N|$ by the 2-factor relaxation, and if equality held then the presumed tour must use all of the cost-1 edges, and contracting them would give a tour over D of length $|N|$, a contradiction to $\text{TSP}(D) > |N|$ from inequality (2). So, the result follows by Proposition 1. □

Table 2 compares the bounds $2^{n/5}$ and F_k on the number of required SECs for the Ball Bearing construction and the symmetric Fibonacci construction for small n. There is little difference at $n = 20$. However, at $n = 100$, the Ball Bearing construction requires (at least) 1,048,576 SECs while the Fibonacci construction requires (at least) 4,807,526,976. These lower bounds match the values $R(G)$ reported by the exact algorithm of Vercesi and Buchanan [15] across all tested values up to $n = 45$ for the Ball Bearing construction and $n = 28$ for the Fibonacci construction, suggesting that our lower bounds are in fact tight.

Table 2: Comparing $2^{n/5}$ versus F_k for $n = 2(k + 2)$.

n	k	$2^{n/5}$	F_k
20	8	16	21
40	18	256	2,584
60	28	4,096	317,811
80	38	65,536	39,088,169
100	48	1,048,576	4,807,526,976

The 1,2,3-TSP instances from Corollary 2 have tiny integrality gaps, as noted in Proposition 3 and proved in Appendix A.

Proposition 3. *For $k \geq 6$, the 1,2,3-TSP instances G derived from* FS(k) *have*

$$\frac{\text{TSP}(G)}{\text{SEP}(G)} = \frac{\frac{3}{2}n + 1}{\frac{3}{2}n} \to 1 \text{ as } n \to \infty.$$

4 Computational Experiments

In this section, we assess the empirical hardness of the proposed instances by applying the Concorde TSP solver to them. Experiments were conducted on the fat partition of a shared HPC cluster. Each node has two Intel Xeon E5-2650 v3 CPUs (2.30 GHz, 10 cores each, 20 cores total) and 128 GB of RAM. Jobs were limited to 48 hours. The code was written in Python, using PyConcorde [14] (and, as such, QSOpt as an optimizer). The code (including generators for the planar embeddings) and results are available at this link.

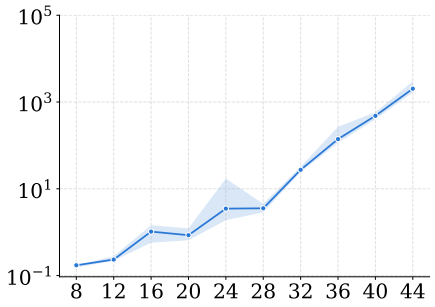
We consider five classes of TSP instances, including four Ball Bearing constructions (1,2-TSP, Euclidean, Maximum, Rectilinear) and one Fibonacci construction (1,2,3-TSP) for different values of n. We apply Concorde to each instance using 10 different seeds and plot the median solve time and interquartile range in Fig. 5. For the Fibonacci construction, the job was killed when $k > 20$ ($n > 44$) due to a memory issue; all other runs either solved to optimality or exceeded the time limit. Running times are consistent with many other hard instances of similar size in the metric context (e.g., HardTSPLIB [16]). In comparison to the tetrahedron instances of Hougardy and Zhong [5], the Euclidean embedding of the Ball Bearing construction (scaled by 1000) gives results that are two orders of magnitude slower (for 70 nodes >10,000s here versus $\leq$ 100s in Fig. 12 and 13 of [5]), leading to the "hardest" Euclidean instances currently available in the literature. Fig. 5d and 5e give runtimes for the Maximum and Rectilinear embeddings. The Rectilinear times are faster than those reported by Zhong [17] but are still much slower than those of TSPLIB instances.

We remark that the proposed instances are not hard in the usual computational complexity sense. As we saw, the 1,2-weighting of the Ball Bearing construction has $\mathrm{TSP}(G) = n + 1$, and a matching lower bound is achieved by exploiting the fact that $\mathrm{BB}(k)$ is non-Hamiltonian and adding the cut $x(E(\mathrm{BB}(k))) \leq n - 1$ to the LP relaxation over the degree-two constraints and 0,1-bounds. This specialized "hack" would speed up Concorde but is not algorithmically satisfying.

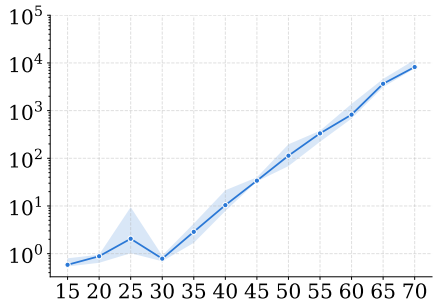
5 Conclusion and Future Work

This paper constructs pathological TSP instances that require exponentially many subtour elimination constraints, with growth rates of at least $2^{n/5}$ for instances in the plane under the Euclidean, Maximum, and Rectilinear metrics, and growth rates of φ^n for 1,2-ATSP and $\varphi^{n/2}$ for metric TSP.

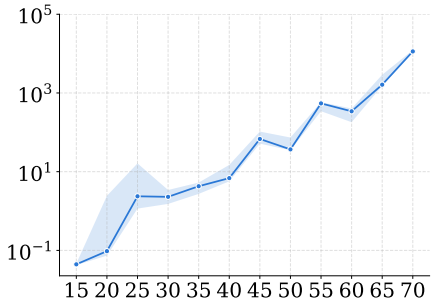
Future work may consider questions like: What is the expected number of required SECs for randomly-generated TSP instances (e.g., in a unit square) or for "typical" instances (however defined)? Is the growth polynomial or exponential? Is the φ^n growth in SECs for ATSP optimal, or do there exist instances with faster growth rate? Can any nontrivial upper bound, say, c^n for $c < 2$, be proven?



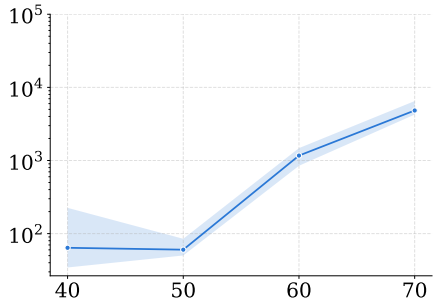
(a) Fibonacci instances, 1,2,3-weighting.



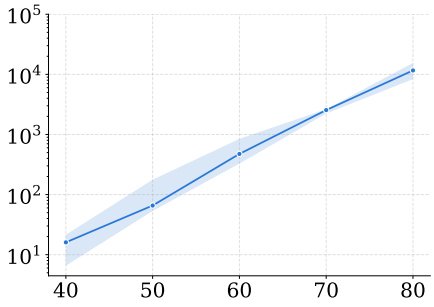
(b) BB instances, 1,2-weighting.



(c) BB instances, Euclidean metric.



(d) BB instances, Maximum metric.



(e) BB instances, Rectilinear metric.

Fig. 5: Behavior of Concorde: number of vertices versus time in sec ($\log_{10}$ scale).

Acknowledgments. The research of E. Vercesi was supported by the Swiss National Science Foundation [Grant 200021_212929/1] "Computational methods for integrality gaps analysis." The research of A. Buchanan is based on work supported the Air Force Office of Scientific Research [Grant FA9550-25-1-0277]. The authors developed and refined the constructions in concert with Claude (e.g., Opus 5.5). Claude was also used to generate TikZ figures. For the experiments, GitHub Copilot was used for code completion and for writing the documentation. However, all text in this paper, including prose, theorem statements, and proofs, was written solely by the (human) authors.

References

1. David Applegate, Robert Bixby, Vašek Chvátal, and William Cook. Implementing the Dantzig-Fulkerson-Johnson algorithm for large traveling salesman problems. *Mathematical Programming*, 97(1):91–153, 2003.
2. David L. Applegate, Robert E. Bixby, Vašek Chvátal, and William J. Cook. *The Traveling Salesman Problem: A Computational Study*. Princeton University Press, 2007.
3. Thomas H Cormen, Charles Eric Leiserson, Ronald L Rivest, and Clifford Stein. *Introduction to Algorithms*. MIT Press, Cambridge, 3 edition, 2009.
4. G Dantzig, R Fulkerson, and S Johnson. Solution of a large-scale traveling-salesman problem. *Operations Research*, 2(4):393–410, 1954.
5. Stephan Hougardy and Xianghui Zhong. Hard to solve instances of the Euclidean traveling salesman problem. *Mathematical Programming Computation*, 13(1):51–74, 2021.
6. Roy Jonker and Ton Volgenant. Transforming asymmetric into symmetric traveling salesman problems. *Operations Research Letters*, 2(4):161–163, 1983.
7. Anna Karlin, Nathan Klein, and Shayan Oveis Gharan. A (slightly) improved bound on the integrality gap of the subtour LP for TSP. In *2022 IEEE 63rd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 832–843. IEEE, 2022.
8. Anna R Karlin, Nathan Klein, and Shayan Oveis Gharan. A (slightly) improved approximation algorithm for metric TSP. In *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing*, pages 32–45, 2021.
9. Anna R Karlin, Nathan Klein, and Shayan Oveis Gharan. A deterministic better-than-3/2 approximation algorithm for metric TSP. In *International Conference on Integer Programming and Combinatorial Optimization*, pages 261–274. Springer, 2023.
10. László Lovász. *Graphs and geometry*, volume 65. American Mathematical Society, 2019.
11. P Miliotis. Integer programming approaches to the travelling salesman problem. *Mathematical Programming*, 10(1):367–378, 1976.
12. Ulrich Pferschy and Rostislav Staněk. Generating subtour elimination constraints for the TSP from pure integer solutions. *Central European Journal of Operations Research*, 25(1):231–260, 2017.
13. Richard P Stanley. *Enumerative Combinatorics*, volume 1. Cambridge University Press, second edition, 2011. Available at https://math.mit.edu/~rstan/ec/ec1.pdf.
14. Joris Vankerschaver. Pyconcorde. https://github.com/jvkersch/pyconcorde, 2026. Accessed: 2026-10-22.

15. Eleonora Vercesi and Austin Buchanan. The Dantzig-Fulkerson-Johnson TSP formulation is easy to solve for few subtour constraints. *INFORMS Journal on Optimization*, 2026. To appear.
16. Eleonora Vercesi, Stefano Gualandi, Monaldo Mastrolilli, and Luca Maria Gambardella. On the generation of metric TSP instances with a large integrality gap by branch-and-cut. *Mathematical Programming Computation*, 15(2):389–416, 2023.
17. Xianghui Zhong. Lower bounds on the integrality ratio of the subtour LP for the traveling salesman problem. *Discrete Applied Mathematics*, 365:109–129, 2025.

Appendix

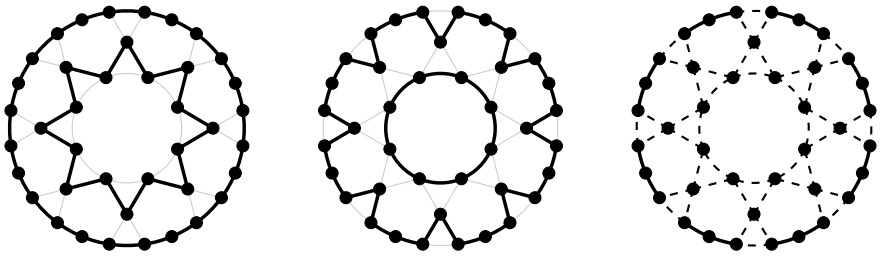
A Integrality Gap Proofs

Proof (of Proposition 2). We know $\mathrm{TSP}(G) \geq n+1$ by the proof of Theorem 1. For $\mathrm{TSP}(G) \leq n+1$, consider a gadget $g = \{i, b, p_1, p_2, p_3\}$ and a tour constructed by following the outer ring clockwise $(p_2, p_3, p_1', \ldots, p_1)$, crossing into the inner ring at ball vertex b, and then following the inner ring starting from i', picking up all other ball vertices along the way until reaching back to i, and closing the tour at p_2. All tour edges exist in $\mathrm{BB}(k)$ except final edge ip_2, giving cost $n+1$.

We prove $\mathrm{SEP}(G) = n$. The degree-two constraints (1b) imply $\sum_{e \in E} x_e = n$, and 1,2-weighting gives cost coefficients $c_e \geq 1$, so all $x \geq 0$ satisfy

$$\sum_{e \in E} c_e x_e \geq \sum_{e \in E} x_e = n.$$

In particular, this holds for any x in the subtour LP. To prove $\mathrm{SEP}(G) \leq n$, consider the 2-factors from Fig. 6. The 2-factor F puts the ball vertices on the inner cycle, while F' puts them on the outer cycle. Their midpoint $x^* = \frac{1}{2}(x^F + x^{F'})$ is half-integral and by convexity has the same cost n and satisfies the degree-two constraints and 0,1-bounds.



(a) F: balls on inner cycle. (b) F': balls on outer cycle. (c) their midpoint x^*.

Fig. 6: Subtour LP solution as midpoint of 2-factors (Ball Bearing).

It remains to show that x^* satisfies an arbitrary SEC $x(\delta(S)) \geq 2$. In the first case, suppose that S coincides with the vertex set of a cycle of F or a cycle of F'. In this case, the *other* 2-factor crosses the cut $(S, \bar{S})$ a total of $2k$ times (twice for each ball), in which case

$$x^*(\delta(S)) = \frac{1}{2}\left(x^F(\delta(S)) + x^{F'}(\delta(S))\right) = \frac{1}{2}(2k) = k \geq 3 > 2.$$

Otherwise, if S does *not* coincide with a cycle of F nor a cycle of F', then F and F' each cross the cut $(S, \bar{S})$ at least twice, so

$$x^*(\delta(S)) = \frac{1}{2}x^F(\delta(S)) + \frac{1}{2}x^{F'}(\delta(S)) \geq 1 + 1 = 2.$$

Thus, the midpoint solution x^* proves $\text{SEP}(G) \leq n$. □

Proof (of Proposition 3). As before, let $D = (N, A)$ be the 1,2-ATSP instance and $G = (V, E)$ be the 1,2,3-TSP instance with $n = |V| = 2|N|$ vertices. By the proof of Corollary 2 we have $\text{TSP}(G) > \frac{3}{2}n$. The integrality of the cost coefficients c_e and even parity of n then imply $\text{TSP}(G) \geq \frac{3}{2}n + 1$. To see $\text{TSP}(G) \leq \frac{3}{2}n + 1$, consider the straightforward tour $(s, s', 1, 1', 2, 2', \ldots, k, k', t, t', s)$, which uses $n/2$ edges of cost 1, $n/2 - 1$ edges of cost 2, and one final edge $\{t', s\}$ of cost 3, giving a total cost of $1(n/2) + 2(n/2 - 1) + 3(1) = \frac{3}{2}n + 1$, as desired.

We prove $\text{SEP}(G) = \frac{3}{2}n$. For $\text{SEP}(G) \geq \frac{3}{2}n$, see that each vertex v has only one incident edge of cost 1; all others have cost 2 or 3. So, the degree-two constraints (1b) and 0,1-bounds force a local cost of $\sum_{e \in \delta(v)} c_e x_e \geq 2 + 1 = 3$ and total cost of

$$\sum_{e \in E} c_e x_e = \frac{1}{2} \sum_{v \in V} \sum_{e \in \delta(v)} c_e x_e \geq \frac{1}{2} \sum_{v \in V} 3 = \frac{3}{2}n.$$

In particular, this holds for points in the subtour LP. To prove $\text{SEP}(G) \leq \frac{3}{2}n$, consider the 2-factors from Fig. 7. The 2-factor F is based on the all-ones composition $1 + 1 + \cdots + 1$, while F' uses $3 + 3 + 1 + \cdots + 1$. (Hence the $k \geq 6$ assumption.) Their midpoint $x^* = \frac{1}{2}(x^F + x^{F'})$ is half-integral and by convexity has the same cost $\frac{3}{2}n$ and satisfies the degree-two constraints and 0,1-bounds.

It remains to show that x^* satisfies an arbitrary SEC $x(\delta(S)) \geq 2$. Observe that F has two cycles: one over the odd row

$$V_1 \coloneqq \{s, s', 1, 1', 3, 3', \ldots, (k-1), (k-1)'\},$$

and one over the even row

$$V_2 \coloneqq \{2, 2', 4, 4', \ldots, k, k', t, t'\}.$$

Meanwhile, F' is similar, with the only changes being that 2 and $2'$ are moved to the odd row, and that 5 and $5'$ are moved to the even row. Denote by V_1' and V_2' the associated vertex sets. If $S \notin \{V_1, V_2, V_1', V_2'\}$, then each of F and F' crosses the cut $(S, \bar{S})$ at least twice, so

$$x^*(\delta(S)) = \frac{1}{2}x^F(\delta(S)) + \frac{1}{2}x^{F'}(\delta(S)) \geq 1 + 1 = 2.$$

Otherwise, we have $S \in \{V_1, V_2, V_1', V_2'\}$. Now, if $S \in \{V_1, V_2\}$, then F' crosses the cut four times, specifically at edges $\{1', 2\}, \{2', 3\}, \{4', 5\}, \{5', 6\}$, so

$$x^*(\delta(S)) = \frac{1}{2}\left(x^F(\delta(S)) + x^{F'}(\delta(S))\right) = \frac{1}{2}(4) = 2.$$

The case $S \in \{V_1', V_2'\}$ is analogous, at edges $\{t', 2\}, \{2', 4\}, \{3', 5\}$ and lastly $\{5', s\}$ if $k = 6$ (or $\{5', 7\}$ if $k \geq 8$). Thus, $\text{SEP}(G) \leq \frac{3}{2}n$ by the midpoint x^*. □

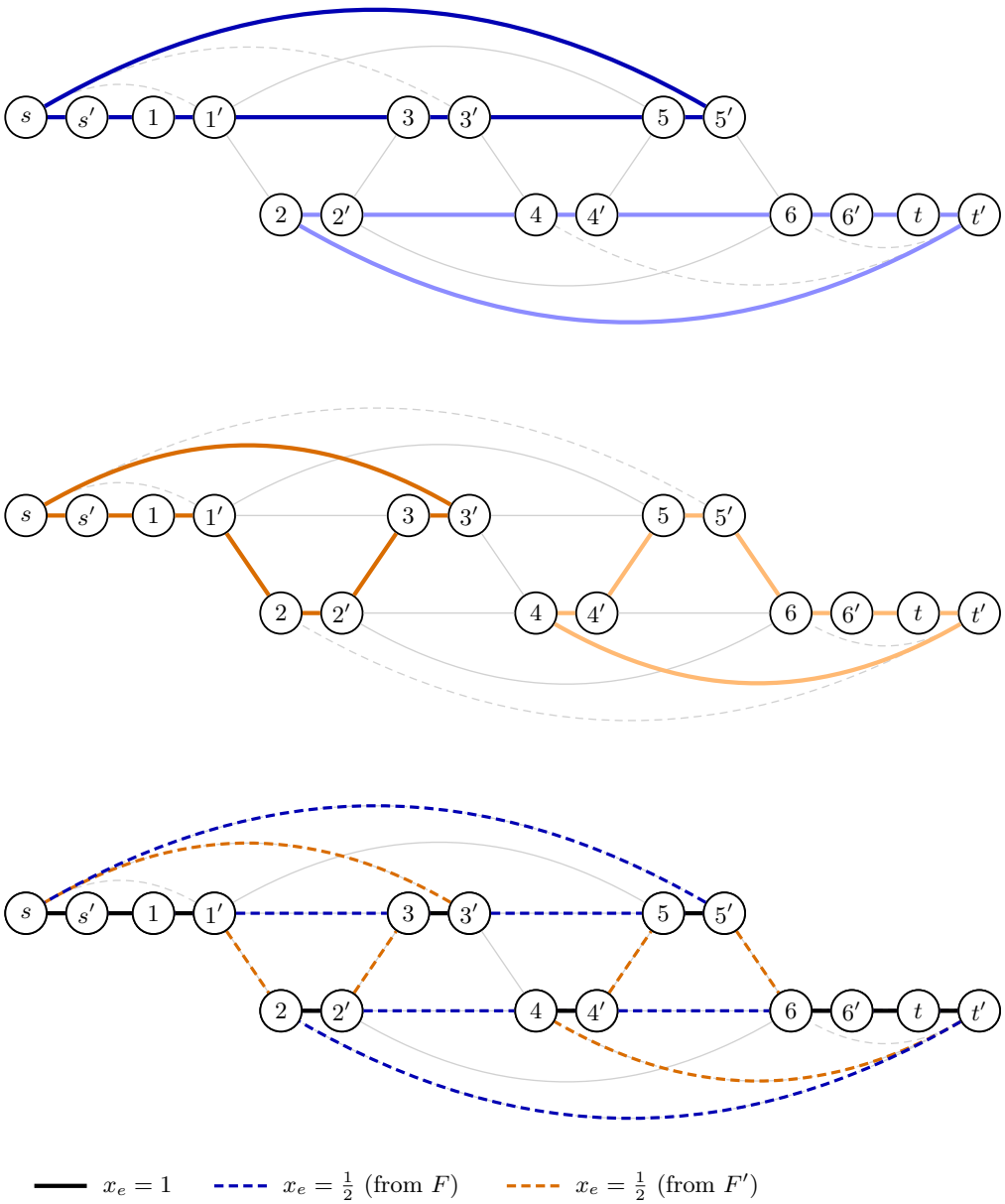


Fig. 7: Subtour LP solution as midpoint of 2-factors (Fibonacci).

B Extra Figures

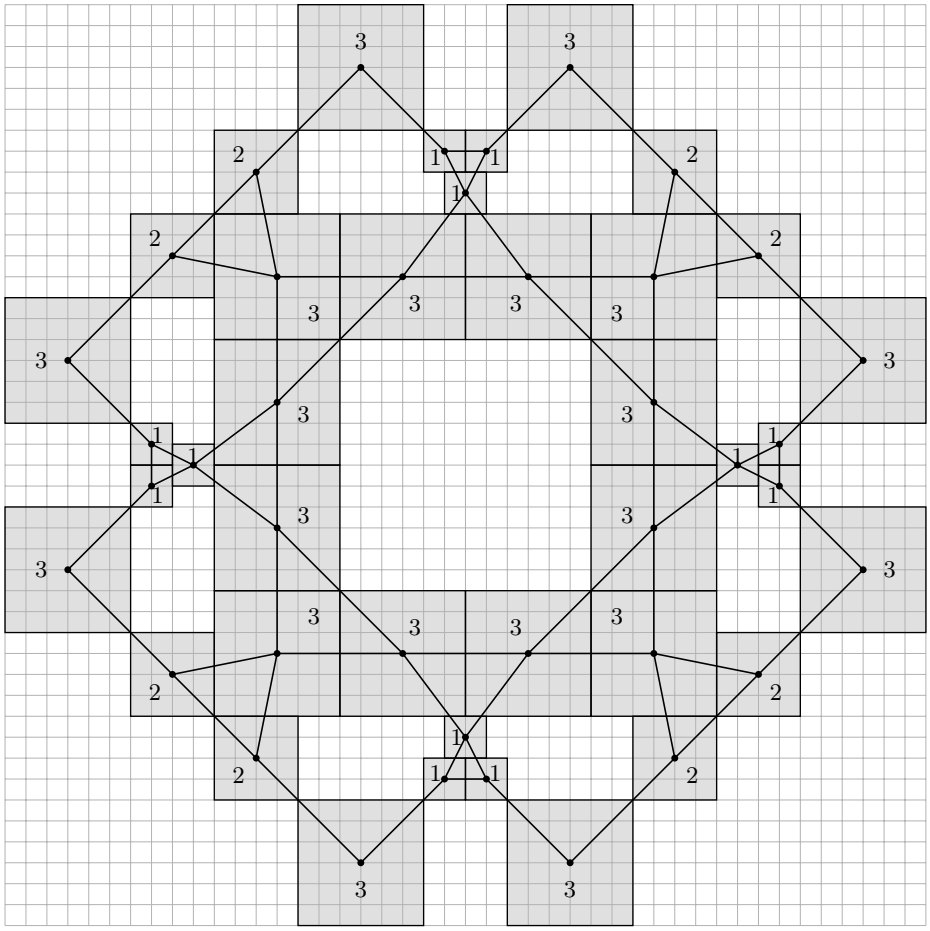


Fig. 8: A square coin embedding of BB(8) for Maximum metric.

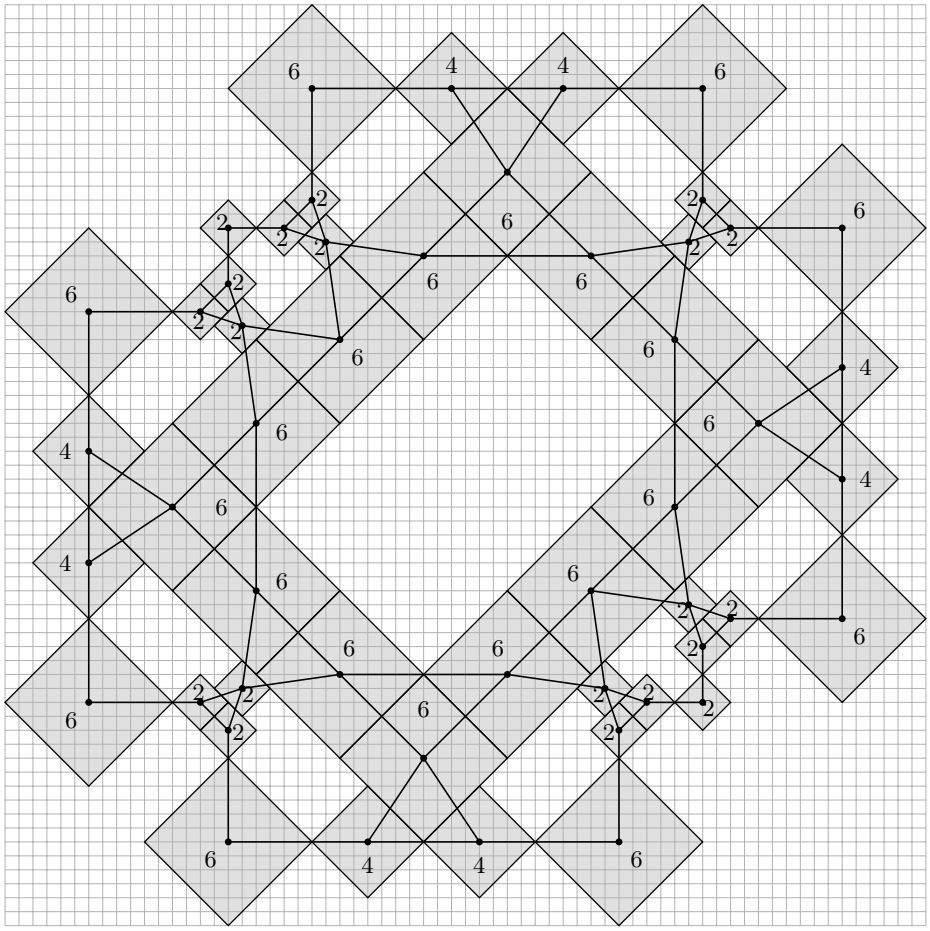


Fig. 9: A diamond coin embedding of BB(10) for Rectilinear metric. Integer coordinates are preserved via doubling.