Weighted amplifiers and inapproximability results for Travelling Salesman problem

Miroslav Chlebík[1] · **Janka Chlebíková**[2]

Abstract

The expander graph constructions and their variants are the main tool used in gap preserving reductions to prove approximation lower bounds of combinatorial optimisation problems. In this paper we introduce the weighted amplifiers and weighted low occurrence of CONSTRAINT SATISFACTION problems as intermediate steps in the NP-hard gap reductions. Allowing the weights in intermediate problems is rather natural for the edge-weighted problems as TRAVELLING SALESMAN or STEINER TREE. We demonstrate the technique for TRAVELLING SALESMAN and use the parametrised weighted amplifiers in the gap reductions to allow more flexibility in fine-tuning their expanding parameters. The purpose of this paper is to point out effectiveness of these ideas, rather than to optimise the expander's parameters. Nevertheless, we show that already slight improvement of known expander values modestly improve the current best approximation hardness value for TSP from $\frac{123}{122}$ (Karpinski et al. in J Comput Syst Sci 81(8):1665–1677, 2015) to $\frac{117}{116}$. This provides a new motivation for study of expanding properties of random graphs in order to improve approximation lower bounds of TSP and other edge-weighted optimisation problems.

Keywords Travelling Salesman problem · Approximation hardness · Probabilistic graph constructions

✉ Miroslav Chlebík
   m.chlebik@sussex.ac.uk

✉ Janka Chlebíková
   janka.chlebikova@port.ac.uk

1   Department of Mathematics, University of Sussex, Brighton, UK

2   School of Computing, University of Portsmouth, Portsmouth, UK

1 Introduction

The TRAVELLING SALESMAN problem (TSP) is undoubtedly one of the most famous combinatorial optimisation problems. In its standard version, we are given an edge-weighted (undirected) graph and the goal is to find a closed tour with a minimum cost that visits each vertex at least once. This is equivalent to the GRAPHIC TRAVELLING SALESMAN problem (where the cost between any two vertices corresponds to their shortest path in a graph) and where exactly one visit per vertex is allowed.

The GRAPHIC TSP plays an important role in understanding complexity of more general the METRIC TSP problem where cost function $c : V \times V \to \mathbb{R}^+$ is defined by a metric. The approximability of the METRIC TSP is a long-standing open problem, Christofides's approximation algorithm with ratio 1.5 (Christofides 1976) hasn't been improved for more than four decades. It is generally believed that the approximation ratio can be close to 4/3 due to known integrality gap for the Held-Karp LP relaxation (Held and Karp 1970).

In the last decade, some significant progress has been done in the GRAPHIC TSP. Gharan et al. (2011) made first breakthrough with a $(1.5 - \varepsilon)$-approximation algorithm where ε being of the order of 10^{-12}. Following that, Mömke and Svensson (2011) obtained a significantly better approximation factor of $\frac{14(\sqrt{2}-1)}{12\sqrt{2}-13} \approx 1.461$, which was improved further to $\frac{13}{9} \approx 1.444$ by Mucha (2014). To our best knowledge, currently the best known approximation ratio is 1.4 due to Sebő and Vygen (2014). The overview about this recent development can also be found in Svensson (2013).

However, there is still a significant gap between the ratio of the best approximation algorithm and the approximation ratio that provably can't be achieved unless P = NP. The first APX-hardness result showed the NP-hardness to approximate the TSP problem within $1 + \varepsilon$ without any explicit value for ε (Papadimitriou and Yannakakis 1993). The first explicit value 5381/5380 was set by Engebretsen (1998), further improved to 3813/3812 by Böckenhauer et al. (2000) and 220/219 by Papadimitriou and Vempala (2006). The further progress in the reductions and amplifiers increased the threshold to 185/184 by Lampis (2014) and to our best knowledge the currently best value is 123/122 by Karpinski et al. (2015).

Main contribution The main novelty of this paper is using weighted amplifiers and weighted low occurrence of CONSTRAINT SATISFACTION problems (CSP) as intermediate steps in the NP-hard gap reductions to the TRAVELLING SALESMAN problem. Allowing the weights in intermediate problems to TSP (or the STEINER TREE problem) is rather natural, as the problems themselves are using edge weights. We demonstrate the technique for TSP and use the parametrised weighted amplifiers in the gap reductions to allow more flexibility in fine-tuning their expanding parameters. In this paper we don't aim to optimise the parameters of amplifiers that provably exist, but show that already slight improvement of known values modestly improve the hardness of approximation for TSP from the current best value $\frac{123}{122}$ (Karpinski et al. 2015) to the new value $\frac{117}{116}$. This provides a new motivation for study of expanding properties of random graphs in order to improve approximation lower bounds of TSP and other edge-weighted optimisation problems.

Preliminaries

All graphs in this paper are undirected and connected. Let $G = (V, E)$ be an edge-weighted graph with cost edge-function $c \colon E \to \mathbb{R}^+$. For an edge $e = \{u, v\} \in E$ we also use the notation uv as an shorthand. In a narrow sense tour in the graph $G = (V, E)$ is an alternating sequence of vertices and edges, starting and ending at a vertex, where each vertex is incident with the previous and the following edge in the sequence. If a starting and ending vertex is the same, the tour is closed. Any solution of TSP is a closed tour spanning V. The multiset T of its edges is itself called a 'tour' (in a slightly broader sense adopted in this paper) even if it doesn't capture the order in which the edges are traversed by that tour. This multiset creates an Eulerian multigraph (edges are taken with their multiplicities if they are used multiple times) spanning V. It is well known (from the theory of Eulerian multigraphs) that from a multiset of edges of a tour one (or possibly many) such traversing the edges can be easily constructed.

A quasi-tour T_0 in G can be obtained from a tour T removing edges of a finite set of closed tours. It is always viewed as a multgraph with the original vertex set V, and now with edge multiset T_0. Each connected component of this multigraph is an Eulerian multigraph or an isolated vertex. A multiset T_0 of edges from E corresponds to a quasi-tour if and only if when all vertices V are balanced with respect to T_0 (each vertex from V is incident with even number of edges from T_0, possibly 0).

Our inapproximability results for the TRAVELLING SALESMAN problem use reductions from Håstad's NP-hard gap type result for MAX-E3-LIN-2, the MAXIMUM SATISFIABILITY problem for linear equations modulo 2 with *exactly* 3 variables per equation Håstad (2001). In fact, Håstad's tight inapproximability results can be stated in the form in which every variable occurs the same number of times in the system of equations, see e.g. Chlebík and Chlebíková (2003).

Theorem 1 *For every $\varepsilon \in \left(0, \frac{1}{4}\right)$ and every fixed sufficiently large integer $k \geq k(\varepsilon)$, the following partial decision subproblem $Q(\varepsilon, k)$ of MAX-E3-LIN-2 is NP-hard: given an instance of MAX-E3-LIN-2 with m equations and exactly k occurrences of each variable, to decide if at least $(1 - \varepsilon)m$ or at most $(\frac{1}{2} + \varepsilon)m$ equations are satisfied by the optimal assignment.*

The results of such form were already used to prove the inapproximability results for other optimisation problems, e.g., the STEINER TREE problem (Chlebík and Chlebíková 2008).

For some optimisation problems it is more convenient to use reductions if all equations of MAX-E3-LIN-2 have the same right hand side. The NP-hard gap results in such a case can be easily enforced if we allow flipping some occurrences of variables, so also the literal $\overline{x} := 1 - x$ can be used for a variable x. The canonical gap versions $Q_b(\varepsilon, 2k)$, for any fixed $b = 0$ or $b = 1$, of MAX-E3-LIN-2 are as follows:

THE $Q_b(\varepsilon, 2k)$ PROBLEM, $b \in \{0, 1\}$
Input An instance of MAX-E3-LIN-2 with m equations of the form $x \oplus y \oplus z = b$, each variable occurring exactly k times as unnegated and k times negated.

Task To decide if at least $(1 - \varepsilon)m$ or at most $(\frac{1}{2} + \varepsilon)m$ equations are satisfied by the optimal assignment.

The corresponding 'fixed occurrence' NP-hard gap result reads as follows [see Chlebík and Chlebíková (2003) for the details of the following theorem]:

Theorem 2 *For every* $\varepsilon \in (0, \frac{1}{4})$ *and every sufficiently large integer* k, $k \geq k(\varepsilon)$, *the partial decision subproblems* $Q_0(\varepsilon, 2k)$ *and* $Q_1(\varepsilon, 2k)$ *of MAX-E3-LIN-2 are NP-hard.*

Weighted amplifiers

Amplifier graphs are useful in proving inapproximability results for CSPs in which every variable appears a bounded (and, typically, very low) number of times. Such CSPs are often used as intermediate steps in proving approximation hardness results for many combinatorial optimisation problems. For problems like TRAVELLING SALES-MAN, or STEINER TREE which are based on edge weights, it is natural to consider the intermediate low degree CSPs with their edge weights as well.

For a graph $G = (V, E)$, a cut is a partition of V into two subsets U and $\overline{U} := V \backslash U$. The cut set $E(U, \overline{U})$ is defined as $E(U, \overline{U}) = \{uv \in E, u \in U \text{ and } v \in \overline{U}\}$ and the cut size as $|E(U, \overline{U})|$. If edges are weighted with $p \colon E \to \mathbb{R}^+$, then $p(E(U, \overline{U}))$ is weight of the cut set $E(U, \overline{U})$, hence $p(E(U, \overline{U})) = \sum_{uv \in E, u \in U, v \in \overline{U}} p(uv)$.

Definition 1 Let $G = (V, E)$ be a graph with edge weights $p \colon E \to \mathbb{R}^+$, and $D \subseteq V$, $|D| \geq 2$. We say that a weighted graph (G, p) is an amplifier for D if for every vertex set $A \subseteq V$

$$p(E(A, \overline{A})) \geq \min\{|D \cap A|, |D \cap \overline{A}|\}.$$

The vertices of the given set D are called the *contacts*, the rest of the vertices ($= V \backslash D$) is the set of *checkers*. We say that an amplifier (G, p) for the set D is a d-regular amplifier if, additionally, all contacts have degree $(d - 1)$ and all checkers have degree d (in G).

In full generality, one could also allow distinct weights for vertices of D to replace the sizes $|D \cap A|, |D \cap \overline{A}|$ with their weighted version, but for our purposes the vertices of D are uniformly weighted each with weight 1.

Note If $G = (V, E)$ is a connected graph and $D \subseteq V$ with $|D| \geq 2$ then we can define edge weights $p \colon E \to \mathbb{R}^+$ in such a way that the graph (G, p) is an amplifier. Indeed, if we set

$$\beta = \min \left\{ \frac{|E(A, \overline{A})|}{|D \cap A|} : A \subseteq V, \ 1 \leq |D \cap A| \leq \frac{1}{2}|D| \right\}$$

then clearly, $\beta > 0$, and any weights $p \colon E \to \mathbb{R}^+$ such that $p(e) \geq \frac{1}{\beta}, \forall e \in E$, produce an amplifier (G, p) for D.

2 Intermediate weighted CSPs

In this section we extend the NP-hard gap results from a system of linear equations with exactly 3 variables to a low occurrence version of w-MAX-3-LIN-2, a weighted hybrid system of linear equations over $\mathbb{Z}_2$ with either 2 or 3 variables. Similarly to MAX-E3-LIN-2, the task of the w-MAX-3-LIN-2 problem is to find an assignment that maximizes weight of the satisfied equations in the hybrid system.

To prove the NP-hard gap results for the w-MAX-3-LIN-2 problem, we extend Håstad's results for MAX-E3-LIN-2 using the amplifiers defined in Sect. 1.

Reduction from $Q(\varepsilon, k)$ to w-MAX-3-LIN-2

Let $\varepsilon \in (0, \frac{1}{4})$, and $k > 0$ be an integer such that the problem $Q(\varepsilon, k)$ is NP-hard. Let an instance I of $Q(\varepsilon, k)$ be given, denote by $v(I)$ the set of variables of I, $v := |v(I)|$. Let's assume that $G = (V, E)$ with the edge weights $p \colon E \to \mathbb{R}^+$ be an amplifier for a set $D \subseteq V$ with $|D| = k$.

Now we describe a gap preserving reduction from $Q(\varepsilon, k)$ to the w-MAX-3-LIN-2 problem with an amplifier (G, p) as a parameter. The instance I of $Q(\varepsilon, k)$ is transformed to a weighted hybrid instance J of w-MAX-3-LIN-2.

- For each variable $x \in v(I)$ take a copy of the amplifier (G, p), let (G_x, p) denote that copy:

  – Inside (G_x, p) the vertices correspond to the variables in J and each edge vv' represents the equation $v \oplus v' = 0$ with weight $p(vv')$ in J.
    The contact vertices of (G_x, p) represent k occurrences of the variable x in the equations of I. Distinct occurrences of a variable x in I are represented by the distinct contact vertices in G_x.

- Every equation $x \oplus y \oplus z = b$ from I, $b \in \{0, 1\}$, also belongs to J with weight 1.

Remark 1 Observe that the above reduction from an instance I of $Q(\varepsilon, k)$ to an instance J of w-MAX-3-LIN-2 preserves the NP-hard gap of $Q(\varepsilon, k)$. Indeed, there is a simple dependence of an optimal value for J on that of I.

In the following we show that if we look at these problems as MINIMUM UNSATIS-FIABILITY problems, where OPT′ is the corresponding minimum weight of unsatisfied equations over all assignments, then OPT′(I) = OPT′(J). Clearly, any assignment to variables from $v(I)$ generate an assignment to variable of J in a natural way; the value of a variable $x \in v(I)$ is assigned to all variables of G_x. Such assignments to variables of J are called **standard**. Hence, obviously OPT′(J) ≤ OPT′(I).

The observation that the optimum OPT′(J) is achieved on standard assignments is based on the amplifier's properties. Any assignment φ to the variables of J can be converted to a standard one in such a way that the weight of unsatisfied equations doesn't increase as follows: consider a variable x from $v(I)$. Assign to all variables in G_x the same value as it is assigned to the majority of contact vertices in G_x by the assignment φ. The fact that (G_x, p) is the amplifier ensures that the weight of unsatisfied equations in J doesn't increase. Now if we repeat the same operation for each variable from $v(I)$, one after another, the result will be a standard assignment without increase of the weight of unsatisfied equations in J. Consequently, OPT′(J)

is achieved on the standard assignments. But for every standard assignment the weight of unsatisfied equations of J is the same as the number of unsatisfied equations of I by that assignment, hence $\text{OPT}'(I) = \text{OPT}'(J)$.

Reduction from $Q_b(\varepsilon, 2k)$ to w-MAX-3-LIN-2

Now we slightly modify the previous reduction from $Q(\varepsilon, k)$ to deal with the instances of $Q_b(\varepsilon, 2k)$ for any fixed $b = 0$ or $b = 1$.

Let $\varepsilon \in (0, \frac{1}{4})$ and $k > 0$ be an integer such that $Q_b(\varepsilon, 2k)$ is NP-hard. Assume that $G = (V, E)$ with edge weights $p \colon E \to \mathbb{R}^+$ is an amplifier for a set $D \subseteq V$ with $|D| = 2k$. Let $\{V^u, V^n\}$ be a partition of V balanced in D, namely $|D \cap V^u| = |D \cap V^n| = k$. Denote further G^u and G^n the induced subgraph of G with the vertex sets V^u and V^n, respectively. In what follows we describe the reduction from $Q_b(\varepsilon, 2k)$ to w-MAX-3-LIN-2 parametrised by an amplifier (G, p) for $D \subseteq V$ with $|D| = 2k$ and with chosen balanced partition $\{V^u, V^n\}$ of V.

Let an instance I of $Q_b(\varepsilon, 2k)$ be given, $\nu(I)$ be the set of variables of I, $\nu = |\nu(I)|$.

- For each variable x from $\nu(I)$ take a copy of an amplifier (G, p), let G_x denote such a copy.

    - Any edge vv' inside either G_x^u or G_x^n represents the cycle equation $v \oplus v' = 0$ taken with weight $p(vv')$.
    - Any edge between $v \in V_x^u$ and $v' \in V_x^n$ in G_x represents the matching equation $v \oplus v' = 1$ taken with weight $p(vv')$.

- The contact vertices of G_x^u (resp. G_x^n) represent k occurrences of unnegated (resp. negated) variable x in the equations of I. Every equation $x \oplus y \oplus z = b$ from I, $b \in \{0, 1\}$, also belongs to J with weight 1.

This way we produce an instance J of the w-MAX-3-LIN-2 problem. Any assignment to variables from $\nu(I)$ generates an assignment to variables of J in a natural way: the value of a variable x is assigned to all variables of G_x^u, and the value opposite to x, $\bar{x} = 1 - x$, is assigned to all vertices of G_x^n. Such assignment to the variables of J is called **standard**. Any assignment to variables of J can be converted to a standard one without increasing the weight of unsatisfied equations as it follows from properties of an amplifier. The arguments from Remark 1 applied to the problem $Q(\varepsilon, 2k)$ can be transformed to the $Q_b(\varepsilon, 2k)$ by reversing all variables from V_x^n to their opposite.

3 The weighted bi-wheel amplifiers

The previous reductions were based on a general theoretical model of weighted amplifiers. In this section we introduce a class of weighted graphs with such expanding properties that generalise the bi-wheel amplifiers from Karpinski et al. (2015). Further we describe in the details the properties of the instances of the subproblem of w-MAX-3-LIN-2, called the Hybrid bi-wheel instances.

Definition 2 Let an integer $k > 0$ and a rational number $\tau > 1$ be such that τk is an integer. The **weighted $(2k, \tau)$-bi-wheel amplifier** $W_{k,\tau} = (V, E)$, $p \colon E \to \mathbb{R}^+$,

is a (weighted) 3-regular amplifier with a specific balanced partition constructed as follows: Take two disjoint cycles, each on τk vertices (connected in consecutive order), $V^u = \{1^u, 2^u, \ldots, (\tau k)^u\}$ and $V^n = \{1^n, 2^n, \ldots, (\tau k)^n\}$, respectively. Select the sets of k contacts $D^u \subseteq V^u$ and $D^n \subseteq V^n$ as $D^u = \{c_1^u, c_2^u, \ldots, c_k^u\}$, $D^n = \{c_1^n, c_2^n, \ldots, c_k^n\}$. The remaining vertices of both cycles, $V^u \backslash D^u$ and $V^n \backslash D^n$, are checkers.

To complete the construction, consider a perfect matching between the checkers of these two cycles where each matching edge has one vertex in the first cycle $V^u \backslash D^u$ and another one in the second cycle $V^n \backslash D^n$.

We assume that in each cycle of the bi-wheel consecutive contacts are separated by a chain of several (at least 1) checkers. Hence, in particular, $\tau \geq 2$.

Remark 2 Let us denote by $\mathscr{C}^u$ ($\mathscr{C}^n$, resp.) the set of edges contained in the first (the second, resp.) cycle in $W_{k,\tau}$, so $\mathscr{C}^u = \{\{i^u, (i+1)^u\} : i = 1, 2, \ldots, \tau k\}$ and $\mathscr{C}^n = \{\{i^n, (i+1)^n\}\} : i = 1, 2, \ldots, \tau k\}$ (the vertex $\tau k + 1$ is the vertex 1), and by $\mathscr{M} \subseteq E$ the associated perfect matching on the set of checkers. Clearly, $|\mathscr{C}^u| = |\mathscr{C}^n| = \tau k$, $|\mathscr{M}| = |V^u \backslash X^u| = |V^n \backslash X^n| = (\tau - 1)k$.

In this paper we consider only bi-wheel amplifiers $(W_{k,\tau}, p)$ whose weights have uniform cycle weight p_c for all cycle edges of both $\mathscr{C}^u$ and $\mathscr{C}^n$, and another uniform matching weight p_m for all matching edges from $\mathscr{M}$.

Now we are ready to describe the specific properties of the Hybrid bi-wheel instances of w-MAX-3-LIN-2 based on a fixed $(2k, \tau)$-bi-wheel amplifier $W_{k,\tau}$ with weights p_c and p_m.

Theorem 3 *For every $\varepsilon \in (0, \frac{1}{4})$ and $b \in \{0, 1\}$ there exist instances of w-MAX-3-LIN-2, called Hybrid$(W_{k,\tau}, p)$, with the following properties:*

(i) *each variable of the system equations Hybrid$(W_{k,\tau}, p)$ occurs exactly 3 times;*
(ii) *m equations are of the form $x \oplus y \oplus z = b$, each of weight 1;*
(iii) *$3\tau m$ equations are of the form $x \oplus y = 0$ each of weight p_c;*
(iv) *$\frac{3}{2}m(\tau - 1)$ equations are of the form $x \oplus y = 1$ each of weight p_m,*

for which it is NP-hard to decide whether there is an assignment to the variables that leaves unsatisfied equations of weight at most εm, or every assignment to the variables leaves unsatisfied equations of weight at least $(0.5 - \varepsilon)m$.

Proof Let $\varepsilon \in (0, \frac{1}{4})$, $b \in \{0, 1\}$, and $k > 0$ be an integer such that $Q_b(\varepsilon, 2k)$ is NP-hard. Let an instance I of $Q_b(\varepsilon, 2k)$ be given, $v(I)$ be its set of $v = |v(I)|$ variables, so $m := |I| = \frac{2kv}{3}$.

For any fixed variable $x \in v(I)$ we construct a $(2k, \tau)$-bi-wheel W_x (a copy of $W_{k,\tau}$) with $2\tau k$ variables $Var(x) = \{x_j^u, x_j^n\}_{j=1}^{\tau k}$. Following the notation from Definition 2, each j^u is mapped to the variable x_j^u, j^n to x_j^n. Let $\mathscr{C}^u(W_x), \mathscr{C}^n(W_x)$ denote the cycle edges, $\mathscr{M}(W_x)$ the corresponding matching edges.

– For each matching edge $j^u l^n \in \mathscr{M}(W_x)$ we create the equation $x_j^u \oplus x_l^n = 1$ weighted by p_m. Therefore, we have $v(\tau - 1)k = \frac{3}{2}m(\tau - 1)$ such equations.

– For each cycle edge $\{j^q, (j+1)^q\} \in \mathscr{C}^q(W_x)$, $j = 1, 2, \ldots, \tau k$ (the vertex $\tau k + 1$ is the vertex 1) with $q \in \{u, n\}$ we introduce the equation $x_j^q \oplus x_{j+1}^q = 0$ weighted by p_c. Hence we have $2k\tau\nu = 3\tau m$ such equations.

Finally, we replace the l-th unnegated appearance of x in I by the contact variable $x_{c^u(l)}^u := c_l^u$, whereas the l-th negated appearance is replaced by $x_{c^n(l)}^n := c_l^n$, for $l = 1, \ldots, k$. This yields m equations of the form $x \oplus y \oplus z = b$, each weighted by 1.

Obviously, each of the variables $Var(x) = \{x_j^u, x_j^n\}_{j=1}^{\tau k}$ where $x \in \nu(I)$ occurs exactly 3 times and the hardness results follows directly from a properties of the reduction discussed in Sect. 2. □

The reduction from Hybrid($W_{k,\tau}$, p), presented later in Sect. 4, is a gap preserving reduction to TSP parametrised by a $(2k, \tau)$-bi-heel amplifier with cycle weights p_c and matching weights p_m. The trade-off between parameters p_c, p_m and τ is crucial for quality of approximation lower bounds.

Definition 3 We call the triple (p_c, p_m, τ) *admissible* if for every k_0 there exists $k \geq k_0$ and a $(2k, \tau)$-bi-wheel that is an amplifier with cycle weights p_c and matching weights p_m.

The bi-wheel amplifiers introduced by Karpinski et al. (2015) are based on the fact that the triple ($p_c = 1$, $p_m = 1$, $\tau = 7$) is admissible. This leads to NP-hardness to approximate TSP to within any constant approximation ratio less than $\frac{123}{122}$. They also observed (Karpinski et al. 2015) that their proof (of amplification properties) doesn't seem to work with $\tau = 6$ instead $\tau = 7$. However, there is an opportunity for fine-tuning here if we allow non-integral τ. If, e.g., 90% of pairs of consecutive contacts in bi-wheel cycles are separated by 6 checkers, and 10% of such pairs are separated by a chain of 5 checkers only, then the proof of required amplification properties still works. The detailed explanation together with all computations for wheel amplifiers can be found in the paper Chlebík and Chlebíková (2003). The proof for bi-wheels is very similar, so along these lines one can argue that the triple ($p_c = 1$, $p_m = 1$, $\tau = 6.9$) is admissible. This itself would (very modestly) improve on the lower approximation bound for TSP given in Karpinski et al. (2015).

Introducing weighted amplifier graph constructions seems to have paid off even more compared to improvement of parameters for unweighted amplifiers. In this case we have more freedom in fine-tuning the approximation hardness lower bounds obtained in parametric way, if we can prove that bi-wheel amplifiers with certain parameters (p_c, p_m, τ) exist.

Let us explain trade-off between parameters (p_c, p_m, τ) of bi-wheels in a simple scenario with $p_m = 1$ fixed. Our contribution allows to use weighted amplifiers with $p_c < 1$ (strengthening of amplifiers) or with $p_c > 1$ (relaxing of amplifiers). One can achieve amplifiers with $p_c < 1$ by increasing τ from $\tau = 7$. On the other hand, to relax to $p_c > 1$ can be achieved with $\tau < 7$. These ideas indicate importance to better understand the exact trade-off between (p_c, p_m, τ) triples for bi-wheel amplifiers that provably exist.

Based on the definition of the admissible triples, the main result of this paper presented as Theorem 4 in Sect. 4 can be formulated as:

Theorem *If (p_c, p_m, τ) is an admissible triple then it is NP-hard to approximate the* Travelling Salesman *problem to within any constant approximation ratio less than*

$$1 + \frac{1}{3(\tau - 1)(4p_c + \max\{1, p_m\}) + 12p_c + 20\max\{1, p_m\}}.$$

In this paper we only want to demonstrate effectiveness of weighted parametrised amplifiers and address the question of fine-tuning in (p_c, p_m, τ) triples for bi-wheel amplifier. This can be summarises as [see Theorem 5 in Sect. 5 for the proof]:

Theorem *The triple $(p_c = \frac{1}{2}, p_m = 1, \tau = 11)$ is admissible, hence for every large enough $k \geq k_0$ there is a $(2k, 11)$-bi-wheel that is an amplifier with cycle weights $p_c = \frac{1}{2}$ and matching weights $p_m = 1$.*

Therefore we can conclude

Corollary *It is NP-hard to approximate the* Travelling Salesman *problem within any constant approximation ratio less than $\frac{117}{116}$.*

4 Gap preserving reduction from Hybrid($W_{k,\tau}$, p) to TSP

In this section we describe a gap preserving reduction from the system of equations Hybrid($W_{k,\tau}$, p) to the Travelling Salesman problem. In the reduction we suppose that all equations of Hybrid($W_{k,\tau}$, p) with three variables are of the form $x \oplus y \oplus z = 0$ to simplify a discussion later (hence Hybrid($W_{k,\tau}$, p) was obtained via reduction from $Q_0(\varepsilon, 2k)$). We also introduce a real parameter $\theta > 0$ set to $\theta = \frac{1}{\max\{1, p_m\}}$, in order to simultaneously capture different scenarios $p_m \leq 1$ and $p_m > 1$.

The gap preserving reduction is similar to the reduction presented in Karpinski et al. (2015), the main difference is in using a parametrised weighted $(2k, \tau)$-bi-wheel amplifier ($W_{k,\tau}$, p) introduced in Sect. 3. We use the concept of forced edges introduced by Lampis (2014) (used also in Karpinski et al. (2015)). The idea is based on the observation that we are able to stipulate that some edges, called *forced* edges, are to be used at least once in any valid tour. It can be achieved by replacing such an edge with a path of many edges of the same total weight. With this trick we may assume without loss of generality that we can force some edges to be used at least once [see Karpinski et al. (2015) for the details]. If u and v are vertices that are connected by a forced edge e, we write $\{u, v\}_F$ or simply uv_F. The construction contains some forced edges, all other edges in the constructed graph are *unforced* edges with edge weight 1.

We start with an instance I of $Q_0(\varepsilon, 2k)$ with ν variables, m equations of the form $x \oplus y \oplus z = 0$ and use the reduction from Sect. 2 to create an instance J of Hybrid($W_{k,\tau}$, p). Using the same notation as in Theorem 3 we construct an instance $G[J]$ of TSP in the following way: for each copy $W_j := (W_{k,\tau}, p)$, $1 \leq j \leq \nu$, of a $(2k, \tau)$-bi-wheel we construct a subgraph of $G[J]$:

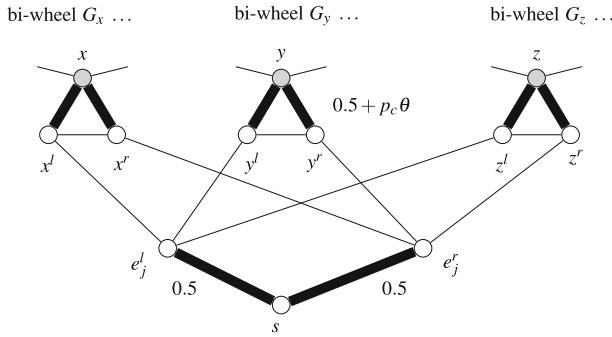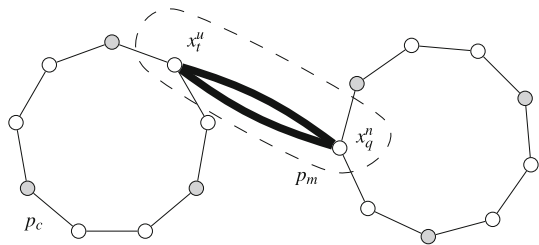(i) each variable x of the bi-wheel W_j, corresponds to a vertex x in the subgraph,

Fig. 1 An example of a 3-variable gadget H_j^{3Q} including the central vertex s, which is not part of the gadget. Thick lines represent forced edges.

Fig. 2 A gadget H^{2M} inside the bi-wheel G_x for the equations $x_t^u \oplus x_q^n = 1$ contains only two forced edges, represented as thick lines.

(ii) for each cycle equation $x \oplus y = 0$, we create an unforced edge xy with weight 1.

Now we add the edges among the vertices of 'bi-wheel' subgraphs using two types of gadgets:

- a 3-*variable gadget* H^{3Q}:
  for each equation j, $1 \le j \le m$, of the form $x \oplus y \oplus z = 0$ we add a 3-variable gadget H_j^{3Q} connecting the contacts x, y, z, where each contact vertex x, y, and z is part of its own $(2k, \tau)$-bi-wheel. Each gadget H_j^{3Q} contains two new vertices γ^l, γ^r for every vertex $\gamma \in \{x, y, z\}$ and two additional vertices e_j^l and e_j^r, see Fig. 1 how the vertices are connected. All edges $\{\gamma^\alpha, \gamma\}_F$ with $\alpha \in \{r, l\}$ and $\gamma \in \{x, y, z\}$ are forced edges with weight $w(\{\gamma^\alpha, \gamma\}_F) = 0.5 + p_c\theta$. All remaining edges of H_j^{3Q} are unforced with weight 1.
- a *matching gadget* H^{2M}:
  for each equation $x_t^u \oplus x_q^n = 1$ we add a matching gadget H^{2M} connecting the checkers x_t^u and x_q^n via two forced edges $\{x_t^u, x_q^n\}_F^1$ and $\{x_t^u, x_q^n\}_F^2$, each of the same weight $2p_c\theta$ (Fig. 2).

At the end of the construction, we add a new central vertex s that is connected to every gadget H_j^{3Q} with two forced edges $\{e_j^l, s\}_F$ and $\{e_j^r, s\}_F$, both with weight 0.5, $w(\{e_j^\alpha, s\}_F) = 0.5$ for both $\alpha \in \{r, l\}$.

Observe that the construction doesn't need gadgets for the cycle edges, the connections between the matching edge gadgets are sufficient to encode these constraints.

Now in the following we describe in the details the properties of the gap preserving reduction from the Hybrid($W_{k,\tau}$, p) to the TRAVELLING SALESMAN problem.

Local edge cost To count the cost $c(T)$ of a tour T, we use the local edge cost counting based on the ideas from Karpinski et al. (2015): the cost $w(uv)$ of any edge uv of T is split into two nonnegative parts, one attached to u and the second one to v. If an edge uv doesn't contain s then cost is split equally with contribution $0.5w(uv)$ for each vertex, but for edges of the form us, the full cost contributes to u, and none to s.

Let T be a multi-set of edges from E that defines a quasi-tour in $G[J](V, E)$. Then for a set $V' \subseteq V$, the local edge cost of V' is formally defined as

$$c_T(V') = \sum_{u \in V' \backslash \{s\}} \sum_{uv \in T} 0.5w(uv) + \sum_{e_j^\alpha \in V'} \sum_{e_j^\alpha s \in T} 0.5w(e_j^\alpha s).$$

Note that for two vertex sets V_1, V_2 we have $c_T(V_1 \cup V_2) \leq c_T(V_1) + c_T(V_2)$ (with equality for disjoint sets), and $c_T(V) = \sum_{e \in T} w(e)$.

In Sect. 4.2 we also use the full local cost of the quasi-tour T for the set V', $c_T^*(V')$, which is defined as follows: if $\#_T(V')$ is the number of connected components induced by T which are fully contained in V', then

$$c_T^*(V') = c_T(V') + 2\#_T(V').$$

Intuitively, $c_T^*(V')$ captures the cost of the full tour restricted to V': it includes the local edge cost and the cost of a connection of the components on V' of the lowest possible price (using two unforced edges), to the rest of the tour. Note that for two disjoint vertex sets V_1, $V_2 \subseteq V$ we have $c_T^*(V_1 \cup V_2) \geq c_T^*(V_1) + c_T^*(V_2)$.

4.1 How to construct a tour from an assignment

Given an instance J of the Hybrid($W_{k,\tau}$, p) and an assignment φ to its variables, we describe a construction of a tour T in $G[J]$ with cost related to φ.

Lemma 1 *Let J be an instance of Hybrid($W_{k,\tau}$, p) from Theorem 3. If there exists an assignment φ to the variables of J with unsatisfied equations of total weight Δ, then there exists a tour in $G[J]$ with cost at most*

$$\left(\frac{3}{2}(\tau - 1)(4p_c\theta + 1) + 6p_c\theta + 10\right)m + 2v + \Delta.$$

Proof **Step 1** First we describe a construction of a quasi-tour T that may have several connected components, but uses all forced edges, each exactly once. For each vertex the selection of edges chosen for T is based on type of the gadget to which the vertex belongs to, either a 3-variable gadget H^{3Q} or a matching gadget H^{2M}.

Part A: A matching gadget H^{2M}

Let $\{W_x, x \in \nu(I)\}$ be the associated set of bi-wheels of J. For a fixed $x \in \nu(I)\}$, let $\{x_j^u\}_{j=1}^{\tau k}$, $\{x_j^n\}_{j=1}^{\tau k}$ be the sets of the bi-wheel variables of W_x. Due to the amplifier properties, a given assignment φ can be converted into a standard one satisfying the same upper bound on the weight of unsatisfied equations [see Sect. 2 for the details]. Hence, we can assume that all equations with two variables are satisfied by the given assignment φ. It means, $x_t^u = x_q^u$, $x_t^n = x_q^n$, $x_t^u \neq x_q^n$ for all $t, q = 1, 2, \ldots, \tau k$. Assuming $x_t^\alpha = 1$ for some $\alpha \in \{u, n\}$, we include in T all unforced edges $x_t^\alpha x_{t+1}^\alpha$ with $t = 1, 2, \ldots, \tau k - 1$ (each edge once), the edge $x_{\tau k}^\alpha x_1^\alpha$, and both adjacent forced edges for each checker corresponding to matching equations (a copy of matching gadget H^{2M}). Informally, for each bi-wheel we select the cycle that corresponds to the assignment 1 and use all the unforced cycle edges including both forced edges from each checker. This creates a component of the quasi-tour T that contains all checkers from both cycles, and all contacts from one cycle.

The contribution of the vertices and chosen incident edges to the local edge cost of the quasi-tour T is $4p_c\theta$ for two forced edges of matching, and 1 for two unforced edges of one (checker) vertex of matching edge (one with $x_t^\alpha = 1$ for $\alpha \in \{u, n\}$), hence $4p_c\theta + 1$ for each matching edge $x_t^u x_q^n \in \mathcal{M}$.

Part B: A 3-variable gadget H^{3Q}

In this part we discuss how the quasi-tour T can traverse through the all vertices of the gadget H_j^{3Q}, $j = 1, 2, \ldots, m$. Following the construction of $G[J]$ each H_j^{3Q} simulates the equation $x \oplus y \oplus z = 0$. A selection of edges belonging to T is based on a given standard assignment φ to the contact variables x, y, and z. Depending on their values, we consider the following cases:

- Case $x + y + z = 2$: Suppose $y = z = 1$, $x = 0$, the other cases are symmetrical. Then we add the edges $z^l z^r$, $y^l y^r$, $e_j^l x^l$, and $e_j^r x^r$ to T which contributes 4 to the local edge cost of T restricted to the gadget H_j^{3Q}. The unforced edges adjacent to y and z contributes 2, hence 6 together for unforced edges.
- Case $x + y + z = 1$ or $x + y + z = 0$: In both cases we add for each vertex $\gamma \in \{x, y, z\}$ the edges $\gamma^l e_j^l$, $\gamma^r e_j^r$ to T with local unforced edge cost contribution of 6. If $x + y + z = 1$, we have to add 1 for edges adjacent to a vertex with assignment 1.
- Case $x + y + z = 3$: This case is similar to $x + y + z = 2$. We add the edges $z^l z^r$, $y^l y^r$, $e_j^l x^l$, and $e_j^r x^r$ to T with local unforced edge cost contribution 4 for the chosen edges inside H_j^{3Q}, 3 for edges adjacent to x, y, and z, hence 7 together.

The local edge cost contribution of the gadget H_j^{3Q} to T can be split into contribution of forced and unforced edges. The forced edges contribute $6(p_c\theta + 0.5) + 1 = 6p_c\theta + 4$ (our local count method allow to omit the vertex s, as weight of its adjacent edges is fully counted at the vertices e_j^l, e_j^r) and following the detailed discussion above, if the equation is satisfied the local unforced edge contribution to cost of T is 6, otherwise 7.

To summarise, in this way we construct a quasi-tour T for which we can count cost summing all local edge costs through all the gadgets (except s):

- there are $\frac{3}{2}(\tau - 1)m$ gadgets of size two (matching gadget H^{3Q}) with local edge cost $4p_c\theta + 1$, and
- there are m copies of 3-variable gadget H^{3Q}, each contributes $6p_c\theta + 10$ if satisfied, $6p_c\theta + 11$ otherwise.

The total edge cost of the quasi-tour T is thus at most

$$\left(\frac{3}{2}(\tau - 1)(4p_c\theta + 1) + 6p_c\theta + 10\right)m + \Delta.$$

Step 2 Observe that the quasi-tour T constructed in Step 1 may have at most $\nu + 1$ connected components (one for each bi-wheel and one containing s), which needs to be connected to create a tour.

Since all unforced edges have cost 1, we can connect components by double using an unforced edge that connects them (this is possible as $G[J]$ was a connected graph). This doesn't increase the cost of quasi-tour but decreases the number of components. Repeating this will result in a connected tour T of the cost

$$\left(\frac{3}{2}(\tau - 1)(4p_c\theta + 1) + 6p_c\theta + 10\right)m + 2\nu + \Delta. \qquad \square$$

4.2 How to define an assignment from a tour

Now we need to prove the opposite direction of the gap preserving reduction: given a tour in $G[J]$ the task is to define an assignment to the variables of the system equations I of Hybrid$(W_{k,\tau}, p)$ such that weight of unsatisfied equations is in a correlation with cost of a given tour.

Lemma 2 *If there is a tour in $G[J]$ with cost*

$$\left(\frac{3}{2}(\tau - 1)(4p_c\theta + 1) + 6p_c\theta + 10\right)m + \Delta - 2,$$

then there is an assignment to the instance J that leaves unsatisfied equations of weight at most $\Delta \cdot \max\{1, p_m\} = \frac{\Delta}{\theta}$, where $\theta = \frac{1}{\max\{1, p_m\}}$.

The high-level idea of the proof is to partition the vertex set of $G[J]$ into the gadget-based subgraphs similarly as in the proof of Lemma 1. For each such subgraph we give a lower bound on the local edge cost of any quasi-tour restricted to it, which in fact corresponds to cost of the tour constructed in Lemma 1. If a given quasi-tour behaves inside a gadget differently, its cost must be obviously higher. The difference between the full tour's local cost and the lower bound is called the *credit* of the gadget. Based on the tour we define an assignment for J and show that the total sum of credits can be used to bound from above the weight of unsatisfied equation, where the total sum of credits is at most Δ.

Proof Let $G[J](V, E)$ be the graph constructed at the beginning of Sect. 4. Obviously, we can suppose that T contains at most two copies of each edge. Otherwise we can

simply remove two copies of the same edge from T and decrease the cost of the initial tour. If T contains two copies of the same unforced edge, removing both copies from the tour results in a quasi-tour, where the number of the connected component may increase by 1. If this is the case, then the cost of 2, which corresponds to the cost of removed copies of unforced edges, is paid to reconnect the tour. Therefore removing two copies of the same edge doesn't increase the cost of the tour, but may result in a quasi-tour.

Thus, we can transform the initial tour T without increasing its initial cost into a quasi-tour without multiple unforced edges and with at most two copies of forced edges.

Consider the following partition of V: (A) $\frac{3}{2}(\tau-1)m$ sets of size 2 vertices ('matching pairs') corresponding to the gadgets H^{2M} covering all checker vertices, (B) m sets of size 11 vertices corresponding to the gadgets H^{3Q} for size-three equations, (C) a single vertex $\{s\}$.

Obviously, local edge cost of the subgraph from part (C) is 0 due to the method how we calculate the local cost of T (cost of all edges adjacent to s is fully contributed to the neighbourhoods of s), therefore part (C) doesn't need to be considered further.

In the following we analyse how the quasi-tour T can locally traverse the vertices inside the subgraphs in (A) and (B) and based on that we define an assignment to J.

As it follows from the construction, only the vertices from the bi-wheels correspond to the variables of J. Observe that each vertex of a bi-wheel is adjacent to exactly two forced and two unforced edges. For a given quasi-tour T, a bi-wheel vertex v (at the same time also variable) of $G[J]$ is called *honest* if the quasi-tour T uses both forced edges adjacent to v exactly once, otherwise the vertex is called *dishonest*. Now we define the *honest assignment* used for all honest variables (checkers and contacts): if the quasi-tour T for the honest vertex v also contains both unforced edges adjacent to it, then we assign 1 to v, and 0 otherwise.

Case A: The matching gadgets H^{2M}

In this part we define *the assignment to the checker variables* of the bi-wheels. Let $W := (W_{k,\tau}, p)$ be a fixed copy of the bi-wheel gadget H^{2M} for a variable $x \in v(I)$ and $M := \{x_t^u, x_q^n\}$ be a fixed matching pair from W.

The variables from M appear in the following equations: a matching equation $x_t^u \oplus x_q^n = 1$ and four incident cycle equations $x_{t-1}^u \oplus x_t^u = 0$, $x_t^u \oplus x_{t+1}^u = 0$, $x_{q-1}^n \oplus x_q^n = 0$, $x_q^n \oplus x_{q+1}^n = 0$ (indices are modulo $k\tau$).

The full local cost of any quasi-tour on M is $c_T^*(M) \geq 4p_c\theta + 1$: $4p_c\theta$ for two forced edges and 1 for two unforced edges, hence

$$cr_T(M) = c_T^*(M) - (4p_c\theta + 1). \tag{1}$$

Clearly, x_t^u is the honest vertex if and only if x_q^n is the honest vertex, therefore we split the discussion into two independent cases: honest matching pairs and dishonest matching pairs.

Case A.1 Honest checker variables

We use the honest assignment for all honest checker vertices: it means if the quasi-tour T contains both unforced edges adjacent to x_t^u (resp. x_q^n), then we assign 1 to x_t^u (resp. x_q^n), 0 otherwise.

If we use such assignment for the honest checkers in all matching pairs in W, the cycle equation $x_t^u \oplus y = 0$ (resp. $x_q^n \oplus y = 0$) is satisfied, when the adjacent vertex y is a honest checker. In other two cases, when y is a dishonest checker or a contact vertex, such equations are included in case of dishonest checkers (Case A.2) or contacts (case B).

Therefore we only need to discuss a matching equation $x_t^u \oplus x_q^n = 1$ in this part.

Obviously, if T selects two unforced edges for only one variable, hence one checker variable is set to 1 and the second to 0, $cr_T(M) = 0$, but the matching equation is satisfied, so no credit is needed.

Otherwise, a quasi-tour on M can either contain two unforced edges adjacent to both x_t^u, x_q^n or to be a connected component without any unforced edges. The full local cost of the quasi-tour T is in both cases $c_T^*(M) = 4p_c\theta + 2$, hence $cr_T(M) = 1$. In both these cases the honest assignment sets the same value to both variables, therefore the equation $x_t^u \oplus x_q^n = 1$ is not satisfied, but there is enough credit to pay for it.

To summarise, the cycle equations incident to M are either satisfied (if a neighbourhood is a honest checker) or they are considered later. If the matching equation is satisfied, then $cr_T(M) = 0$, if not then $cr_T(M) = 1$:

- If $p_m \leq 1$ the credit $cr_T(M)$ is sufficient to pay for the unsatisfied matching equation,
- If $p_m > 1$ then the credit is not sufficient, the weight of an unsatisfied equation is $p_m \cdot cr_T(M)$. That is why we aim in this scenario to bound from above the weight of unsatisfied equations by $\max\{1, p_m\} \cdot cr_T$.

Case A.2 Dishonest checker variables

If $M := \{x_t^u, x_q^n\}$ is a dishonest matching pair, then the quasi-tour must use one of the incident forced edges twice. Therefore, the full local cost of T is at least $6p_c\theta + 1$, hence the credit is $cr_T(M) \geq 2p_c\theta$ due to (1).

In this case we define an assignment to the variables from M in such a way that the matching equation $x_t^u \oplus x_q^n = 1$ is satisfied, and the credit is used for unsatisfied incident cycle equations.

Now we consider an assignment to the all dishonest matching pairs in W all it once and show that there is an assignment for dishonest checker variables for which at most half of the cycle equations incident to the dishonest checker variables is not satisfied.

Let $\mathscr{M}_{DH}$ be a set of dishonest matching pairs in W. Clearly,

$$\sum_{e \in \mathscr{M}_{DH}} cr_T(e) \geq 2p_c\theta|\mathscr{M}_{DH}|.$$

Let us denote $\mathscr{C}_{DH}$ the union of quadruples of cycle edges incident to all $e \in \mathscr{M}_{DH}$. Obviously, $|\mathscr{C}_{DH}| \leq 4|\mathscr{M}_{DH}|$.

Now consider random assignments to all pairs $\mathscr{M}_{DH}$ satisfying $x_t^u \oplus x_q^n = 1$ for every $x_t^u x_q^n \in \mathscr{M}_{DH}$. The expected number of unsatisfied cycle equations in $\mathscr{C}_{DH}$ is at most $\frac{1}{2}|\mathscr{C}_{DH}| \leq 2|\mathscr{M}_{DH}|$ so the weight of unsatisfied equations is at most

$$2p_c|\mathscr{M}_{DH}| \leq \frac{1}{\theta} \cdot \sum_{e \in \mathscr{M}_{DH}} cr_T(e) = \max\{1, \, p_m\} \sum_{e \in \mathscr{M}_{DH}} cr_T(e).$$

Hence similarly to Case A.1, the weight of unsatisfied equations is bounded from above by $\max\{1, p_m\} \cdot cr_T$.

Case B: The 3-variable gadget H^{3Q}

In this part we define the assignment to the contact variables of the bi-wheels.

Let $H^{3Q} := H_j^{3Q}$, $j \in \{1, \dots, m\}$, be a fixed 3-variable gadget on the vertex set $V^{3Q} = \{x, y, z, x^r, x^l, y^r, y^l, z^r, z^l, e^r, e^l\}$. We'll discuss the individual cases depending on how many vertices from $\{x, y, z\}$ are dishonest in H^{3Q} with regard to T.

Case B.1. All three vertices x, y, z are honest Any unforced edge is contained in the quasi-tour T at most once, therefore to keep the degree of the vertices x, y, z balanced in T, the quasi-tour T either contains both unforced edges or none for every x, y, z.

To balance the degree of the vertices γ^l, γ^r in T, where $\gamma \in \{x, y, z\}$, either the edges $\gamma^l e^l$, $\gamma^r e^r$ are in T, or T contains the edge $\gamma^l \gamma^r$. In the latter case the additional cost at least 1 must be added for connecting the quasi-tour on the vertex set γ, γ^l, γ^r to the rest of the tour either using two different unforced edges incident to γ or using a multiple unforced edge. Hence, the contribution of the vertex set γ^l, γ^r, γ to the cost of any tour is at least $3 + 2p_c\theta$, where the full cost of the edges $\gamma^l e^l$, $\gamma^r e^r$ is added to the vertices γ^l, γ^r to simplify counting. This holds for any vertex $\gamma \in \{x, y, z\}$, therefore the vertices x, y, z contribute together to the cost of T with at least $(9 + 6p_c\theta)$.

The degree of the vertices e^l, e^r may be balanced in T through the edges $\gamma^l e^l$, $\gamma^r e^r$, whose cost is already counted, but cost at least 1 needs to be added for forced edges $e^l s$, $e^r s$. (If the degree of the vertices e^l, e^r is not balanced in T through the edges $\gamma^l e^l$, $\gamma^r e^r$, then multiple used of forced edges $e^l s$, $e^r s$ contributes at least 2.)

To summarise, $c_T^*(V^{3Q}) \geq 6p_c\theta + 10$, and the credit of H^{3Q} with respect to the quasi-tour T is

$$cr_T(V^{3Q}) = c_T^*(V^{3Q}) - (6p_c\theta + 10). \tag{2}$$

In this subcase we use the honest assignment for the contact honest variables: for each $\gamma \in \{x, y, z\}$, we set $\gamma = 1$ if the quasi-tour T uses both unforced edges incident on γ, and 0 otherwise.

To recall, each variable γ appears in four cycle equations and the size-three equation. Due to the properties of the construction any contact vertex in the cycle is incident only to a checker: (i) if a honest checker is incident with a contact variable in the bi-wheel, the honest assignment sets the same value to both honest variables (1, if they are adjacent by an edge, 0 otherwise), therefore the corresponding cycle equations are

always satisfied; (ii) if a dishonest checker is incident with a contact variable in the bi-wheel, weight of all such unsatisfied cycle equations is considered in Case A.2 for dishonest checkers.

Therefore we need to prove that the credit $cr_T(V^{3Q})$ is enough to pay only for unsatisfied the size-three equation using the honest assignment to its contacts x, y, z. We'll discuss the possible options separately.

- If $x + y + z = 0$ or $x + y + z = 2$ then the equation $x \oplus y \oplus z = 0$ is satisfied. As it has been proved before, the full local cost of any tour T on the vertex V^{3Q} is at least $(6p_c\theta + 10)$, and therefore $cr_T(V^{3Q}) \geq 0$, but no credit is needed to pay for the size-three equation.
- If $x + y + z = 1$, then the assignment failed to satisfy the size-three equation of weight 1. We assume that $x = y = 0 \neq z$ holds, the other cases are symmetrical. Now we want to prove that under such constraints the full local cost of the tour T inside H^{3Q} is always at least $c_T^*(V^{3Q}) \geq 6p_c\theta + 11$, hence $cr_T(V^{3Q}) \geq 1$ to pay for the unsatisfied size-three equation. Let us discuss two subcases:

  - If there is an edge $\gamma^l\gamma^r \in T$ for some $\gamma \in \{x, y\}$, then we need to increase the cost about 2 for connecting a connected component on the vertex set $\{\gamma, \gamma^l, \gamma^r\}$ to the rest of the tour. Hence $c_T^*(V^{3Q}) \geq 6p_c\theta + 11$, therefore $cr_T(V^{3Q}) \geq 1$ in this case as it follows from (2).
  - If this is not a case, then the only way how the vertices x^r, x^l, y^r, y^l are balances with respect to T is $e^l\gamma^l$, $e^r\gamma^r \in T$ for both $\gamma \in \{x, y\}$. Because the vertices e^l, e^r must also be balanced in T, we obtain that either e^rz^r, $e^lz^l \in T$ or alternatively, the forced edges e^rs_F, e^ls_F were used more than once together with z^rz^l. Again, $c_T^*(V^{3Q}) \geq 6p_c\theta + 11$, hence $cr_T(V^{3Q}) \geq 1$.

- If $x = y = z = 1$, then similarly, the assignment failed to satisfy the size-three equation with weight 1 and we need to prove that $c_T^*(V^{3Q}) \geq 6p_c\theta + 11$ for the tour T which includes both unforced edges for each vertex x, y, z.

If T contains all three edges x^lx^r, y^ly^r, z^lz^r, then all the vertices γ, γ^l, γ^r, for each $\gamma \in \{x, y, z\}$ are balanced in T, therefore the edges γ^re^r, γ^le^l can't be in the quasi-tour T and the degree of e^r, e^l must be balanced with multi-use of the forced edges e^rs_F, e^ls_F. Therefore, $c_T^*(V^{3Q}) \geq 6p_c\theta + 11$.

If there exists $\gamma \in \{x, y, z\}$ such that γ^le^l or γ^re^r are in T, then both γ^le^l, γ^re^r must be in T and the vertex set γ, γ^l, γ^r (including the full cost of edges γ^le^l, γ^re^r) contributes at least $4 + 2p_c\theta$ to the cost of T, therefore $c_T^*(V^{3Q}) \geq 6p_c\theta + 11$.

Summary for the case of x, y, z being honest contacts: if size-three equation is not satisfied then $cr_T(V^{3S}) \geq 1$; all incident cycle equations are either satisfied or are considered in Case A.2 of dishonested checkers.

Case B.2. At least one of the vertices x, y, z is dishonest

We start with the following observation.

Observation Suppose that a quasi-tour T is dishonest at a vertex $\gamma \in \{x, y, z\}$. Then it is possible to extend the initial honest assignment defined for the set of honest variables from $\{x, y, z\}$ (this can also be an empty set) such that the size-three equation $x \oplus y \oplus z = 0$ is satisfied, and there is sufficient credit to pay for one cycle equation incident to γ. □

Proof of Observation If γ is incident to at least one dishonest checker then the corresponding cycle equation is covered in Case A.2, hence at most one incident cycle equation needs to be discussed here

If both cycle neighbours adjacent to γ are honest, as γ is the dishonest vertex, due to the balanced degree at γ in T, only one of the unforced edges from γ can be used in the quasi-tour T. Consequently, one of cycle neighbours of γ is set to 1 and the other to 0. Therefore we can chose any assignment for γ (e.g. the one satisfying the size-three equation) and show that we have enough credit to pay for one cycle equation associated with γ.   □

In what follows we explain that for each dishonest variable $\gamma \in \{x, y, z\}$ there is always sufficient credit to pay for one cycle equation. The discussion will depend on the number of the dishonest variables from $\{x, y, z\}$.

Obviously, for any dishonest variable $\gamma \in \{x, y, z\}$ only one of the two incident forced edges is used twice in a quasi-tour T. If this is not a case then one copy of the edges $\gamma^r \gamma_F$, $\gamma^l \gamma_F$ can be replaced by $\gamma^r \gamma^l$ reducing the cost of T by $2p_c\theta > 0$.

- Suppose all three vertices x, y, z use a forced edge twice in T. The cost of forced edges inside H^{3Q} is at least $9(p_c\theta + 0.5) + 1$, and the unforced edges incident to x, y, z contribute 1.5 to the local edge cost in order to balance the degree at these vertices in T. Three vertices of the form γ^α, for $\gamma \in \{x, y, z\}$ and $\alpha \in \{l, r\}$ have odd degree in T and any quasi-tour must use three unforced edges inside H^{3Q} to balance their degrees. Hence, $c_T^*(V^{3Q}) \geq 9p_c\theta + 10$, and $cr_T(V^{3Q}) \geq 3p_c\theta$ due to (2). Therefore, if we chose an assignment satisfying the size-three equation, there is sufficient credit $cr_T(V^{3Q}) \geq 3p_c\theta$ to pay for 3 cycle equations which is enough due to the observation above.
- Suppose exactly two of the vertices x, y, z use a forced edge twice in T, let's assume x, y (the other cases are symmetrical). For the honest variable z we use the *honest assignment*: 1, if both unforced edges incident to z are used in T; 0 otherwise. Then obviously, the cycle equations for edges incident with z are either satisfied or are considered in Case A.2. The contribution of the forced edges inside H^{3Q} to the cost of T is at least $8(p_c\theta + 0.5) + 1$, and unforced edges incident to x, y contributes to the local edge cost by 1. One of the vertices x^l, x^r and one of y^l, y^r have odd degree in T, therefore any quasi-tour needs to use two unforced edges from H^{3Q} to balance the degrees in T which increases the cost by 2.

Finally, the degree of the vertices z^l, z^r must be balanced in the quasi-tour T. If two unforced edges incident to z^l, z^r are in T, we have the full local cost $c_T^*(V^{3Q}) \geq 8p_c\theta + 10$, hence $cr_T(V^{3Q}) \geq 2p_c\theta$. Otherwise if the edge $z^l z^r$ is in T, then either two unforced edges incident on z are used or the quasi-tour pays 2 extra for a component, so again $c_T^*(V^{3Q}) \geq 8p_c\theta + 10$ and $cr_T(V^{3Q}) \geq 2p_c\theta$.

Therefore we can select an assignment for x, y that together with the honest assignment for z satisfies the size-three equation in such a way that there is sufficient credit to pay for at most 2 cycle equations incident on x, y, which is enough following the observation above.

- Suppose that only one of x, y, z has an incident forced edge used twice in T, let's assume x (the other cases are symmetrical). For y and z we set the honest assign-

ment and similarly to the previous case, the cycle equations for edges incident to y and z are either satisfied or are considered in Case A.2.

The contribution of the forced edges to the full local cost of T is at least $7(p_c\theta + 0.5) + 1$, and the local edge cost 0.5 for one unforced edge incident to x.

Also, at least one unforced edge incident on x^l or x^r is used, since one of them has odd degree in T. For y^l, y^r either two unforced edges are used to balance their degree in T, or if $y^l y^r \in T$, then either unforced edges incident on y contribute 1 (or the quasi-tour pays extra for a component), hence the contribution to the cost to balance the degree of the vertices y^l, y^r is at least 2. With the same reasoning for z^l, z^r, we get that the full local cost is $c_T^*(V_j^{3Q}) \geq 7p_c\theta + 10$, hence $cr_T(V^{3Q}) \geq p_c\theta$.

Now selecting an assignment for x that satisfies the size-three equation, we have sufficient credit to pay for one cycle equation incident on x which is enough due to observation above.

Summary for the quasi-tour with a dishonest vertex: we have an option to choose an assignment for x, y, z that satisfies the size-three equation; and the weight of unsatisfied incident cycle equations that are not considered in Case A.2 can be bounded from above by $\frac{1}{\theta} cr_T(V^{3Q}) = \max\{1, p_m\} \cdot cr_T(V^{3Q})$.

Let us now conclude our analysis. To calculate the cost of the quasi-tour T, the vertex set is partitioned into two types of gadget-based subgraphs covering the entire set V which were considered separately:

– Case A: the size of 2 gadgets H^{2M} ('matching pairs'), there are $\frac{3}{2}(\tau - 1)m$ such gadgets;
– Case B: the size of 11 gadgets H^{3Q} ('size-three equation'), there are m such gadgets; the edges incident to s are counted in this case even if s is not part of the gadget.

Following the discussion in both cases, the cost of any quasi-tour must be at least $\left(\frac{3}{2}(\tau - 1)(4p_c\theta + 1) + 6p_c\theta + 10\right)m$ and following the assumption of the theorem, the sum of full local costs must be

$$c_T^*(V) \leq \left(\frac{3}{2}(\tau - 1)(4p_c\theta + 1) + 6p_c\theta + 10\right)m + \Delta,$$

where the additional cost 2 may need to be added to connect the component with the singleton s to the rest of the tour. Therefore the sum of all credits cr_T is at most Δ. Since we have already argued that the $\max\{1, p_m\} \sum cr_T$ bounds from above the weight of all equations unsatisfied by the defined assignment in both cases locally, this concludes the proof. $\qquad\square$

Theorem 4 *If (p_c, p_m, τ) is an admissible triple then it is* NP-*hard to approximate the* TRAVELLING SALESMAN *problem to within any constant approximation ratio less than*

$$1 + \frac{1}{3(\tau - 1)(4p_c + \max\{1, p_m\}) + 12p_c + 20\max\{1, p_m\}}.$$

Proof Let $\varepsilon \in (0, \frac{1}{4})$. Consider a $(2k, \tau)$-bi-wheel with large enough k, which is an amplifier with cycle weights p_c and matching weights p_m. We have instances of Hybrid$(W_{k,\tau}, p)$ with v copies of a bi-wheel $(W_{k,\tau}, p)$, m equations of the form $x \oplus y \oplus z = 0$ each of weight 1, $3\tau m$ equations of the form $x \oplus y = 1$ each of weight p_m with the following NP-hard gap results: It is NP-hard to decide whether there is an assignment to the variables that leaves unsatisfied equations of weight at most εm, or every assignment to the variables leaves unsatisfied equations of weight at least $(0.5 - \varepsilon)m$. Due to Lemma 1 and 2 we now know that for produced instances $G[J]$ of TSP it is NP-hard to decide whether there is a tour with cost at most $\left(\frac{3}{2}(\tau - 1)(4p_c\theta + 1) + 6p_c\theta + 10\right)m + 2v + \varepsilon m$, where $\theta = \frac{1}{\max\{1, p_m\}}$ or all tours have cost at least $\left(\frac{3}{2}(\tau - 1)(4p_c\theta + 1) + 6p_c\theta + 10\right)m + (0.5 - \varepsilon)m \cdot \theta - 2$.

The ratio between these two cases can get arbitrarily close to

$$1 + \frac{1}{3(\tau - 1)(4p_c + \max\{1, p_m\}) + 12p_c + 20\max\{1, p_m\}}$$

by appropriate choices of $\varepsilon > 0$ and large enough k. $\qquad\square$

5 Random construction of bi-wheel amplifiers

As it follows directly from Theorem 4, the better bounds for the random constructions of bi-wheel amplifiers can directly improve lower bounds for approximability of the TSP. In this section we only focus on case $p_c < 1$ which will be enough to improve known lower bounds for TSP.

Theorem 5 *The triple $(p_c = \frac{1}{2}, p_m = 1, \tau = 11)$ is admissible, hence for every large enough $k \geq k_0$ there is a $(2k, 11)$-bi-wheel that is an amplifier with cycle weights $p_c = \frac{1}{2}$ and matching weights $p_m = 1$.*

Proof Let an integer $k > 0$ be large enough. Consider a bi-wheel $W_{k,11}$ with two disjoint cycles, each on $11k$ vertices,

$$V^u = \{1^u, 2^u, \ldots, (11k)^u\} \text{ and } V^n = \{1^n, 2^n, \ldots, (11k)^n\}$$

with sets of k contacts $D = D^u \cup D^n$, where

$$D^u = \{11^u, (11 \cdot 2)^u, \ldots, (11 \cdot k)^u\} \text{ and } D^n = \{11^n, (11 \cdot 2)^n, \ldots, (11 \cdot k)^n\},$$

the remaining vertices $V^u \backslash D^u$ and $V^n \backslash D^n$ are checkers. That means, that any two consecutive contacts in the bi-wheel cycles are separated by a chain of 10 checkers. For a given subset $A \subseteq V$, the fragments of A are all connected components of A within cycles of V^u and V^n together.

Now select uniformly at random a perfect matching from checkers $V^u \backslash D^u$ to checkers $V^n \backslash D^n$. We claim that with a high probability a bi-wheel created in this way is a weighted amplifier with weights $p_c = \frac{1}{2}$ and $p_m = 1$.

Such a randomly constructed bi-wheel fails to have required amplification properties if the system of so-called *bad sets*,

$$\mathscr{B} := \{B \subseteq V : p(E(B, \overline{B})) < \min\{|D \cap B|, |D \cap \overline{B}|\}$$

is nonempty. If $\mathscr{B} \neq 0$, it also contains a *minimal bad set*, where by *minimal* we mean an element of $\mathscr{B}$ that is minimal with respect to the partial order $\preceq$ defined for any $A, B \subseteq V$ as:

$$A \preceq B \text{ iff } |(|(A \triangle B) \cap D| \leq p(E(B, \overline{B})) - p(E(A, \overline{A}))$$

and if $A \cap D = B \cap D$ and $p(E(B, \overline{B})) = p(E(A, \overline{A}))$ then the number of fragments of A is at most the number of fragments of B.

These minimal bad sets have certain more restrictive properties that allow better estimation of the upper bound of probability that such sets B exist in a randomly constructed bi-wheel. These results are derived in details in Chlebík and Chlebíková (2003) for the case of wheel-amplifiers. The methods and results can be extended in straightforward way to the weighted bi-wheel case as well.

Let us mention some of these properties: given a bad set B, then

- Fact 1: Any fragment of a minimal bad set B (and of $\overline{B}$) contains at least 2 vertices.
- Fact 2: If an end vertex of a fragment of a minimal bad set B is a checker x, then this checker is matched with a vertex from B as well. The same is true for fragments of $\overline{B}$.

In our strategy to prove that probability that a minimal bad set exists is < 1, we need to understand how for a given subset $S \subseteq V \backslash D$ of checkers to estimate from above the probability (over the random matchings of $V^u \backslash D^u$ to $V^n \backslash D^n$) that S is the trace of a minimal bad set B in the checker vertices $V \backslash D$, i.e., $S = B \backslash D$.

We want to take the sum of these probabilities over all potentially minimal bad sets and prove that the sum is < 1. It follows that with high probability, no set is a minimal bad set and therefore a randomly constructed bi-wheel is a weighted amplifier with parameters ($p_c = \frac{1}{2}$, $p_m = 1$, $\tau = 11$).

Now suppose that $S \subseteq V \backslash D$ with $|S| = w$, what is the probability that there are exactly c matching edges with exactly one endpoint in S?

We can observe that the most interesting case is the balanced one, namely when S contains equal (or nearly equal) number of checkers from each side V^u, V^n, as follows:

Suppose that S contains $\frac{w}{2} + l$ checkers on one side, and $\frac{w}{2} - l$ checkers on the side, $l \geq 0$. The probability that exactly c matching edges have exactly one endpoint in S is

$$P(w, c, l) =$$
$$= \binom{\frac{w}{2} + l}{\frac{c}{2} + l} \binom{\frac{w}{2} - l}{\frac{c}{2} - l} \binom{10k - \frac{w}{2} - l}{\frac{c}{2} - l} \binom{10k - \frac{w}{2} + l}{\frac{c}{2} + l} \times$$
$$\times \left(\frac{c}{2} + l\right)! \left(\frac{c}{2} - l\right)! \frac{\left(\frac{w}{2} - \frac{c}{2}\right)! \left(10k - \frac{w}{2} - \frac{c}{2}\right)!}{(10k)!}$$

We observe that in such scenario to create exactly c matching edges between S and $\overline{S}$ we need to select more endpoints, $\frac{c}{2} + l$, on the side where S is larger, since after we remove checkers matched from S to $\overline{S}$ we must have a perfect matching for the rest of S (and $\overline{S}$).

Interestingly, $P(w, c, l)$ achieves maximum for $l = 0$. To see that, one can simplify $\frac{P(w,c,l+1)}{P(w,c,l)}$ as $\left(1 - \frac{2l+1}{\frac{c}{2}+l+1}\right)\left(1 + \frac{2l+1}{\frac{11}{2}-l}\right)\left(1 + \frac{2l+1}{10k-\frac{w}{2}-l}\right)$ and to prove that this quotient is for relevant range of parameters less than < 1 (using the fact $1 + x \le e^x$).

We consider several parameters for a potentially minimal bad set such as $\min\{|D \cap B|, |D \cap \overline{B}|\}$; the number of fragments of B in cycles V^u and V^n together (this is the same as the number of fragments of $\overline{B}$). As $p(E(B, \overline{B})) < \min\{|D \cap B|, |D \cap \overline{B}|\}$, it also gives estimate $f < \min\{|D \cap B|, |D \cap \overline{B}|\}$.

Also, with $S := B \backslash D$ we have an upper bound on the possible c for which we have exactly c matching edges with exactly one endpoint in S, namely

$$c + f < \min\{|D \cap B|, |D \cap \overline{B}|\}.$$

For any feasible vector of parameters we have in hand the probability $P(w, c, l)$ computed above. Using restrictions of a minimal bad set we estimate from above the number of ways how such a bad set can be generated. We use simple sum of probabilities for the union bound. Using Stirling's formula and the binary entropy function for bounding $\frac{1}{k}$ log of that probability, we can see that it stays negative and bounded away from zero.

This will complete the proof that with high probability we constructed bi-wheel above amplifiers with parameters ($p_c = \frac{1}{2}$, $p_m = 1$, $\tau = 11$). □

6 Conclusion

The methods of this paper provide a new motivation for the study of expanding properties of random graphs. As we have demonstrated, introducing the parametrised weighted amplifiers and weighted low occurrence CONSTRAINT SATISFACTION problems as intermediate steps in the NP-hard gap reductions, allows more flexibility in fine-tuning their expanding parameters. We show that already slight improvement of known expander values modestly improve the hardness of approximation for TSP from the current best value $\frac{123}{122}$ (Karpinski et al. 2015) to the new value $\frac{117}{116}$. The introduced method of weighted amplifiers (or expanders) can be of independent interest. Such technique could be used in the gap preserving reductions for other edge-weighted optimisation problems to improve their approximation hardness results.

References

Böckenhauer HJ, Hromkovič J, Klasing R, Seibert S, Unger W (2000) An improved lower bound on the approximability of metric TSP and approximation algorithms for the TSP with sharpened triangle inequality. In: Proceedings of the 17th annual symposium on theoretical aspects of computer science , STACS'00. Springer, Berlin, pp 382—394

Chlebík M, Chlebíková J (2003) Approximation hardness for small occurrence instances of NP-hard problems. In: Algorithms and complexity, 5th Italian conference, CIAC 2003, Lecture Notes in Computer Science, vol 2653, pp 152–164

Chlebík M, Chlebíková J (2008) The Steiner Tree problem on graphs: inapproximability results. Theor Comput Sci 406(3):207–214

Chlebík M, Chlebíková J (2019) Approximation hardness of Travelling Salesman via weighted amplifiers. In: Computing and combinatorics—25th international conference, COCOON 2019, Springer, Lecture Notes in Computer Science, vol 11653, pp 115–127

Christofides N (1976) Worst-case analysis of a new heuristic for the Travelling Salesman problem. Technical Report 388, Graduate School of Industrial Administration, Carnegie Mellon University

Engebretsen L (1998) An explicit lower bound for TSP with distances one and two. Algorithmica 35:301–319

Gharan SO, Saberi A, Singh M (2011) A randomized rounding approach to the Traveling Salesman problem. In: 2011 IEEE 52nd annual symposium on foundations of computer science. IEEE, pp 550–559

Håstad J (2001) Some optimal inapproximability results. J ACM 48(4):798–859

Held M, Karp RM (1970) The traveling-salesman problem and minimum spanning trees. Oper Res 18(6):1138–1162

Karpinski M, Lampis M, Schmied R (2015) New inapproximability bounds for TSP. J Comput Syst Sci 81(8):1665–1677

Lampis M (2014) Improved inapproximability for TSP. Theory Comput 10:217–236

Mömke T, Svensson O (2011) Approximating graphic TSP by matchings. In: 2011 IEEE 52nd annual symposium on foundations of computer science. IEEE, pp 560–569

Mucha M (2014) 13/9-approximation for graphic TSP. Theory Computi Syst 55(4):640–657

Papadimitriou CH, Vempala SS (2006) On the approximability of the traveling salesman problem. Combinatorica 26(1):101–120

Papadimitriou CH, Yannakakis M (1993) The traveling salesman problem with distances one and two. Math Oper Res 18(1):1–11

Sebő A, Vygen J (2014) Shorter tours by nicer ears: 7/5-approximation for the graph-TSP, 3/2 for the path version, and 4/3 for two-edge-connected subgraphs. Combinatorica 34(5):597–629

Svensson O (2013) Overview of new approaches for approximating TSP. In: International workshop on graph-theoretic concepts in computer science. Springer, pp 5–11