On modelling hard combinatorial optimisation problems as linear programs: refutations of the 'unconditional impossibility' claims

Moustapha Diaby*

OPIM Department,
University of Connecticut,
Storrs, CT 06268, USA
Email: moustapha.diaby@uconn.edu
*Corresponding author

Mark H. Karwan and Lei Sun

Department of Industrial and Systems Engineering,
SUNY at Buffalo,
Amherst, NY 14260, USA
Email: mkarwan@buffalo.edu
Email: leisun@buffalo.edu

Abstract: There has been a series of developments in the recent literature (by essentially a same 'circle' of authors) with the absolute/unconditioned (implicit or explicit) claim that there exists no abstraction of an *NP-complete* combinatorial optimisation problem in which the defining combinatorial configurations [such as 'tours' in the case of the travelling salesman problem (TSP) for example] can be modelled by a *polynomial-sized* system of linear constraints. The purpose of this paper is to provide general as well as specific refutations for these recent claims.

Keywords: linear programming; combinatorial optimisation; computational complexity; travelling salesman problem; TSP; P vs. NP.

Reference to this paper should be made as follows: Diaby, M., Karwan, M.H. and Sun, L. (2021) 'On modelling hard combinatorial optimisation problems as linear programs: refutations of the 'unconditional impossibility' claims', *Int. J. Operational Research*, Vol. 40, No. 2, pp.261–278.

Biographical notes: Moustapha Diaby is an Associate Professor of Production and Operations Management, University of Connecticut. He received his PhD in Management Science/Operations Research, MS in Industrial Engineering and BS in Chemical Engineering from University at Buffalo – The State University of New York. His areas of interest are mathematical programming, manufacturing systems modelling and analysis, operations and logistics management, and project management. He has many (60+) refereed publications in books, journals, and conference proceedings. He serves/has served as a reviewer and/or ad-hoc editorial team member for many of the professional journals and for government agencies in the USA and Canada.

Mark H. Karwan, Praxair Professor in Operations Research and SUNY Distinguished Teaching Professor, has served 42 years in the Department of Industrial and Systems Engineering, University at Buffalo; Department Chair (1987–1992) and Dean of Engineering and Applied Sciences (1994–2006). His 32 PhD students have been guided in modeling and algorithmic development mathematical programming, multiple criteria decision making. His 100+ publications show diverse application areas: logistics, production planning, hazardous waste routing, military path planning and analytics. Funding has come from NSF, ONR, AFRL and industry. His industry consulting includes: 1) production planning, routing, forecasting, cost/pricing; 2) defense industry – logistics and dynamic resource allocation.

Lei Sun is a Lead Data Scientist at Praxair Inc. and a Research Assistant Professor in the Department of Industrial and Systems Engineering at University at Buffalo – The State University of New York. He received his PhD in Industrial Engineering from University at Buffalo in 2013. His research interests include network optimisation, machine learning, supply chain, logistics and business analytics.

1 Introduction

Combinatorial optimisation problems (COPs) occupy a central place in operations research (OR) and in mathematics in general. Part of the reason for this is the wide practical applicability of COPs, as virtually every scheduling, sequencing, or routing problem that arises in industry involves a COP. COPs are also central for the field of theoretical computer science (TCS) in particular, because COPs serve as a foundation for the study of *computatibity,* which in turn, has to do with the very foundations of mathematics in general, and specifically the issues of the generalisability of the *axiomatic approach* (Hilbert, 1901) and the so-called *incompleteness theorem* (Gödel, 1931). Since the beginnings of their field, operations researchers have approached COPs from an 'engineering' perspective, with a focus on developing/'engineering' methods aimed at obtaining ideal or satisfactory solutions to the practical problems they model. The perspectives in TCS and mathematics in general, on the other hand, have been much more abstract, with a focus on what is/is not possible in terms of *computability* (Gödel, 1931).

A convergence between the OR and TCS perspectives was brought about by the landmark invention of the theory of *NP-completeness* (Cook, 1971; see also Garey and Johnson, 1979) on one hand, and the landmark developments of *interior-point methods* (Katchian, 1979; Karmarkar, 1984) on the other hand. *NP-completeness* delineated the so-called 'NP' class of problems, and within it, the subclass 'P' of problems that are easily 'computable' according to some theoretical measure, thereby essentially reducing the question of *computability* to the question of whether the 'P' and 'NP' classes are equal. With *interior-point methods* came the discovery that the 'engineering' problems called 'linear programs (LPs)' (see Bazaraa et al., 2006) fall within the 'P' class of problems, and this reduced the *computability* question for engineers to that of whether or not any of the representatives of the 'hard' NP-class of problems (i.e., the so-called *NP-complete* problems) could be modelled/'engineered' as a *polynomial-sized* LP. A

COP is said to have been 'modelled as an LP' if it has been abstracted into an LP which has integral extreme points corresponding to the combinatorial configurations being decided upon. An LP model is said to be of *polynomial size* if its numbers of variables and constraints are respectively bounded by polynomial functions of a measure of the *size* of the input data for the problem being modelled.

As far as we know, the first *polynomial-sized* LP models to be proposed for an *NP-complete* problem [specifically, the travelling salesman problem (TSP); see Lawler et al., 1985] are those of Swart (1986, 1987). Not having/having had access to Swart's papers, we do not know of any of the details of his models. However, the concensus of the research communities is that their validities were refuted by Yannakakis (1991). The essence of the Yannakakis (1991) developments is the showing that for the TSP, the polytope which is stated in terms of its extreme points in the space of the 'natural' city-to-city variables (i.e., the so-called '*the* TSP polytope'; see Lawler et al., 1985) does not have a *polynomial-sized extended formulation* (EF) which is *symmetric* (see Yannakakis, 1991). Roughly, a system of linear constraints (say, 'system A') is an EF (or an *extension*) of another system of linear constraints (say, 'system B') if the polytopes they respectively induce in the space of the 'system B' variables coincide (see Yannakakis, 1991; Balas, 2005; among others). If 'system A' is an EF of 'system B', the polytope induced by 'system A' is said to be an EF of the polytope induced by 'system B', and also, 'system B' (or the polytope it induces) is said to be the *projection of 'system A'* (or the polytope it induces) *on the space of 'system B'*. Clearly, a given system of equations cannot induce a polytope in the space of variables which are not part of that system. Hence, as developed in Diaby and Karwan (2016, 2017), the addition of redundant variables (and possibly, constraints also) to a system of constraints leads to *degenerate*/non-meaningful EF relationships with respect to the task of making inferences about model *sizes*.

Yannakakis' (1991) work is carefully scoped. However, it fails to make appropriate explicit exceptions for cases involving zero-matrices in its analyses, thereby failing to explicitly distinguish between *degenerate*/non-meaningful EFs from meaningful ones with respect to the purposes of those analyses. We argue that the lack of recognition of this distinction is what has led to the increasingly over-scoped claims in the recent EF literature which attempts to build on Yannakakis' (1991) approach. Specifically, the claim in Fiorini et al. (2011, 2012) is that the natural descriptions of the travelling salesman and stable set polytopes do not have *polynomial-sized EF*s, regardless of the *symmetry* condition of Yannakakis (1991). Theoretical and numerical refutations of these Fiorini et al. (2011, 2012) developments are provided in Diaby and Karwan (2016, 2017). In Avis and Tiwary (2015), Braun et al. (2015), Fiorini et al. (2015) [which is essentially the 'journal version' of Fiorini et al. (2011, 2012)] and Averkov et al. (2018) respectively, the claim is the absolute/unconditioned (implicit or explicit) statement that there exists no abstraction of an *NP-complete* COP in which the defining combinatorial configurations (such as 'tours' in the case of the TSP for example) can be modelled by a *polynomial-sized* system of linear constraints. The purpose of this paper is to provide general as well as specific refutations for these recent claims.

The plan of the paper is as follows. We will discuss general-level refutations in Section 2, where we will provide a *polynomial-sized* system of linear constraints which correctly abstracts TSP tours, and also derive conditions for the existence of an affine map (which a linear map is a special case of) between disjoint sets of variables with no implication nor need for *extended formulations* (EF) relationships. Specific refutations will be discussed in Section 3, using the Fiorini et al. (2015) developments. Finally, some concluding remarks will be offered in Section 4.

2 General refutations

The fundamental presumption upon which the developments in the 'unconditional impossibility' papers (Avis and Tiwary, 2015; Braun et al., 2015; Fiorini et al., 2015; Averkov et al., 2018) rest is that the *size* of the description of a polytope in the space of its variables can be inferred from that of another polytope stated in a disjoint space of variables. The objective of this section is to provide general, direct refutations of this misconception, and using the TSP, of the over-reaching 'impossibility claim' itself. We will first recall the definition of the standard TSP polytope (i.e., the so-called '*the* TSP polytpe'). Then, we will present an alternate, *polynomial-sized* system of linear constraints which correctly abstracts TSP tours. Finally, we will show that the existence of an affine map between disjoint sets of variables [as was brought to our attention by Kaibel et al. (2013)] is not a sufficient condition for the existence of EF relationships from which valid comparisons between the descriptions of polytopes stated in those (disjoint) spaces can be made. For the discussions about the TSP polytopes, we will use the following conventions.

Assumption 1

We assume without loss of generality that:

1 a city designated as '0' is the beginning and ending point of all travels

2 the TSP tours have been ordered, with the k^{th} one designated by T_k ($k \in \{1, \ldots, (n-1)!\}$).

Notation 1

1 n: number of cities.

2 $m := n - 1$.

3 $\Omega := \{0, \ldots, m\}$: index set of the cities.

4 $\mathcal{A} := \left\{(i, j) \in \Omega^2 : i \neq j\right\}$: set of arcs of the TSP (city-to-city) graph; set of possible TSP 'travel legs'.

5 $M := \Omega \backslash \{0\}$: set of cities to visit when city '0' is considered the starting and ending point of the travels.

6 $S := \{1, \ldots, m\}$: index set for the 'times-of-visit' for the cities in M (see Picard and Queyranne, 1978).

7 $\forall (i, j) \in \mathcal{A}$, x_{ij}: variable indicating whether city i is visited immediately before city j ($x_{ij} = 1$), or not ($x_{ij} = 0$).

8 $\forall (i,s) \in (M,S)$, w_{is} : variable indicating whether city i is visited at 'time' s ($w_{is} = 1$), or not ($w_{is} = 0$).

9 $Conv(A)$: convex hull of A.

Definition 1: [Standard TSP polytope: '*the* TSP polytope']

$\forall\ F \subseteq \mathcal{A} : F \neq \varnothing$, let $x^F := \{x \in \{0,1\}^{n(n-1)} : x_{ij} = 1 \text{ iff } (i,j) \in F\}$. The standard TSP polytope (i.e., '*the* TSP polytope') is defined as $Conv(\{x^{T_k},\ (k = 1, \ldots, (n-1)!)\})$.

2.1 There exists an alternate 'TSP polytope'

In this section, we will present a *polynomial-sized* system of linear constraints which correctly abstracts TSP tours, and illustrate it with a numerical example.

Theorem 1: The extreme points of

$$AP := \left\{ \mathbf{w} \in \mathbb{R}^{(n-1)^2} : \sum_{s \in S} w_{is} = 1 \quad \forall i \in M; \quad \sum_{i \in M} w_{is} = 1 \quad \forall s \in S; \quad \mathbf{w} \geq \mathbf{0} \right\}$$

are in one-to-one correspondence with TSP tours which start and end at city '0'.

Proof: Polytope AP is the standard linear assignment problem (or *Birkhoff*) polytope, and has therefore, integral extreme points (see Birkhoff, 1946; Burkard et al., 2009; among others). Moreover, using the assumption that city '0' is the starting and ending point of travel, it is trivial to construct a unique TSP tour from a given extreme point of AP, and vice versa (i.e., it is trivial to construct a unique extreme point of AP from a given TSP tour), as shown below.

Let $\mathbf{w}^k$ $(k \in \{1, \ldots, (n-1)!\})$ denote the k^{th} extreme point of AP, with corresponding set of assignments $C^k := \{((a_p^k, p) \in (M,S), \quad p = 1, \ldots, n-1 : (\forall (p,q) \in S^2 : p \neq q,\ a_p^k \neq a_q^k)\}$. Then, the components of $\mathbf{w}^k$ are specified as follows:

$$\forall\,(i,s) \in (M,S)\,,\ w_{is}^k = \begin{cases} 1 & \text{if } i = a_s^k; \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

The order of visits for the unique TSP tour (T_k) corresponding to $\mathbf{w}^k$ is: $0 \longrightarrow a_1^k \longrightarrow \ldots \longrightarrow a_{n-1}^k \longrightarrow 0$.

Conversely, let T_k $(k \in \{1, \ldots, (n-1)!\})$ denote the k^{th} TSP tour, with order of visits specified as: $0 \longrightarrow a_1^k \longrightarrow \ldots \longrightarrow a_{n-1}^k \longrightarrow 0$ (where $a_p^k \in M$ for $p = 1, \ldots, n-1$). The unique extreme point of AP, $\mathbf{w}^k$, corresponding to T_k is obtained by applying (1) above. ■

Remark 1

1 AP is distinct from the permutahedron (i.e., the convex hull of all vectors that arise from permutations) of the TSP cities.

2 AP does not induce TSP tours *per se* (i.e., Hamiltonian cycles of the TSP cities; see Lawler et al., 1985).

3 It is not possible to make AP induce TSP tours *per se* by adding a 'dummy' city to the set of cities and using it as the starting and ending point of the travels. The reason for this is that it would not necessarily be possible to associate the AP solution thus obtained to a TSP tour, as the actual cities of the first and last *times-of-visit* may be different.

4 The vertices of AP are linear assignment problem (LAP; see Burkard et al., 2009) solutions, whereas the vertices of '*the* TSP polytope' model Hamiltonian cycles. Hence, AP and '*the* TSP polytope' are mathematically-unrelated polytopes.

5 The association that can be made between AP and '*the* TSP polytope' is *cognitive* only.

6 According to the *Minkowski-Weyl theorem* [Minkowski, 1910; Weyl, 1935; see also Rockafellar, (1997), pp.153–172], every polytope can be equivalently described as the intersection of hyperplanes ($\mathcal{H}$-representation/external description) or as a convex combination of (a finite number of) vertices ($\mathcal{V}$-representation/internal description). '*The* TSP polytope' is easy to state in terms of its $\mathcal{V}$-representation. However, no *polynomial-sized* $\mathcal{H}$-representation of it is known. On the other hand, the $\mathcal{H}$-representation of AP is well-known to be of (low-degree) *polynomial size* (see Burkard et al., 2009), and it is trivial to state its $\mathcal{V}$-representation also.

7 In order to model the TSP optimisation problem as an LP in the space of the natural, *travel-leg* x_{ij} variables, an $\mathcal{H}$-representation of '*the* TSP polytope' must be developed. This task has thwarted all efforts so far. In order to model the TSP optimisation problem as an LP in the space of the LAP, *travel-time* w_{ir} variables, a linear function which correctly captures TSP tours costs must be developed. Hence, mathematical developments focused on what is/is not possible to do for the linear system modelling of '*the* TSP polytope' would have no pertinence in abstractions which are based on AP and do not require the natural x_{ij} variables. Examples of such abstractions are given in Diaby (2007) and Diaby and Karwan (2016). □

The results below follow directly from the discussions above.

Corollary 1

1 AP is an alternate 'TSP polytope' from the standard TSP polytope (i.e., the so-called '*the* TSP polytope').

2 AP is a refutation of the claim in the recent literature (Avis and Tiwary, 2015; Braun et al., 2015; Fiorini et al., 2015; Averkov et al., 2018) that the combinatorial configurations which define an *NP-complete* problem cannot be abstracted into a *polynomial-sized* system of linear constraints.

Definition 2

1 The standard TSP polytope may be referred to as the *travel-legs (TL) TSP polytope*, and its extreme points may be referred to as *travel-legs (TL) TSP tours*.

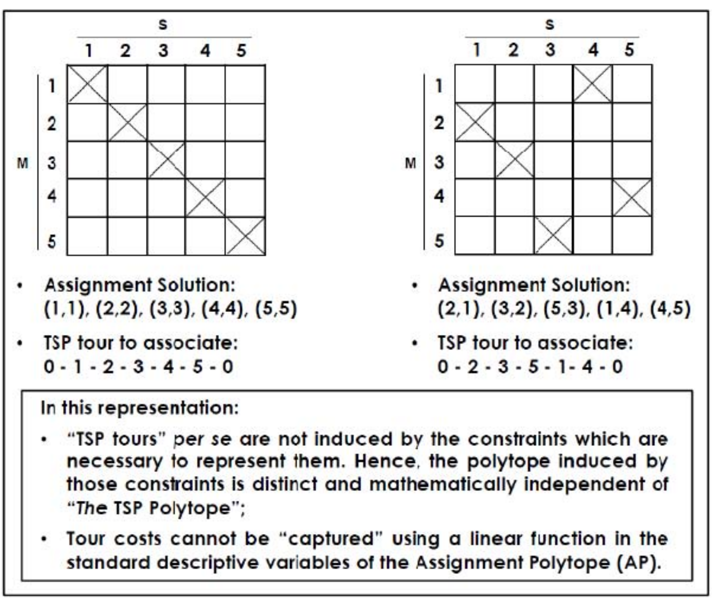
2 AP may be referred to as the *travel-times (TT) TSP polytope (rooted at '0')*, and its extreme points may be referred to as *travel-times (TT) TSP tours*.

We will now provide a numerical illustration of the discussions above.

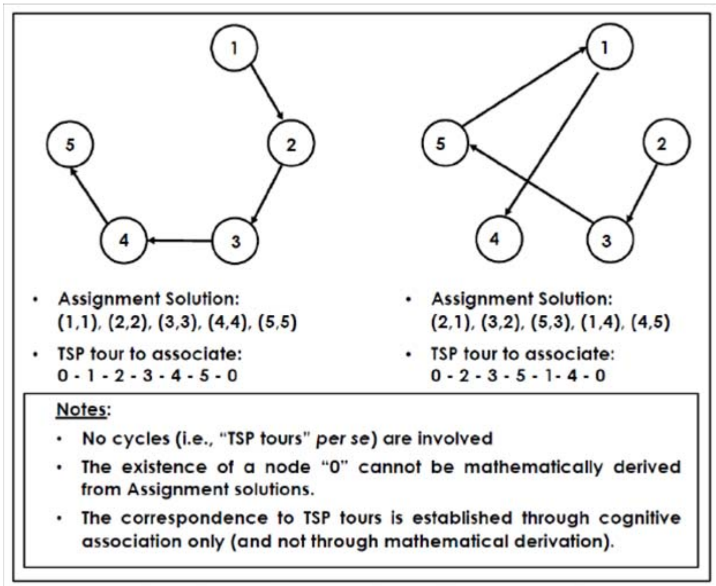
Example 1

Some of the differences between the *travel-times TSP polytope* (AP) and the *travel-legs TSP polytope* ('*the* TSP polytope') discussed above will now be illustrated using a 6-city TSP with node set $\{0, 1, 2, 3, 4, 5\}$.

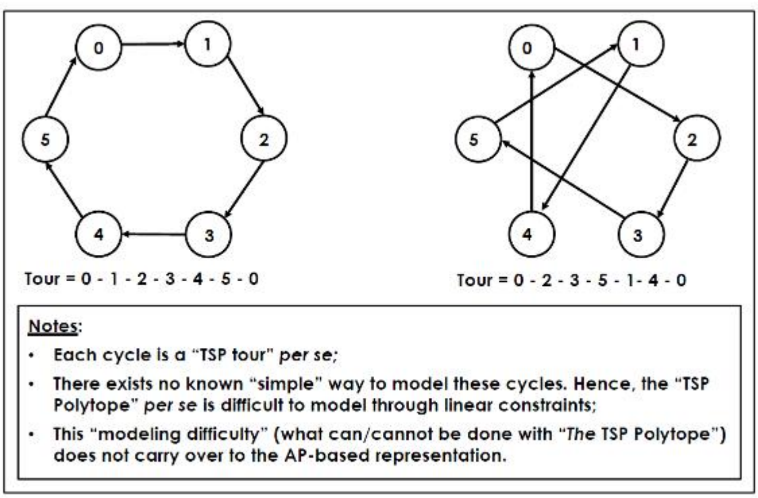
- Illustration of the *travel-times TSP tours* on the AP graphical tableau:



- Illustration of the *travel-times TSP tours* on the TSP graph:



- Illustration of the *travel-legs TSP tours* on the TSP graph:



□

2.2 *The existence of an affine map does not necessarily imply extension relationships*

In the field of OR, it is common for two formulations of a problem expressed in terms of disjoint sets of variables (say x and y) to be equivalent. In such a case, if an affine mapping between the sets of variables is known (say from y to x), one may add its expression to the model stated in the space of the 'domain' variables (y) and project the resulting *augmented* model onto the space of the 'range' variables (x) for the purposes of comparing the *strengths* of the bounds that can be obtained from the two formulations in a same space of variables [see Balas, (2005), pp.132–136]. The validity of this approach comes from the fact that the addition of redundant variables and constraints to (i.e., an *augmentation* of) an optimisation model does not change the objective function value of that model. Note however, that a polytope cannot have a constraints description in the space of variables which are not part of its set of descriptive variables. Hence, as shown in Diaby and Karwan (2016, 2017), it is not valid to use the projection of an *augmentation* of a model in order to make inferences about the *size* of the constraints description of that model if the vector space being projected to is disjoint from that of the model. Hence, as has been mentioned earlier in this paper and also in Diaby and Karwan (2016, 2017), the belief/presumption that the existence of an affine map between disjoint sets of variables describing different polytopes is sufficient to imply EF relationships from which valid inferences about model *sizes* can be made is a misconception. We will now demonstrate this by deriving a sufficient condition for such an existence with no necessity nor implication of *extension* relations. We will also offer an alternate interpretation of such an existence in an optimisation context.

Theorem 2

Let:

- $x \in \mathbb{R}^p$ and $y \in \mathbb{R}^q$ be disjoint vectors of variables
- $X := \{x \in \mathbb{R}^p : Ax \leq a\} \neq \varnothing$ (where: $A \in \mathbb{R}^{k \times p}$; $a \in \mathbb{R}^k$)
- $L := \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^{p+q} : Bx + Cy = b \right\} \neq \varnothing$ (where: $B \in \mathbb{R}^{m \times p}$; $C \in \mathbb{R}^{m \times q}$; $b \in \mathbb{R}^m$)
- $Y := \{y \in \mathbb{R}^q : Dy \leq d\} \neq \varnothing$ (where: $D \in \mathbb{R}^{l \times q}$; $d \in \mathbb{R}^l$).

Then, the following are true:

1 Provided $B^T B$ is non-singular, there exists a one-to-one affine function from $\mathbb{R}^q$ to $\mathbb{R}^p$ which maps y onto x.

2 Assume the following conditions are true:

   a $B^T B$ is non-singular

   b the constraints of L are *redundant* for X and Y respectively

   c L is the graph of a one-to-one correspondence between the points of X and the points of Y [see Beachy and Blair, (2006), pp.47–59].

Then, the optimisation of any linear function of x over X can be done without any reference to the constraints description of X.

Proof:

1 From the definition of L, we have:

$$Bx = b - Cy. \tag{2}$$

Pre-multiplying (2) by B^T gives:

$$B^T Bx = B^T b - B^T Cy. \tag{3}$$

Using the non-singularity of $B^T B$ (according to the premise) and pre-multiplying (3) by $(B^T B)^{-1}$, we get:

$$x = (B^T B)^{-1} B^T b - (B^T B)^{-1} B^T Cy. \tag{4}$$

Expression (4) can be written as:

$$\begin{aligned} &x = \overline{C}y + \overline{b}, \\ &\text{where } \overline{C} := -(B^T B)^{-1} B^T C, \text{ and } \overline{b} := (B^T B)^{-1} B^T b. \end{aligned} \tag{5}$$

2 Consider the task of minimising the function $\alpha^T x$ ($\alpha \in \mathbb{R}^p$) over X. This optimisation problem can be expressed as:

Problem LP_0:

$$\left| \begin{array}{ll} \text{Minimise:} & \alpha^T x \\ \text{Subject To:} & x \in X. \end{array} \right.$$

Since the constraints of L are *redundant* for X and Y respectively (according to premise b), LP_0 is equivalent to:

Problem LP_1:

$$\left| \begin{array}{ll} \text{Minimise:} & \alpha^T x \\ \text{Subject to:} & \begin{pmatrix} x \\ y \end{pmatrix} \in L; \;\; x \in X; \;\; y \in Y. \end{array} \right.$$

Using (5) (since $B^T B$ is non-singular according to premise a) to eliminate x from the objective of *Problem LP_1*; using the fact that the constraints of L are *redundant* for X and Y respectively (according to premise b) to eliminate the constraints of L from *Problem LP_1*; and using the fact that L is the graph of a one-to-one correspondence between the points of X and the points of Y (according to premise c) to eliminate X from the constraints set of *Problem LP_1*, gives that *Problem LP_1* can be solved using the following two-step procedure:

Step 1 Solve *Problem* LP_2:

$$\left| \begin{array}{ll} \text{Minimise:} & \left(\alpha^T \overline{C}\right) y + \alpha^T \overline{b} \\ \text{Subject To:} & y \in Y \end{array} \right.$$

Let the solution be y^*.

Step 2 Use Graph L to 'retrieve' the optimal solution x^* to *Problem* LP_1:

$$x^* = \overline{C} y^* + \overline{b}.$$

Note that the second term of the objective function of *Problem* LP_2 can be ignored in the optimisation process of *Problem* LP_2, since that term is a constant.

Also, if L is derived from knowledge of the vertex ('$\mathcal{V}$-') representation of X only (as would be the case if X were '*the* TSP polytope' for example), then this would mean that the $\mathcal{H}$-representation of X is not involved in the 'two-step' solution process above, but rather, that only the $\mathcal{V}$-representation of X is involved.

Part 2 of Theorem 2 is similar to Proposition 2 of Padberg and Sung (1991, p.323).

3 Specific refutations

In all of the recent EF developments with the claim that an *NP-complete* COP cannot be abstracted into a *polynomial-sized* LP (i.e., Avis and Tiwary, 2015; Braun et al., 2015; Fiorini et al., 2015; Averkov et al., 2018), results are developed for the natural polytopes of COPs using 'slack matrices'-based concepts (such as 'extension complexity'; 'approximation complexity'; 'relaxation complexity'; etc.), and their generalisations to all arbitrary abstractions of COPs hinges (invariably) on the use of the notion that the existence of linear maps (which are special-cases of affine maps) between disjoint sets of variables implies EF relationships between polytopes stated in the spaces of those sets of variables. As discussed in Diaby and Karwan (2016, 2017), this notion is *degenerate* in the sense that it allows for every conceivable pair of polytopes to be EFs of each other provided they are non-empty and are stated in terms of disjoint sets of variables. Hence, the key to pinpointing the inherent mathematical flaws in all of these recent EF papers is to focus on their generalisation steps (which invariably need this *degenerate* EF notion). We will illustrate this in this section using the Fiorini et al. (2015) developments, after providing a brief overview of the background definitions which are used interchangeably in the 'line of research'.

3.1 Background definitions

Definition 3 ('Standard EF Definition'): An *extended formulation* for a polytope $P = \{x \in \mathbb{R}^d \ : Ax \leq b\} \subseteq \mathbb{R}^d$ is a polyhedron $Q = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^{d+k} : Ex + Fy \leq g \right\}$, the projection of which onto x-space, $\varphi_x(Q) := \left\{ x \in \mathbb{R}^d : \left(\exists y \in \mathbb{R}^k : \begin{pmatrix} x \\ y \end{pmatrix} \in Q \right) \right\}$, is equal to P (where $A \in \mathbb{R}^{n \times d}$, $b \in \mathbb{R}^n$, $E \in \mathbb{R}^{m \times d}$, $F \in \mathbb{R}^{m \times k}$, and $g \in \mathbb{R}^m$) (Yannakakis, 1991).

Definition 4 ('Fiorini et al. Definition 1'): A polyhedron $Q = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^{d+k} : Ex + Fy \leq g \right\}$ is an *extended formulation* of a polytope $P \subseteq \mathbb{R}^d$ if there exists a linear map $\pi : \mathbb{R}^{d+k} \longrightarrow \mathbb{R}^d$ such that P is the image of Q under π (i.e., $P = \pi(Q)$; where $E \in \mathbb{R}^{m \times d}$, $F \in \mathbb{R}^{m \times k}$, and $g \in \mathbb{R}^m$) [Fiorini et al., (2015), p.17:3, lines 20–21; p.17:9, lines 22–23].

Definition 5 ('Fiorini et al. Definition 2'): An *extended formulation* of a polytope $P \subseteq \mathbb{R}^d$ is a linear system $Q = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^{d+k} : Ex + Fy \leq g \right\}$ such that $x \in P$ if and only if there exists $y \in \mathbb{R}^k$ such that $\begin{pmatrix} x \\ y \end{pmatrix} \in Q$. (In other words, Q is an EF of P if $\left(x \in P \Longleftrightarrow \left(\exists y \in \mathbb{R}^k : \begin{pmatrix} x \\ y \end{pmatrix} \in Q \right) \right)$ (where $E \in \mathbb{R}^{m \times d}$, $F \in \mathbb{R}^{m \times k}$, and $g \in \mathbb{R}^m$) [Fiorini et al., (2015), p.17:2, last paragraph; p.17:9, line 20–21].

Remark 2

1 Because every equality constraint in an optimisation problem can be replaced by a pair of inequality constraints, the expression of Q in the definitions above is general. However, the equality constraints are sometimes separated out in the EF papers, and E, F, and g are partitioned by rows as $E^{=}$, $E^{\leq}$, $F^{=}$, $F^{\leq}$, $g^{=}$, and $g^{\leq}$, respectively, so that Q is written as:

$$Q = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^{d+k} : E^{\leq}x + F^{\leq}y \leq g^{\leq};\ E^{=}x + F^{=}y = g^{=} \right\} \tag{6}$$

(where $E^{\leq} \in \mathbb{R}^{m^{\leq} \times d}$, $E^{=} \in \mathbb{R}^{m^{=} \times d}$, $F^{\leq} \in \mathbb{R}^{m^{\leq} \times k}$, $F^{=} \in \mathbb{R}^{m^{=} \times k}$, $g^{\leq} \in \mathbb{R}^{m^{\leq}}$, and $g^{=} \in \mathbb{R}^{m^{=}}$, with $m^{\leq} + m^{=} = m$.)

2 Definition 3 is the standard, reference EF definition.

3 Definitions 4 and 5 are alternate EF definitions which are used interchangeably in Fiorini et al. (2015) and other recent EF papers.

4 Definition 5 is consistent with Definition 3.

5 Definition 4 is inconsistent with Definition 3 (and therefore, with Definition 5 also), when the description of Q does not involve the x-variables (i.e., when $E = \mathbf{0}$; see Diaby and Karwan, 2016, 2017).

6 The use of Definition 4 in the recent EF papers is what allows them to generalise their results beyond the natural polytopes of COPs. However, this use of Definition 4 is also what makes the mathematics in those EF papers inherently flawed. This will be illustrated in the remainder of this section.

□

3.2 Inherently-flawed mathematics: Illustration using Fiorini et al. (2015)

Example 2: Let $\mathbf{x} \in \mathbb{R}^3$ and $\mathbf{y} \in \mathbb{R}$ be disjoint vectors of variables. Let P be a polytope in the space of $\mathbf{x}$, and Q, a polytope in the space of $\begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix}$, with:

$$P := Conv\left(\left\{\begin{pmatrix} 8 \\ 10 \\ 6 \end{pmatrix}, \begin{pmatrix} 12 \\ 15 \\ 9 \end{pmatrix}\right\}\right), \text{ and} \tag{7}$$

$$Q := \left\{\begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} \in \mathbb{R}^{3+1} : 2 \leq \mathbf{0} \cdot \mathbf{x} + \mathbf{y} \leq 3\right\}. \tag{8}$$

We will now discuss some key results of Fiorini et al. (2015) which are refuted by P and Q.

1 *Refutation of the validity of Definition 4.*

a Note that the following is true for P and Q:

$$\mathbf{x} \in P \not\Leftrightarrow \left(\exists \mathbf{y} \in \mathbb{R} : \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} \in Q\right). \tag{9}$$

For example,

$$\left(\exists \mathbf{y} \in \mathbb{R} : \begin{pmatrix} 22.5 \\ -50 \\ 100 \\ y \end{pmatrix} \in Q\right) \not\Rightarrow \left(\begin{pmatrix} 22.5 \\ -50 \\ 100 \end{pmatrix} \in P\right). \tag{10}$$

Hence, Q **is not** an *extended formulation* of P according to Definition 5.

b Observe that the following is also true for P and Q:

$$P = \left\{\mathbf{x} \in \mathbb{R}^3 : \mathbf{x} = \pi \cdot \begin{pmatrix} \mathbf{z} \\ \mathbf{y} \end{pmatrix}, \mathbf{z} \in \mathbb{R}^3, \begin{pmatrix} \mathbf{z} \\ \mathbf{y} \end{pmatrix} \in Q\right\} \tag{11}$$

$$= \pi(Q); \tag{12}$$

$$\text{where } \pi = \begin{bmatrix} 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & 3 \end{bmatrix}. \tag{13}$$

In other words, P is the image of Q under the linear map π. Hence, Q **is** an *extended formulation* of P according to Definition 4.

c It follows from a and b above, that Definitions 4 and 5 are in contradiction of each other with respect to P and Q.

Hence P and Q are a refutation of the validity of Definition 4, since Definition 5 is equivalent to Definition 3, which is the standard/reference definition (as indicated in Remarks 2.2 and 2.4).

2 *Refutation of 'Theorem 3' of Fiorini et al. (2015, p.17:10).*

The proof of the theorem ('Theorem 3') hinges on Definition 4. The specific statement in Fiorini et al. (2015, p.17:10, lines 26–28) is:

> "*...Because*
>
> $$Ax \leq b \iff \exists y : E^{=}x + F^{=}y = g^{=},\ E^{\leq}x + F^{\leq}y \leq g^{\leq}, \tag{14}$$
>
> *each inequality in* $Ax \leq b$ *is valid for all points of* Q. ..."

The equivalent of (14) in terms of P and Q [using the partitioned form (6) for Q] is:

$$\mathbf{x} \in P \iff \exists \mathbf{y} \in \mathbb{R} : \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} \in Q. \tag{15}$$

Clearly, (15) is **not true**, since it is in contradiction of expressions (9)-(10) above. Hence, the proof of 'Theorem 3' (and therefore, 'Theorem 3') of Fiorini et al. (2015) is refuted by P and Q.

3 *Refutation of 'Lemma 9'* of Fiorini et al. (2015, pp.17:13–17:14).

The first part of the lemma is stated (in Fiorini et al., 2015) thus:

> "*Lemma 9. Let* P, Q*, and* F *be polytopes. Then, the following hold:*
>
> *(i) if F is an extension of P, then* xc(F) $\geq$ xc(P); …"

The proof of this (in Fiorini et al., 2015) is stated as follows:

> "*Proof: The first part is obvious because every extension of F is in particular an extension of P.* …"

The notation '$xc(\cdot)$' in these statements (of Fiorini et al., 2015) stands for '*extension complexity* of $(\cdot)$', which is defined as [Fiorini et al., (2015), p.17:9, lines 24–25]:

> "*...the extension complexity of P is the minimum size (i.e., the number of inequalities) of an EF of P.*"

The refutation of the Fiorini et al. (2015) *Lemma 9* for P [as shown in (7) above] and Q [as shown in (8) above] is as follows.

First, note that (8) can be re-written in its explicit form as:

$$Q = \left\{(\mathbf{x}, \mathbf{y}) \in (\mathbb{R}^3, \mathbb{R}) : 2 \leq y \leq 3\right\}. \tag{16}$$

As shown in Part (1) above (in this proof), Q is an *extension* of P according to Definition 4 [which is central in Fiorini et al. (2015)]. Accordingly, therefore, this means that Q is an *extended formulation* of every one of the infinitely-many possible $\mathcal{H}$-descriptions of P. This would be true in particular for the $\mathcal{H}$-description below for P:

$$P = \left\{ \begin{array}{l} \mathbf{x} \in \mathbb{R}^3 : \\ -5x_1 + 4x_2 \leq 0; \\ 3x_2 - 5x_3 = 0; \\ 3x_1 - 4x_3 \leq 0; \\ 8 \leq x_1 \leq 12; \\ 10 \leq x_2 \leq 15; \\ 6 \leq x_3 \leq 9 \end{array} \right\}. \tag{17}$$

Clearly, the number of inequalities in (17) is greater than the number of inequalites in (16). In other words, for P and Q, we have that:

$$xc(Q) \ngeq xc(P). \tag{18}$$

Hence, P and Q are a refutation of 'Lemma 9' of Fiorini et al. (2015), being that Q is the *extension*, and P, the *projection,* according to the definitions used in Fiorini et al. (2015).

□

Remark 3

1 According to Fiorini et al. (2015, p.17:7, Section 1.4, first sentence; p.17:11, lines 6–11; p.17:14, lines 5–6; p.17:16, lines 13–14 after 'Figure 4' caption), their 'Theorem 3' and 'Lemma 9' play pivotal, foundational roles in the rest of their developments. Note that 'Lemma 9' (of Fiorini et al., 2015) does not depend on any one of the *extended formulations* definitions used in Fiorini et al. (2015) in particular. Hence, we believe the numerical illustration we have provided above represents a simple-yet-complete refutation of the developments in Fiorini et al. (2015).

2 A 'feature' of Q in our counter-example above is that its minimal inequality ('outer' or '$\mathcal{H}$-') description does not require the x-variables. Hence, P and Q in the example essentially have disjoint sets of descriptive variables. Hence, as shown in Diaby and Karwan (2016, 2017), the EFs relationship which would be created between the two polytopes by the addition of the expression of the linear map in (11)-(13) to the description of Q is *degenerate*/meaningless with respect to the task of making valid inferences about the *size* of the $\mathcal{H}$-description of Q from the *size* of the $\mathcal{H}$-description of P. The reason for this is that the derivation of the linear map involves the extreme-point ('inner-' or '$\mathcal{V}$-') description of P only, as detailed in Diaby and Karwan (2016, 2017).

□

4 Conclusions

In this paper, we have provided a multi-levelled refutation of the claim in some of the recent *extended formulations* (EFs) papers (Avis and Tiwary, 2015; Braun et al., 2015; Fiorini et al., 2015; Averkov et al., 2018) that an *NP-complete* problem cannot be abstracted into a *polynomial-sized* linear program (LP). One of the two fundamental misconceptions in those papers is the (sometimes-implicit) belief that all abstractions of a combinatorial optimisation problem (COP) *must* involve the polytope stated in terms of the natural variables for that COP (for example, '*the* TSP Polytope' for the TSP). We have provided a direct refutation of this misconception by exhibiting a *polynomial-size* LP model which we have shown to correctly abstract TSP tours. The other misconception in the 'unconditional impossibility' papers is the belief that the existence of an affine map between disjoint sets of variables implies EF relationships from which valid inferences about model *sizes* can be made. In order to refute this misconception on a general level, we have developed conditions for such existence independently of any implication or need for an *extension* relationship. We have also clarified the meaning of the existence of such a mapping in an optimisation context.

A consequence of the presumption that the existence of an affine map between disjoint sets of variables implies meaningful/non-degenrate EF relationships between polytopes in the spaces of those sets of variables is that it introduces inherent flaws in the mathematics of the papers that use it in their efforts to make inferences about model *sizes*. Focusing on this, we have provided counter-example refutations of the key foundational results of Fiorini et al. (2015) (namely, their 'Theorem 3' and 'Lemma 9') upon which (according to them) their claims that:

> "We solve this question by proving a super-polynomial bound on the number of inequalities in every LP for the TSP." [Fiorini et al., (2015), p.17:2, lines 9–10]

> "We also prove such unconditional super-polynomial bounds for the maximum cut and the maximum stable set problems." [Fiorini et al., (2015), p.17:2, lines 10–12]

and

> "... it is impossible to prove $P = NP$ by means of a polynomial-sized LP that expresses any of these problems." [Fiorini et al., (2015), p.17:2, lines 12–13]

are based. In other words, we have shown that these claims of Fiorini et al. (2015) are not supported by the mathematics in their paper. The approach we have used in order to do this can be readily applied to all of the other papers in this 'line of research', including Braun et al. (2015) and Averkov et al. (2018).

As far as we know, very few papers have been published, which offer alternate abstractions of *NP-complete* COPs which do not require their natural variables. Some exceptions are Diaby (2007, 2010a, 2010b, 2010c), Maknickas (2015), and Diaby and Karwan (2016). Our suggestion for future directions is that research focus be shifted away from developments aimed at showing negative results for the natural polytopes of COPs in favor of efforts directed at the exploration of novel approaches which may yield low-dimensional alternate abstractions instead.

References

Averkov, G., Kaibel, V. and Weltge, S. (2018) 'Maximum semidefinite and linear extension complexity of families of polytopes', *Mathematical Programming A*, Vol. 167, No. 2, pp.381–394.

Avis, D. and Tiwary, H.R. (2015) 'On the extension complexity of combinatorial polytopes', *Mathematical Programming A*, Vol. 153, No. 1, pp.95–115.

Balas, E. (2005) 'Projection, lifting, and extended formulation in integer and combinatorial optimization', *Annals of Operations Research*, Vol. 140, No. 1, pp.125–161.

Bazaraa, M., Sherali, H.D. and Shetty, C.M. (2006) *Nonlinear Programming – Theory and Algorithms*, Wiley, New York, NY.

Beachy, J.A. and Blair, W.D. (2006) *Abstract Algebra*, Waveland Press, Inc., Long Grove, IL.

Birkhoff, G. (1946) 'Three observations on linear algebra', *Universidad Nacional de Tucumán. Revista A.*, Vol. 5, pp.147–151.

Braun, G., Fiorini, S., Pokutta, S. and Steurer, D. (2015) 'Approximation limits of linear programs (beyond hierarchies)', *Mathematics of Operations Research*, Vol. 40, No. 3, pp.756–772.

Burkard, R., Dell'Amico, M. and Martello, S. (2009) *Assignment Problems*, SIAM, Philadelphia.

Cook, S.A. (1971) 'The complexity of theorem-proving procedures', *Proceedings of the Third Annual ACM Symposium on Theory of Computing*, pp.151–158.

Diaby, M. (2007) 'The traveling salesman problem: a linear programming formulation', *WSEAS Transactions on Mathematics*, Vol. 6, No. 6, pp.745–754.

Diaby, M. (2010a) 'Linear programming formulation of the multi-depot multiple traveling salesman problem with differentiated travel costs', in Davendra, D. (Ed.): *Traveling Salesman Problem, Theory and Applications*, pp.257–282, InTech, New York, NY.

Diaby, M. (2010b) 'Linear programming formulation of the set partitioning problem', *International Journal of Operational Research*, Vol. 8, No. 4, pp.399–427.

Diaby, M. (2010c) 'Linear programming formulation of the vertex coloring problem', *International Journal of Mathematics in Operational Research*, Vol. 2, No. 3, pp.259–289.

Diaby, M. and Karwan, M.H. (2016) *Advances in Combinatorial Optimization: Linear Programming Formulations of the Traveling Salesman and Other Hard Combinatorial Optimization Problems*, World Scientific Publishing, Singapore.

Diaby, M. and Karwan, M.H. (2017) 'Limits to the scope of applicability of extended formulations', *International Journal of Mathematics in Operational Research*, Vol. 10, No. 1, pp.18–33.

Fiorini, S., Massar, S., Pokutta, S., Tiwary, H.R. and de Wolf, R. (2011) *Linear vs. Semidefinite Extended Formulations: Exponential Separation and Strong Bounds*, Unpublished [online] http://arxiv.org/pdf/1111.0837.pdf (accessed 1 March 2012).

Fiorini, S., Massar, S., Pokutta, S., Tiwary, H.R. and de Wolf, R. (2012) 'Linear vs. semidefinite extended formulations: exponential separation and strong bounds', *Proceedings of the 44th ACM Symposium on the Theory of Computing (STOC '12)*, New York, NY, pp.95–106.

Fiorini, S., Massar, S., Pokutta, S., Tiwary, H.R. and de Wolf, R. (2015) 'Exponential lower bounds for polytopes in combinatorial optimization', *Journal of the ACM*, Vol. 62, No. 2, Article No. 17.

Garey, M.R. and Johnson, D.S. (1979) *Computers and Intractability: A Guide to the Theory of NP-Completeness*, Freeman and Company, New York, NY.

Gödel, K. (1931) 'Über Formal Unentscheidbare Sätze der Principia Mathematica und Verwandter Systeme, I', *Monatshefte für Mathematik und Physik*, Vol. 38, No. 1, pp.173–198.

Hilbert, D. (1901) 'Mathematische Probleme', *Archiv für Mathematik und Physik*, Vol. 1, pp.213–237.

Kaibel, V. (2011) 'Extended formulations in combinatorial optimization', *Optima*, Vol. 85, No. 2, pp.2–7.

Kaibel, V., Pokutta, S. and Yannakakis, M. (2013) *Private Email Communications*.

Karmarkar, N. (1984) 'A new polynomial-time algorithm for linear programming', *Combinatorica*, Vol. 4, pp.373–395.

Katchian, L.G. (1979) 'A polynomial algorithm in linear programming', *Soviet Mathematics Doklady*, Vol. 20, pp.191–194.

Lawler, E.L., Lenstra, J.K., Rinnooy Kan, A.H.G. and Shmoys, D.B. (Eds.) (1985) *The Traveling Salesman Problem: A Guided Tour of Combinatorial Optimization*, Wiley, New York.

Maknickas, A.A. (2015) 'How to solve kSAT in polynomial time', *International Journal of Operational Research*, Vol. 23, No. 3, pp.257–267.

Minkowski, H. (1910) *Geometrie der Zahlen*, Teubner, Leipzig.

Padberg, M. and Sung, T-Y. (1991) 'An analytical comparison of different formulations of the travelling salesman problem', *Mathematical Programming*, Vol. 52, No. 2, pp.315–357.

Picard, J.C. and Queyranne, M. (1978) 'The time-dependent traveling salesman problem and its application to the tardiness in one-machine scheduling', *Operations Research*, Vol. 26, No. 1, pp.86–110.

Rockafellar, R.T. (1997) *Convex Analysis*, Princeton University Press, Princeton, New Jersey, USA.

Swart, E.R. (1986, 1987 revision) $P = NP$, Technical report, University of Guelph, Canada.

Weyl, H. (1935) 'Elementare Theorie der konvexen Polyheder', *Commentarii Math. Helvetici*, Vol. 7, pp.290–306.

Yannakakis, M. (1991) 'Expressing combinatorial optimization problems by linear programming', *Journal of Computer and System Sciences*, Vol. 43, No. 3, pp.441–466.