QUBO formulation for the Snake-in-the-box and Coil-in-the-box problems

Federico Fuidio
Facultad de Ingeniería
Universidad de Montevideo
Montevideo, Uruguay
ffuidio@correo.um.edu.uy

Eduardo Canale
Facultad de Ingeniería
Universidad de la República
Montevideo, Uruguay
canale@fing.edu.uy

Rafael Sotelo
Facultad de Ingeniería
Universidad de Montevideo
Montevideo, Uruguay
rsotelo@um.edu.uy

Abstract—**This paper present the first QUBO formulations for the Snake-in-the-box (SITB) and Coil-in-the-box (CITB) problems. Both formulations are also capable of solving the NP-Hard problems of Maximum induced path and Maximum induced cylce respectively. In the process we also found a new QUBO formulation for the Maximum Common Induced Sub-graph problem. We proved the correctness of our formulations for the SITB, CITB and Maximum Common Sub-graph problem, and tested the formulations of the SITB and CITB in both classical and quantum solvers, being able to get the best solution for up to 5 dimensions.**

Index Terms—**QUBO, Quantum Annealer, Snake in the box, Coil in the box, Maximum common induced sub-graph problem**

I. INTRODUCTION

A. Snake-In-The-Box and Coil-In-The-Box problems

Both the Snake-in-the-box problem (SITB problem for short) and the Coil-in-the-box problem (CITB problem for short) were first introduced by Kautz in 1958 [1].

The SITB problem involves finding the longest induced path in the n-dimensional hypercube Q_n. (See section II for definitions of induced graphs and Q_n). In other words, it entails discovering the longest path that is isomorphic to an induced sub-graph of Q_n. Similarly, the CITB problem involves finding the longest induced cycle in the hypercube graph Q_n.

Note that in both cases, by *longest* we mean the graph with the maximum number of edges. Since in a path the number of edges is one less than the number of vertices, and in a cycle both numbers are the same, both the SITB and CITB problems are equivalent if we aim to maximize the number of vertices instead than the number of edges. Therefore, in all the formulations presented here, we are trying to maximize the number of vertices. We call n-snake and n-coil to an induced path and cycle of Q_n respectively. We often follow the notation introduced in [5] and call an n-snake and n-coil simply as a snake and a coil respectively.

Both problems have a wide range of applications, first, solutions of these problems can be used for error-detection codes [1]. Also, they can be used for scalar quantizers in digital communication systems [11], rank modulation [14] or genetics [12].

Usual methods to solve these problems involve brute force approaches [1], genetic algorithms [10] and Monte-Carlo Tree search [13]. In table I, we present the best known values for snake and coil lengths for dimensions 1 to 13. Notice that the length refer to the number of edges. Also, for dimentions 1 to 8, it is proved that the values shown are indeed the best ones.

Table I
BEST KNWON VALUES FOR SNAKE AND COIL IN THE BOX

Dimension	SITB	CITB
1	**1**	**0**
2	**2** [2]	**4** [1]
3	**4** [2]	**6** [1]
4	**7** [2]	**8** [1]
5	**13** [2]	**14** [1]
6	**26** [2]	**26** [2]
7	**50** [10]	**48** [8]
8	**98** [20]	**96** [18]
9	190 [15]	188 [15]
10	370 [13]	366 [21]
11	712 [21]	692 [21]
12	1373 [21]	1344 [21]
13	2687 [21]	2594 [21]

Values in bold type are proven to be the best

B. Quantum Annealer and QUBO formualtion

QUBO (Quadratic Unconstrained Binary Optimization) [23] is a mathematical model for optimization problems. It entails finding the minumum value of a binary quadratic form (i.e. finding the minimum value of a function of the form $\sum_{i,j} \alpha_{i,j} x_i x_j$ for $x_i \in \{0,1\}$, $\alpha_{i,j} \in \mathbb{R}$).

Quantum annealing [9] represents a potential advantage in tackling the problem at hand compared to traditional algorithms. It is a method aimed at finding the global minimum of an objective function, drawing inspiration from the process of annealing in metallurgy where a material is heated and then slowly cooled to remove defects and attain a more stable structure.

D-Wave Systems [29] is a notable producer of commercial quantum annealers. These devices are specifically designed for quantum annealing and have found application in various research and commercial projects.

C. Quantum Annealers and NP-Hard problems

NP-hard problems [4] are distinguished by their substantial complexity, where the time required to find a solution increases exponentially with the size of the input. These problems are particularly challenging in computational theory due to their intractability using classical algorithms, especially as the problem size expands.

In classical computing, solving NP-hard problems typically involves heuristic or approximation methods, such as genetic algorithms [7] or Monte-Carlo Tree search [16], to find near-optimal solutions within a reasonable timeframe. Exact solutions generally require brute-force approaches that are computationally expensive and impractical for large instances.

Quantum computing introduces a new paradigm for addressing NP-hard problems, leveraging the principles of quantum mechanics to process information in fundamentally different ways compared to classical computers. One promising approach is quantum annealing.

Quantum annealers offer potential advantages over classical solvers by exploring a vast solution space more efficiently due to quantum superposition and tunneling effects [19]. Specifically, quantum annealers can escape local minima more effectively than classical methods, potentially finding better solutions for NP-hard problems.

D. Structure of the paper

The paper is organized as follows: Section II provides the necessary background in graph theory, introducing key concepts and definitions used throughout the paper. Section III introduces the QUBO formulations for the SITB and CITB problems, as well as formulations for the Induced Sub-graph problem and the Maximum Common Induced Sub-graph problem, which are foundational to the SITB and CITB formulations. Section IV presents the statements of the theorems related to the correctness of the QUBO formulations presented in section III. Section V discusses the results obtained from applying the QUBO formulations of the SITB and CITB problems using both classical and quantum solvers. Section VI concludes the paper, summarizing the findings and outlining future work, with a particular focus on exploring problem symmetries to reduce the number of variables in the QUBO formulations, aiming to solve larger instances of the problems. The appendix contains the detailed proofs of the theorems stated in Section IV, providing the mathematical rigor behind the correctness of the QUBO formulations.

II. BACKGROUND

In this section we introduce some key concepts and definitions used in this paper.

Induced sub-graph: Given a graph $G = (V, E)$, an induced subgraph by a subset of vertices $A \subseteq V$ is a subgraph of G whose vertex set is equal to A, and whose edges consist of all edges connecting two vertices in A. We denote this graph as $G[A]$.

Graph Isomorphism: We say that two graphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ are isomorphic (denoted as

$G_1 \simeq G_2$) if there exists a bijective function $\phi : V_1 \rightarrow V_2$ that preserves the graph structure, i.e., $\{u, v\} \in E_1$ if and only if $\{\phi(u), \phi(v)\} \in E_2$. If such a function ϕ exists, we say that ϕ is an isomorphism between G_1 and G_2.

n-dimensional hypercube: The hypercube Q_n can be defined as a graph whose vertex set is $V_{Q_n} = \{0, 1\}^n$ and edge set is $E_{Q_n} = \{\{u, v\} \mid u, v \in V_{Q_n}, d_H(u, v) = 1\}$, where $d_H(u, v)$ is the Hamming distance between u and v.

Path and Coil graphs: We define the graph P_n as a path of n vertices, meaning that $V_{P_n} = \{v_0, v_1, \ldots, v_{n-1}\}$ and $E_{P_n} = \{\{v_i, v_{i+1}\} \mid 0 \leqslant i < n - 1\}$. Similarly, the graph C_n is defined as a cycle of n vertices, so that $V_{C_n} = \{v_0, v_1, \ldots, v_{n-1}\}$ and $E_{C_n} = \{\{v_i, v_{(i+1 \bmod n)}\} \mid 0 \leqslant i \leqslant n - 1\}$.

III. FORMULATION

A. QUBO formulation for the Induced sub-graph problem

The Induced sub-graph problem [3] is a decision problem that involves determining if a graph $G_1 = (V_1, E_1)$ is an induced sub-graph of $G_2 = (V_2, E_2)$. Note that this problem is equivalent to determining if there exists an injective function $\phi : V_1 \rightarrow V_2$ such that $\{u, v\} \in E_1$ if and only if $\{\phi(u), \phi(v)\} \in E_2$. Calude et al. (2017) [22] provides a QUBO formulation for this problem.

In this work, we present a similar formulation inspired by the formulation of the Graph Isomorphism Problem given by Lucas (2014) [17]. The idea of the formulation involves defining binary variables to describe the function ϕ. Let $x_{u,i}$ for $u \in V_1$ and $i \in V_2$ be such that $x_{u,i} = 1$ only if the vertex u is mapped to i by the function ϕ. Define s_i for $i \in V_2$ such that $s_i = 1$ only if there exists a vertex $u \in V_1$ that is mapped to i by ϕ. We arrive at the following QUBO formulation:

$$Q = H_A + H_B$$

$$H_A = \sum_{u \in V_1} \left(1 - \sum_{i \in V_2} x_{u,i}\right)^2 + \sum_{i \in V_2} \left(s_i - \sum_{u \in V_1} x_{u,i}\right)^2$$

$$H_B = \sum_{uv \in E_1} \sum_{ij \notin E_2} x_{u,i} x_{v,j} + \sum_{uv \notin E_1} \sum_{ij \in E_2} x_{u,i} x_{v,j}$$

This formulation needs a total of $|V_1||V_2| + |V_2|$ binary variables.

We define ϕ as the relationship $\{(u, i) \in V_1 \times V_2 \mid x_{u,i} = 1\}$.

Note that if $H_A = 0$, then ϕ defines an injective function from V_1 to V_2, and if $H_B = 0$, the function preserves the graph structure. Since $H_A \geqslant 0$ and $H_B \geqslant 0$, $\min(Q) \geqslant 0$, implying that if the minimum value of Q is zero, then $H_A = H_B = 0$, and G_1 is an induced sub-graph of G_2.

B. QUBO formulation for the Maximum Common Induced sub-graph problem

The maximum common induced sub-graph problem [3] consists of finding, given two graphs G_1 and G_2, a graph with the maximum number of vertices that is an induced sub-graph of both G_1 and G_2. Equivalently, we aim to find

the induced sub-graph of G_1 with the greatest order that is isomorphic to an induced sub-graph of G_2. [24] introduces the first QUBO formulation for this problem. In this paper, we provide a different formulation for this problem, that can be naturally extended to solve the SITB and CITB problems.

We modify the formulation of the induced sub-graph problem by adding variables p_u to indicate if a vertex u in V_1 is mapped to a vertex in V_2. Let $x_{u,i}$ and s_i be defined as in the previous formulation. Then, the QUBO formulation is as follows:

$$Q = \alpha H_A + \beta H_B + \gamma H_O \tag{1}$$

$$H_O = -\sum_{u \in V_1} p_u \tag{2}$$

$$H_A = \sum_{u \in V_1} \left(p_u - \sum_{i \in V_2} x_{u,i} \right)^2 + \sum_{i \in V_2} \left(s_i - \sum_{u \in V_1} x_{u,i} \right)^2 \tag{3}$$

$$H_B = \sum_{uv \in E_1} \sum_{ij \notin E_2} x_{u,i} x_{v,j} + \sum_{uv \notin E_1} \sum_{ij \in E_2} x_{u,i} x_{v,j} \tag{4}$$

Where $\alpha, \beta, \gamma > 0$. Also, we found that if $\alpha, \beta > \gamma$ then the minimum value of Q is achieved only if we found a solution for the Maximum Common Induced Sub-graph problem.

This formulation uses $|V_1||V_2| + |V_1| + |V_2|$ binary variables.

As in the previous formulation, we define ϕ as the relationship $\{(u,i) \in V_1 \times V_2 \mid x_{u,i} = 1\}$. Also, we define two sets $A_x \subseteq V_1$ and $B_x \subseteq V_2$ as:

$$A_x = \{u \in V_1 \mid x_{u,i} = 1 \text{ for some } i \in V_2\} \tag{5}$$

$$B_x = \{i \in V_2 \mid x_{u,i} = 1 \text{ for some } u \in V_1\} \tag{6}$$

In this formulation, $H_A = 0$ ensures that ϕ is an injective function from the set A_x to V_2, and $H_B = 0$ ensures that the function preserves the graph structure. H_O represents the objective function of the formulation, which aims to maximize the number of vertices in the induced sub-graph.

If we choose values for α, β, γ such that $\alpha, \beta > \gamma$, then the minimum value of Q is achieved only if $G_1[A_x] \simeq G_2[B_x]$. We will prove this in the appendix.

C. QUBO formulation for the SITB problem

Using the formulation of the Maximum Common Induced sub-graph problem, we can solve both the SITB and CITB problems. The SITB problem can be viewed as an instance of the maximum common induced sub-graph problem, where $G_1 = P_{2^n}$ (the path of 2^n vertices) and $G_2 = Q_n$ (the n-dimensional hypercube graph). With an additional restriction to ensure that the graph $G_1[A_x]$ (the selected induced sub-graph of P_n) is connected.

In order to achieve this, we introduce a new term, H_C, to the previous formulation. The QUBO function Q for the SITB problem is then defined as follows:

$$Q = \alpha H_A + \beta H_B + \gamma H_O + \delta H_C \tag{7}$$

$$H_O = -\sum_{u \in V_1} p_u \tag{8}$$

$$H_A = \sum_{u \in V_1} \left(p_u - \sum_{i \in V_2} x_{u,i} \right)^2 + \sum_{i \in V_2} \left(s_i - \sum_{u \in V_1} x_{u,i} \right)^2 \tag{9}$$

$$H_B = \sum_{uv \in E_1} \sum_{ij \notin E_2} x_{u,i} x_{v,j} + \sum_{uv \notin E_1} \sum_{ij \in E_2} x_{u,i} x_{v,j} \tag{10}$$

$$H_C = (1 - p_{v_0})^2 + \sum_{uv \in E_1} (p_u - p_v)^2 \tag{11}$$

Where $\alpha, \beta, \gamma, \delta > 0$, $\alpha, \beta > \gamma + 2\delta$ and $\delta > 2^n \gamma$.

Since $|V_1| = |V_2| = 2^n$ this formulation uses $|V_1||V_2| + |V_1| + |V_2| = 2^{2n} + 2^{n+1}$ binary variables.

Notice that if H_C equals 1, then the graph $G_1[A_x]$ is connected and contains the vertex v_0, and $H_C = 0$ if and only if all the variables p_u are equal to one.

D. QUBO formulation for the CITB problem

The formulation for the Coil-in-the-box problem is not as straightforward. In the SITB problem, as we showed earlier, we can essentially solve the maximum common induced sub-graph problem by setting the largest path that can fit in the n-dimensional hypercube graph (a path with 2^n vertices) as G_1, setting Q_n as G_2 and introducing an additional term to ensure that the chosen sub-graph of P_{2^n} remains connected.

This approach solves the SITB problem for two reasons. First, any connected sub-graph of a path is also a path. Second, in a path, the term H_C can ensure that the selected sub-graph of the path remains connected. However, it's important to note that in other types of graphs, such as general graphs or graphs with more complex structures, the H_C term may not guarantee contentedness.

Then, in order to define the CITB problem, we cannot simply take G_1 as the largest possible cycle that can fit in the Hypercube graph (a cycle with 2^n vertices) and use the same formulation of the SITB. However we can slightly modify the SITB formulation in the following way:

First, define $G_1 = (V_1, E_1)$ where $V_1 = V_{\text{path}} \cup V_{\text{cycle}}$ and $E_1 = E_{\text{path}} \cup E_{\text{cycle}}$, and:

$$V_{\text{path}} = \{1, 2, \ldots, 2^n - 1\}$$

$$V_{\text{cycles}} = \{(1, 3), (1, 4), \ldots, (1, 2^n - 1)\}$$

$$E_{\text{path}} = \{\{j, j+1\} \mid 1 \leqslant j \leqslant 2^n - 2\}$$

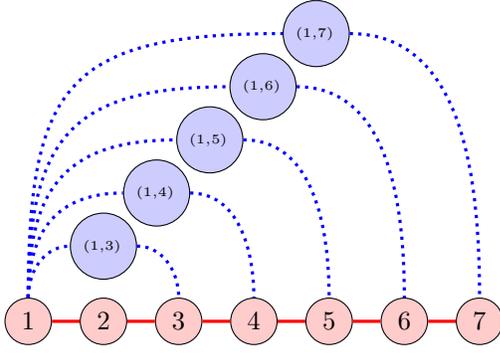$$E_{\text{cycles}} = \{\{i, v\} \mid i \in \{1, n\}, v \in V_{\text{cycles}}\}$$

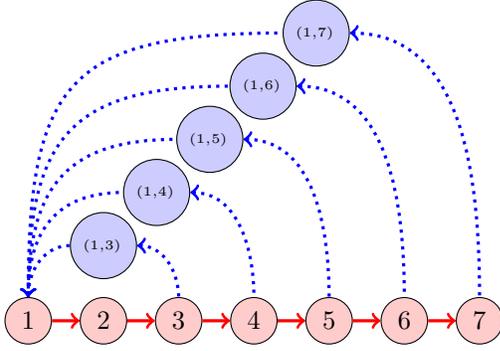Figure 1. Representation of the graph G_1 for $n = 3$

Figure 2. Representation of the directed graph $\overrightarrow{G_1}$ for $n = 3$

In Figure 1, we present a representation of the graph G_1 for $n = 3$. Vertices and edges belonging to V_{cycle} and E_{cycle} are drawn in dotted blue, while vertices and edges belonging to V_{path} and E_{path} are shown in red.

From G_1, we define an oriented graph $\overrightarrow{G_1}$, where the vertex set is V_1, and we assign a direction to each edge as shown in Figure 2. We refer to the edge set of the oriented graph as $\overrightarrow{E_1}$. Additionally, we denote the edge $(u, v) \in \overrightarrow{E_1}$ (with direction from u to v) as $u \rightarrow v$.

Finally, we define G_2 as the n-dimensional hypercube graph Q_n, and we define the binary variables $x_{u,i}$, p_u, and s_i in the same manner as in our previous formulations. With these definitions, we arrive at the following QUBO formulation:

$$Q = \alpha H_A + \beta H_B + \gamma H_O + \delta H_C + \epsilon H_R \quad (12)$$

$$H_O = -\sum_{u \in E_1} p_u \quad (13)$$

$$H_A = \sum_{u \in V_1} \left(p_u - \sum_{i \in V_2} x_{u,i} \right)^2 + \sum_{i \in V_2} \left(s_i - \sum_{u \in V_1} x_{u,i} \right)^2 \quad (14)$$

$$H_B = \sum_{uv \in E_1} \sum_{ij \notin E_2} x_{u,i} x_{v,j} + \sum_{uv \notin E_1} \sum_{ij \in E_2} x_{u,i} x_{v,j} \quad (15)$$

$$H_C = \left(1 - \sum_{u \in V_{cycle}} p_u \right)^2 \quad (16)$$

$$H_R = \sum_{u \in V_{path}} \left(p_u - \sum_{v: \, u \rightarrow v \in \overrightarrow{E_1}} p_v \right)^2 \quad (17)$$

Where $\alpha, \beta, \gamma, \delta, \epsilon > 0$, $\alpha, \beta > \gamma + \delta + \epsilon$ and $\delta, \epsilon > 2^n \gamma$.

In this formulation $|V_1| = 2^{n+1} - 4$ and $|V_2| = 2^n$. Then the formulation uses $|V_1||V_2| + |V_1| + |V_2| = 2^{2n+1} - 2^n - 4$ binary variables.

Here, the term H_C, if null, ensures that exactly one vertex in V_{cycle} is selected, representing the use of a single vertex from the cycle component. Whereas the term H_R, if null, ensures that for every selected vertex u in V_{path}, there is exactly one corresponding vertex $v \in V_1$ such that $u \rightarrow v$ is in $\overrightarrow{E_1}$. Therefore, if both H_C and H_R are zero, the induced sub-graph $G_1[A_x]$ forms a cycle.

E. Longest induced path and Longest induced cycle problems

The longest induced path problem involves finding the longest induced path in an arbitrary graph G. Similarly, the longest induced cycle problem aims to find the longest induced cycle in an arbitrary graph G. Both of these problems are proven to be NP-hard [6].

Note that the SITB and CITB problems are just special cases of these problems when the graph G is the hypercube Q_n. Thus, the formulations presented for the SITB and CITB problems can handle the more general NP-hard problems of the longest induced path and longest induced cycle in a graph, respectively.

Therefore, to solve the longest induced path (respectively, longest induced cycle) problem in a graph $G = (V, E)$, we can use the same formulation employed for the CITB (respectively, SITB) problem, where we set $G_2 = G$ instead of $G_2 = Q_n$. This demonstrates the versatility and broader applicability of our formulations in addressing these complex combinatorial problems.

IV. Correctness of the QUBO formulations

In this section we state four theorems regarding the correctness of the formulations presented in this paper. The first two theorems (IV.1 and IV.2) proves the correctness of the Maximum common induced sub-graph problem, under the assumption that $\alpha, \beta, \gamma > 0$ and $\alpha, \beta > \gamma$. These results are important since the formulation of both the SITB and CITB problems are based on our formulation for the Maximum common induced sub-graph problem.

Then we state two theorems (IV.3 and IV.4) that proves the correctness of the QUBO formulations for the SITB and CITB problem respectively. The four theorems are proven in the Appendix.

A. Correctness of the QUBO formulation of the Maximum Common Induced Sub-graph Problem

Theorem IV.1. *Let $Q(x, p, s)$ be the QUBO defined in (1) to (4) for the Maximum Common Induced Sub-graph problem, where x, p, s are vectors containing the variables $x_{u,i}$, p_u and s_i respectively.*

Then, if $\alpha, \beta > \gamma$ and x, p, s are such that $Q(x, p, s)$ is the minimum value of Q then the relationship ϕ defined by x is an injective function that preserves the graph structure.

Theorem IV.2. *(Correctness of the QUBO formulation for the Maximum Common Sub-graph Problem). Let $Q(x, p, s)$ be the QUBO defined in (1) to (4), where x, p, s are vectors containing the variables $x_{u,i}$, p_u and s_i respectively.*

Then, if $\alpha, \beta > \gamma$ and x, p, s are such that $Q(x, p, s)$ is the minimum value of Q then the relationship ϕ defined by x is the solution of the Maximum Commom sub-graph problem.

B. Correctness of the QUBO formulations for the SITB and CITB problems

Theorem IV.3. *(Correctness of the QUBO formulation for the SITB Problem). Let $Q(x, p, s)$ be the QUBO defined in (7) to (11) for the SITB problem, where x, p, s are vectors containing the variables $x_{u,i}$, p_u and s_i respectively.*

Then, if $\alpha, \beta > \gamma + 2\delta, \delta > 2^n \gamma$ and x, p, s are such that $Q(x, p, s)$ is the minimum value of Q then the relationship ϕ defined by x is an injective function that preserves the graph structure, and the subgraph $P_{2^n}[A_x]$ is a path.

Theorem IV.4. *(Correctness of the QUBO formulation for the CITB Problem). Let $Q(x, p, s)$ be the QUBO defined in (12) to (17) for the CITB problem, where x, p, s are vectors containing the variables $x_{u,i}$, p_u and s_i respectively.*

Then, if $\alpha, \beta > \gamma + \delta + \epsilon, \delta, \epsilon > 2^n \gamma$ and x, p, s are such that $Q(x, p, s)$ is the minimum value of Q then the relationship ϕ defined by x is an injective function that preserves the graph structure, and the subgraph $G_1[A_x]$ is a cycle.

V. RESULTS

We tested both the SITB and CITB formulations in three different solvers. We use one classic solver and two QPU (Quantum Processing Unity) based solvers.

For the classic solver the SimulatedAnnealingSampler [28] provided by D-Wave was used to simulate the implementation of a thermal annealing process. This simulator runs locally and was used on a personal laptop equipped with a 12th Gen Intel(R) Core(TM) i7-1255U CPU.

For the QPU based solvers we use two solvers that make use of the QPU. We used a Quantum-classical hybrid solver provided by D-Wave that uses both classical and quantum resources [27].

For the purely QPU-based solution, we used the D-Wave Advantage_system6.3 quantum computer, with 5614 qubits. We configure the number of shots to 1000, leaving the rest of the parameters unchanged. See [25] (D-Wave documentation) for explanations for all parameters of the QPU.

We accessed the QPU based solvers (both the hybrid and the purely quantum solver) using the *sapi* interface [26].

A. Snake-in-the-box results

Here we present the best solution obtained in the three solvers for the SITB problem. In both, the hybrid solver and the simulated annealer, we were able to get a 5-snake of length 13 in Q_5. This is, we were able to get the best solution for dimension 5. For the QPU solver we only were able to get the best solution for dimension 3.

Hybrid solution for $n = 5$ (note that it has 14 vertices, therefore the length of the snake is 13):

$$(10100) \rightarrow (00100) \rightarrow (00101) \rightarrow (01101) \rightarrow$$
$$(01111) \rightarrow (11111) \rightarrow (10111) \rightarrow (10011) \rightarrow$$
$$(00011) \rightarrow (00010) \rightarrow (01010) \rightarrow (11010) \rightarrow$$
$$(11000) \rightarrow (11001)$$

Simulated solution for $n = 5$:

$$(11001) \rightarrow (10001) \rightarrow (10000) \rightarrow (10100) \rightarrow$$
$$(00100) \rightarrow (01100) \rightarrow (01000) \rightarrow (01010) \rightarrow$$
$$(01011) \rightarrow (00011) \rightarrow (00111) \rightarrow (10111) \rightarrow$$
$$(11111) \rightarrow (11110)$$

QPU solution for $n = 3$:

$$(010) \rightarrow (000) \rightarrow (100) \rightarrow (101) \rightarrow (111)$$

B. Coil-in-the-box results

For the CITB problem we were also able to get the best solution up to dimension 5 using the Simulated Annealing and the Hybrid Solver. For the QPU we only were able to get the best solution for dimension 2.

Hybrid solution for $n = 5$:

$$(00101) \rightarrow (00111) \rightarrow (10111) \rightarrow (11111) \rightarrow$$
$$(11101) \rightarrow (11100) \rightarrow (10100) \rightarrow (10000) \rightarrow$$
$$(10010) \rightarrow (11010) \rightarrow (01010) \rightarrow (01011) \rightarrow$$
$$(01001) \rightarrow (00001)$$

Simulated solution for $n = 5$:

$$(10111) \rightarrow (10110) \rightarrow (00110) \rightarrow (00100) \rightarrow$$
$$(01100) \rightarrow (01000) \rightarrow (11000) \rightarrow (10000) \rightarrow$$
$$(10001) \rightarrow (00001) \rightarrow (00011) \rightarrow (01011) \rightarrow$$
$$(11011) \rightarrow (11111)$$

QPU solution for $n = 2$:

$$(11) \rightarrow (01) \rightarrow (00) \rightarrow (10)$$

VI. CONCLUSIONS AND FUTURE WORK

In this study, we introduced the first QUBO formulations for the Snake-in-the-box (SITB) and Coil-in-the-box (CITB) problems. Our formulations not only address these classic problems but also extend to solving the NP-hard problems of Maximum Induced Path and Maximum Induced Cycle, respectively. Through testing on both classical and quantum solvers, we demonstrated the efficacy of our approach by obtaining optimal solutions for dimensions up to 5.

The results obtained are promising, especially considering the complexity of the problems and the relatively early stage of quantum computing technology. The ability of our formulations to leverage quantum annealers for problem-solving hints

at the potential for significant advancements in combinatorial optimization problems as quantum technology matures.

For future work, we aim to further optimize our QUBO formulations by exploiting the symmetries inherent in the SITB and CITB problems. By reducing the number of variables in the formulations, we hope to increase the scalability of our approach and enable the solving of these problems for higher dimensions on quantum annealers. This would not only advance the state-of-the-art for these specific problems but also contribute to the broader field of quantum optimization and its applications. Additionally, we plan to test our formulations using a wider variety of solvers and apply them to solve the more general problems of the longest induced path and longest induced cycle.

REFERENCES

[1] W. H. Kautz, "Unit-Distance Error-Checking Codes," *IRE Transactions on Electronic Computers*, vol. EC-7, no. 2, pp. 179-180, June 1958. doi: 10.1109/TEC.1958.5222529.

[2] D. W. Davies, "Longest 'Separated' Paths and Loops in an N Cube," *IEEE Transactions on Electronic Computers*, vol. EC-14, no. 2, pp. 261-261, 1965. doi: 10.1109/pgec.1965.264259.

[3] H. G. Barrow and R. M. Burstall, "Subgraph isomorphism, matching relational structures and maximal cliques," *Inform. Process. Lett.*, vol. 4, no. 4, pp. 83–84, 1976.

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

[5] F. Harary, J. Hayes, and P. Wu, "A Survey of the Theory of Hypercube Graphs," *Computers & Mathematics with Applications*, vol. 15, pp. 277-289, Dec. 1988. doi: 10.1016/0898-1221(88)90213-1.

[6] M. R. Garey and D. S. Johnson, "Computers and Intractability: A Guide to the Theory of NP-Completeness,". W. H. Freeman & Co., 1990.

[7] Holland, J.H. (1992). *Adaptation in Natural and Artificial Systems*. MIT Press.

[8] K. J. Kochut, "Snake-in-the-box codes for dimension 7," *J. Combin. Math. Combin. Comput.*, vol. 20, pp. 175–185, 1996.

[9] T. Kadowaki and H. Nishimori, "Quantum annealing in the transverse Ising model," *Physical Review E*, vol. 58, no. 5, p. 5355, 1998.

[10] W. D. Potter, R. W. Robinson, J. A. Miller, K. J. Kochut, and DZ Redys, "Using the Genetic Algorithm to Find Snake-in-the-Box Codes," in *IEA/AIE*, 1994, pp. 421–426.

[11] S. Kim and D. L. Neuhoff, "Snake-in-the-box codes as robust quantizer index assignments," in *Proceedings of the IEEE International Symposium on Information Theory*, p. 402, 2000. doi: 10.1109/ISIT.2000.866700.

[12] I. Zinovik, Y. Chebiryak, and D. Kroening, "Periodic Orbits and Equilibria in Glass Models for Gene Regulatory Networks," *IEEE Transactions on Information Theory*, vol. 56, no. 2, pp. 805–820, 2010. doi: 10.1109/tit.2009.2037078.

[13] D. Kinny, "A new approach to the Snake-In-The-Box problem," *Frontiers in Artificial Intelligence and Applications*, vol. 242, pp. 462-467, Jan. 2012. doi: 10.3233/978-1-61499-098-7-462.

[14] Y. Yehezkeally and M. Schwartz, "Snake-in-the-Box Codes for Rank Modulation," *IEEE Transactions on Information Theory*, vol. 58, no. 8, pp. 5471-5483, Aug. 2012. doi: 10.1109/TIT.2012.2196755.

[15] E. Wynn, "Constructing circuit codes by permuting initial sequences," 2012, arXiv:1201.1647 [math.CO].

[16] Browne, C.B., Powley, E., Whitehouse, D., Lucas, S.M., Cowling, P.I., Rohlfshagen, P., ... & Colton, S. (2012). *A survey of Monte Carlo Tree Search methods*. IEEE Transactions on Computational Intelligence and AI in Games, 4(1), 1-43.

[17] A. Lucas, "Ising formulations of many NP problems," *Frontiers in Physics*, vol. 2, 2014. [Online]. Available: http://dx.doi.org/10.3389/fphy.2014.00005.

[18] P. R. J. Östergård and V. H. Pettersson, "On the maximum length of coil-in-the-box codes in dimension 8," *Discrete Applied Mathematics*, vol. 179, pp. 193–200, 2014.

[19] McGeoch, C. (2014). *Adiabatic quantum computation and quantum annealing: Theory and practice*. Morgan & Claypool Publishers.

[20] P. R. J. Östergård and V. H. Pettersson, "Exhaustive Search for Snake-in-the-Box Codes," *Graphs and Combinatorics*, vol. 31, pp. 1019–1028, 2015. doi: 10.1007/s00373-014-1423-3.

[21] David Allison and Daniël Paulusma, "New Bounds for the Snake-in-the-Box Problem," *ArXiv*, 2016. Available online: https://api.semanticscholar.org/CorpusID:5066905.

[22] C. S. Calude, M. J. Dinneen, and R. Hua, "QUBO formulations for the graph isomorphism problem and related problems," *Theoretical Computer Science*, vol. 701, pp. 54-69, 2017. doi: 10.1016/j.tcs.2017.04.016.

[23] A. P. Punnen, Ed., "The Quadratic Unconstrained Binary Optimization Problem: Theory, Algorithms, and Applications," Cham: Springer International Publishing, 2022. doi: 10.1007/978-3-031-04520-2.

[24] N. Huang and D. Roje, "A New QUBO Objective Function for Solving the Maximum Common Subgraph Isomorphism Problem Via Quantum Annealing," *SN Computer Science*, vol. 2, p. 186, 2021. doi: 10.1007/s42979-020-00431-5.

[25] D-Wave Systems Inc., "The D-Wave Advantage System: An Overview," https://www.dwavesys.com/resources/white-paper/the-d-wave-advantage-system-an-ove (Accessed Apr. 11, 2024)

[26] D-Wave Systems Inc., "Configuring Access to Leap's Solvers," https://docs.ocean.dwavesys.com/en/stable/overview/sapi.html (Accessed Dec. 7, 2023).

[27] D-Wave Systems Inc., "D-Wave Hybrid Release 0.6.10," https://docs.ocean.dwavesys.com/en/latest/docs_hybrid/sdk_index.html (Accessed Apr. 10, 2024).

[28] D-Wave Systems Inc., "Simulated Annealing Sampler," https://docs.ocean.dwavesys.com/en/latest/docs_neal/reference/sampler.html (Accessed Apr. 10, 2024).

[29] D-Wave Systems Inc., "Welcome to D-Wave — D-Wave System Documentation," https://docs.dwavesys.com/docs/latest/c_gs_1.html (Accessed Apr. 2, 2024).

VII. APPENDIX

A. Proof of theorem IV.1

Theorem IV.1: Let $Q(x, p, s)$ be the QUBO defined in (1) for the Maximum Induced Sub-graph problem, where x, p, s are vectors containing the variables $x_{u,i}$, p_u and s_i respectively.

Then, if α, $\beta > \gamma$ and x, p, s are such that $Q(x, p, s)$ is the minimum value of Q then the relationship ϕ defined by x is an injective function that preserves the graph structure.

Proof. We proceed by contradiction. First, suppose that the minimum value of Q is achieved by x, p, s, and $\phi = \{(u, i) \in V_1 \times V_2 \mid x_{u,i} = 1\}$ is either not a function from A_x to V_2 (*case 1*), not injective (*case 2*), or does not preserve the graph structure (*case 3*). Then, we find $\hat{x}$, $\hat{p}$, $\hat{s}$ with a strictly lower value of Q, leading to a contradiction.

Case 1: The relationship ϕ is not a function from A_x to V_2. This implies that there exists a vertex $u \in V_1$ that is mapped to more than one vertex in V_2, meaning that there exist $i, j \in V_2$ with $i \neq j$ such that $x_{u,i} = x_{u,j} = 1$. We can suppose that $p_u = s_i = s_j = 1$. This is because if any of these three variables were equal to 0, then we would immediately obtain a new solution with the corresponding variable equal to 1 and a strictly lower value of Q. Then, we have:

$$\sum_{k \in V_2} x_{u,k} \geq x_{u,i} + x_{u,j} = 2 \tag{18}$$

Using (18) along with the fact that $p_u = 1$ we get:

$$p_u - \sum_{k \in V_2} x_{u,k} = 1 - \sum_{k \in V_2} x_{u,k} \leq -1 \tag{19}$$

We can generate a new solution $\hat{x}, \hat{p}, \hat{s}$ with $\hat{x}_{u,j} = 0$, $\hat{s}_j = 0$, leaving all the other variables unchanged. We claim that this solution satisfy $Q(\hat{x}, \hat{p}, \hat{s}) < Q(x, p, s)$.

$$\begin{aligned}
\sum_{k \in V_2} \hat{x}_{u,k} &= \hat{x}_{u,j} + \sum_{k \in V_2,\, k \neq j} \hat{x}_{u,k} \\
&= (x_{u,j} - 1) + \sum_{k \in V_2,\, k \neq j} x_{u,k} \\
&= \sum_{k \in V_2} x_{u,k} - 1
\end{aligned}$$

Using now (19), along with the fact that $p_u = \hat{p_u}$:

$$\hat{p}_u - \sum_{k \in V_2} \hat{x}_{u,k} = p_u - \sum_{k \in V_2} x_{u,k} + 1 \leq 0$$

By taking absolute value on both sides we get:

$$\left| \hat{p}_u - \sum_{k \in V_2} \hat{x}_{u,k} \right| = \left| p_u - \sum_{k \in V_2} x_{u,k} \right| - 1$$

$$\left(\hat{p}_u - \sum_{k \in V_2} \hat{x}_{u,k} \right)^2 < \left(p_u - \sum_{k \in V_2} x_{u,k} \right)^2 \tag{20}$$

The rest of the terms of H_A (defined in (3)) remains unchanged. The term H_O also remains unchanged, since $p_u = \hat{p}_u$. Finally, H_B may change, but clearly it does not increase (since we are changing a variable that is set to 1 by 0). Therefore:

$$\begin{aligned}
H_A(\hat{x}, \hat{p}, \hat{s}) &< H_A(x, p, s) \\
H_B(\hat{x}, \hat{p}, \hat{s}) &\leq H_B(x, p, s) \\
H_O(\hat{x}, \hat{p}, \hat{s}) &= H_O(x, p, s)
\end{aligned}$$

Therefore $Q(\hat{x}, \hat{p}, \hat{s}) < Q(x, p, s)$, contradicting the minimality of $Q(x, p, s)$.

case 2: The map ϕ is not injective, This implies that there exists a vertex $i \in V_2$ such that two (or more) vertices of V_1 are mapped to it. Meaning that there exists $i \in V_2$ and $u, v \in V_1$, $u \neq v$ such that $x_{u,i} = x_{v,i} = 1$. As in the first case, we can suppose that $s_i = p_u = p_v = 1$.

We take the solution $\hat{x}, \hat{p}, \hat{s}$ with $\hat{x}_{v,i} = 0$, $\hat{p}_v = 0$ and all the other values unchanged. We have then:

$$\begin{aligned}
\beta H_B(\hat{x}, \hat{p}, \hat{s}) &\leq \beta H_B(x, p, s) \\
\gamma H_O(\hat{x}, \hat{p}, \hat{s}) &= \gamma H_O(x, p, s) + \gamma
\end{aligned}$$

With a similar analysis of the term H_A that the one in case 1, we get:

$$\left(\hat{s}_i - \sum_{w \in V_2} \hat{x}_{w,i} \right)^2 < \left(s_i - \sum_{w \in V_2} x_{w,i} \right)^2 \tag{21}$$

Since we are working with natural numbers, we can subtract one from the second term changing the strict inequality for a non-strict one.

$$\left(\hat{s}_i - \sum_{w \in V_2} \hat{x}_{w,i} \right)^2 \leq \left(s_i - \sum_{w \in V_2} x_{w,i} \right)^2 - 1$$

All the other terms in H_A remains unchanged. After multiplying by α we have:

$$\alpha H_A(\hat{x}, \hat{p}, \hat{s}) \leq \alpha H_A(x, p, s) - \alpha$$

Then:

$$Q(\hat{x}, \hat{p}, \hat{s}) \leq Q(x, p, s) + \gamma - \alpha$$

Since $\alpha > \gamma$, $Q(\hat{x}, \hat{p}, \hat{s}) < Q(x, p, s)$, contradicting the minimality of the solution.

Now, using what we proved in cases 1 and 2, we can assume that if x, p, s gives the minimum value of Q, then ϕ is an injective function (implying that $H_A = 0$). This will be useful in the next case.

Case 3: The function ϕ does not preserves the graph structure. Then, we can find $u, v \in V_1$, $i, j \in V_2$ such that $x_{u,i} = x_{v,j} = 1$, $\{u, v\} \in E_1$ but $\{i, j\} \notin E_2$ (The case when

$\{u, v\} \notin E_1$ and $\{i, j\} \in E_2$ is analogous). Notice that this implies that $p_u = p_v = s_i = s_j = 1$.

Like we mentioned earlier, we can assume that ϕ is an injective function. In this case, we consider the solution $\hat{x}, \hat{p}, \hat{s}$ with $\hat{x}_{u,i} = 0$, $\hat{p}_u = 0$, and $\hat{s}_i = 0$. If ϕ is an injective function, then the map $\hat{\phi} = \{(u, i) \in V_1 \times V_2 \mid \hat{x}_{u,i} = 1\}$ is also an injective function, implying that H_A remains equal to zero. Then:

$$\alpha H_A(\hat{x}, \hat{p}, \hat{s}) = \alpha H_A(x, p, s) = 0$$
$$\beta H_B(\hat{x}, \hat{p}, \hat{s}) \leq \beta H_B(x, p, s) - \beta$$
$$\gamma H_O(\hat{x}, \hat{p}, \hat{s}) = \gamma H_O(x, p, s) + \gamma$$

Implying that:

$$Q(\hat{x}, \hat{p}, \hat{s}) \leq Q(s, p, s) + \gamma - \beta$$

Since $\beta > \gamma$, $Q(\hat{x}, \hat{p}, \hat{s}) < Q(x, p, s)$, contradicting the minimality of x, p, s.

Thus, if the minimum value of Q is achieved by the binary variables x, p, s the set ϕ is an structure preserving injective function.

□

B. Proof of theorem IV.2

Theorem IV.2: (Correctness of the QUBO formulation for the Maximum Common Sub-graph Problem). Let $Q(x, p, s)$ be the QUBO defined in (1), where x, p, s are vectors containing the variables $x_{u,i}$, p_u and s_i respectively.

Then, if $\alpha, \beta > \gamma$ and x, p, s are such that $Q(x, p, s)$ is the minimum value of Q then the relationship ϕ defined by x is the solution of the Maximum Commom sub-graph problem.

Proof. Since $Q(x, p, s)$ is the minimum value of Q, using theorem IV.1 we know that ϕ is an injective function from A_x to V_2 that preserves the graph structure, meaning that $G_1[A_x] \simeq G_2[B_x]$. Then, we have $H_A(x, p, s) = H_B(x, p, s) = 0$ and $\gamma H_O(x, p, s) = -\gamma|A_x|$, therefore:

$$Q(x, p, s) = -\gamma|A_x| \tag{22}$$

By contradiction, suppose that there exists $A \subseteq V_1$, $B \subseteq V_2$ such that $G_1[A] \simeq G_2[B]$ with $|A| > |A_x|$.

Since $G_1[A] \simeq G_2[B]$, we know that there exists a bijective function $\psi : A \to B$ that preserves the graph structure. Define the binary variables $\hat{x}_{u,i}$, $\hat{p}_u$, $\hat{s}_i$ for $u \in V_1$, $i \in V_2$ in the following way:

$$\hat{x}_{u,i} = \begin{cases} 1 & \text{if } u \in A \text{ and } \psi(u) = i \\ 0 & \text{otherwise} \end{cases}$$

$$\hat{p}_u = \begin{cases} 1 & \text{if } u \in A \\ 0 & \text{otherwise} \end{cases}$$

$$\hat{s}_i = \begin{cases} 1 & \text{if } i \in B \\ 0 & \text{otherwise} \end{cases}$$

The map $\{(u, i) \in V_1 \times V_2 \mid \hat{x}_{u,i} = 1\}$ is exactly ψ. This means that the map induced by the binary variables $\hat{x}_{u,i}$, $\hat{p}_u$ and $\hat{s}_i$ is an structure preserving injective function. Therefore $H_A(\hat{x}, \hat{p}, \hat{s}) = H_B(\hat{x}, \hat{p}, \hat{s}) = 0$ and $\gamma H_O(\hat{x}, \hat{p}, \hat{s}) = -\gamma|A|$, and:

$$Q(\hat{x}, \hat{p}, \hat{s}) = -\gamma|A|$$

This is a contradiction, since $Q(\hat{x}, \hat{p}, \hat{s}) < Q(x, p, s)$.

□

C. Proof of theorem of SITB

Theorem IV.3: (Correctness of the QUBO formulation for the SITB Problem). Let $Q(x, p, s)$ be the QUBO defined in (7) to (11) for the SITB problem, where x, p, s are vectors containing the variables $x_{u,i}$, p_u and s_i respectively.

Then, if $\alpha, \beta > \gamma + 2\delta$, $\delta > 2^n \gamma$ and x, p, s are such that $Q(x, p, s)$ is the minimum value of Q then the relationship ϕ defined by x is an injective function that preserves the graph structure, and the subgraph $P_{2^n}[A_x]$ is a path.

Proof. We can proceed exactly as in the proof of Theorem IV.1. First, we suppose that $Q(x, p, s)$ is the minimum value of Q and ϕ is not a correct solution of the SITB problem, and then we find a new solution $\hat{x}, \hat{p}, \hat{s}$ with $Q(\hat{x}, \hat{p}, \hat{s}) < Q(x, p, s)$, leading to a contradiction. In this proof, we have 4 cases:

Case 1: If the relationship ϕ is not a function, there exists a vertex $u \in V_1$ that is mapped to more than one vertex in V_2, meaning that there exist $i, j \in V_2$ with $i \neq j$ such that $x_{u,i} = x_{u,j} = 1$. Again, we can suppose that $p_u = s_i = s_j = 1$. With the same change of variables as the one

$$\alpha H_A(\hat{x}, \hat{p}, \hat{s}) < \alpha H_A(x, p, s)$$
$$\beta H_B(\hat{x}, \hat{p}, \hat{s}) \leq \beta H_B(x, p, s)$$
$$\gamma H_O(\hat{x}, \hat{p}, \hat{s}) = \gamma H_O(x, p, s)$$
$$\delta H_C(\hat{x}, \hat{p}, \hat{s}) = \delta H_C(x, p, s)$$

Implying that $Q(\hat{x}, \hat{p}, \hat{s}) < Q(x, p, s)$, contradicting the minimality of $Q(x, p, s)$.

Case 2: ϕ is not injective. Again, we can do the same change as the one performed in the *Case 2* of Theorem IV.1. Notice that in that after changing the value of only one p_u then H_C can change on at most 2. Therefore we obtain:

$$\alpha H_A(\hat{x}, \hat{p}, \hat{s}) \leq \alpha H_A(x, p, s) - \alpha$$
$$\beta H_B(\hat{x}, \hat{p}, \hat{s}) \leq \beta H_B(x, p, s)$$
$$\gamma H_O(\hat{x}, \hat{p}, \hat{s}) = \gamma H_O(x, p, s) + \gamma$$
$$\delta H_C(\hat{x}, \hat{p}, \hat{s}) \leq \delta H_C(x, p, s) + 2\delta$$

Then:

$$Q(\hat{x}, \hat{p}, \hat{s}) \leq Q(x, p, s) + 2\delta + \gamma - \alpha$$

Since $\alpha > \gamma + 2\delta$ we have that $Q(\hat{x}, \hat{p}, \hat{s}) < Q(x, p, s)$, leading to a contradiction.

Case 3: The function ϕ does not preserves the graph structure. Then we can perform the same change as in *Case 3* of Theorem $IV.1$. In this case the effect on H_C is the same, therefore we obtain:

$$H_A(\hat{x}, \hat{p}, \hat{s}) = H_A(x, p, s) = 0$$

$$\beta H_B(\hat{x}, \hat{p}, \hat{s}) \leqslant \beta H_B(x, p, s) - \beta$$

$$\gamma H_O(\hat{x}, \hat{p}, \hat{s}) = \gamma H_O(x, p, s) + \gamma$$

$$\delta H_C(\hat{x}, \hat{p}, \hat{s}) \leqslant \delta H_C(x, p, s) + 2\delta$$

Then:

$$Q(\hat{x}, \hat{p}, \hat{s}) \leqslant Q(x, p, s) + 2\delta + \gamma - \beta$$

Since $\beta > \gamma + 2\delta$ we have that $Q(\hat{x}, \hat{p}, \hat{s}) < Q(x, p, s)$, leading to a contradiction.

Case 4: The selected subgraph of P_{2^n} is not connected. Since $H_A(x, p, s) = H_B(x, p, s) = 0$, we have that $H_C(x, p, s) \geqslant 1$ (this is because $H_C(x, p, s) = 0$ only if all the 2^n vertices are used in the induced path, but we know that for $n \geqslant 2$ there is no induced path of 2^n vertices)

Since $G_1[A_x]$ is not connected then $H_C(x, p, s) > 1$. We can see that the trivial solution $\hat{x}_{u,i} = \hat{p}_u = \hat{s}_i = 0 \ \forall u, i$ (all variables equal to zero) gives a lower value of Q:

$$Q(\hat{x}, \hat{p}, \hat{s}) = \delta$$

$$Q(x, p, s) \geqslant -|A_x|\gamma + 2\delta > -2^n\gamma + 2\delta$$

Therefore:

$$Q(x, p, s) \geqslant -2^n\gamma + Q(\hat{x}, \hat{p}, \hat{s}) + \delta$$

Since $\delta > 2^n\gamma$ we get that $Q(\hat{x}, \hat{p}, \hat{s}) < 0$, leading to a contradiction. $\qquad\square$

D. Proof of theorem CITB

Theorem IV.4: (Correctness of the QUBO formulation for the CITB Problem). Let $Q(x, p, s)$ be the QUBO defined in (12) to (17) for the CITB problem, where x, p, s are vectors containing the variables $x_{u,i}$, p_u and s_i respectively.

Then, if α, $\beta > \gamma + \delta + \epsilon$, $\delta, \epsilon > 2^n\gamma$ and x, p, s are such that $Q(x, p, s)$ is the minimum value of Q then the map ϕ defined by x is an injective function that preserves the graph structure, and the subgraph $G_1[A_x]$ cycle.

Proof. We proceed as in the proof of theorem $IV.1$. First, suppose that the minimum value of Q is achieved by x, p, s, and $\phi = \{(u, i) \in V_1 \times V_2 \,|\, x_{u,i} = 1\}$ is not a correct solution of the CITB problem, then, we find a solution $\hat{x}, \hat{p}, \hat{s}$ such that $Q(\hat{x}, \hat{p}, \hat{s}) < Q(x, p, s)$, leading to a contradiction. We have now 4 cases:

Case 1: The relationship ϕ is not a function from A_x to V_2. This implies that there exists a vertex $u \in V_1$ that is mapped to more than one vertex in V_2, meaning that there exist $i, j \in V_2$ with $i \neq j$ such that $x_{u,i} = x_{u,j} = 1$. Again, we can suppose that $p_u = s_i = s_j = 1$. Then we have:

$$\alpha H_A(\hat{x}, \hat{p}, \hat{s}) < \alpha H_A(x, p, s)$$

$$\beta H_B(\hat{x}, \hat{p}, \hat{s}) \leqslant \beta H_B(x, p, s)$$

$$\gamma H_O(\hat{x}, \hat{p}, \hat{s}) = \gamma H_O(x, p, s)$$

$$\delta H_C(\hat{x}, \hat{p}, \hat{s}) = \delta H_C(x, p, s)$$

$$\epsilon H_R(\hat{x}, \hat{p}, \hat{s}) \leqslant \epsilon H_R(x, p, s)$$

Then $Q(\hat{x}, \hat{p}, \hat{s}) < Q(x, p, s)$ contradicting the minimality of $Q(x, p, s)$.

case 2: The map ϕ is not injective, This implies that there exists a vertex $i \in V_2$ such that two (or more) vertices of V_1 are mapped to it. Meaning that there exists $i \in V_2$ and $u, v \in V_1$, $u \neq v$ such that $x_{u,i} = x_{v,i} = 1$. As in the first case, we can suppose that $s_i = p_u = p_v = 1$.

We take the solution $\hat{x}, \hat{p}, \hat{s}$ with $\hat{x}_{v,i} = 0$, $\hat{p}_v = 0$ and all the other values unchanged. We have then:

$$\alpha H_A(\hat{x}, \hat{p}, \hat{s}) \leqslant \alpha H_A(x, p, s) - \alpha$$

$$\beta H_B(\hat{x}, \hat{p}, \hat{s}) \leqslant \beta H_B(x, p, s)$$

$$\gamma H_O(\hat{x}, \hat{p}, \hat{s}) = \gamma H_O(x, p, s) + \gamma$$

$$\delta H_C(\hat{x}, \hat{p}, \hat{s}) \leqslant \delta H_C(x, p, s) + \delta$$

$$\epsilon H_R(\hat{x}, \hat{p}, \hat{s}) \leqslant \epsilon H_R(x, p, s) + \epsilon$$

Then we have:

$$Q(\hat{x}, \hat{p}, \hat{s}) \leqslant Q(x, p, s) + \epsilon + \delta + \gamma - \alpha$$

Since $\alpha > \epsilon + \delta + \gamma$ we get $Q(\hat{x}, \hat{p}, \hat{s}) < Q(x, p, s)$, contradicting the minimality of $Q(x, p, s)$.

Case 3: The function ϕ does not preserves the graph structure. Then, we can find $u, v \in V_1$, $i, j \in V_2$ such that $x_{u,i} = x_{v,j} = 1$, $\{u, v\} \in E_1$ but $\{i, j\} \notin E_2$ (The case when $\{u, v\} \notin E_1$ and $\{i, j\} \in E_2$ is analogous). Notice that this implies that $p_u = p_v = s_i = s_j = 1$.

Like we mentioned earlier, we can assume that ϕ is an injective function. In this case, we consider the solution $\hat{x}, \hat{p}, \hat{s}$ with $\hat{x}_{u,i} = 0$, $\hat{p}_u = 0$, and $\hat{s}_i = 0$. If ϕ is an injective function, then the map $\hat{\phi} = \{(u, i) \in V_1 \times V_2 \,|\, \hat{x}_{u,i} = 1\}$ is also an injective function, implying that H_A remains equal to zero. Then:

$$\alpha H_A(\hat{x}, \hat{p}, \hat{s}) = \alpha H_A(x, p, s) = 0$$

$$\beta H_B(\hat{x}, \hat{p}, \hat{s}) \leqslant \beta H_B(x, p, s) - \beta$$

$$\gamma H_O(\hat{x}, \hat{p}, \hat{s}) = \gamma H_O(x, p, s) + \gamma$$

$$\delta H_C(\hat{x}, \hat{p}, \hat{s}) \leqslant \delta H_C(x, p, s) + \delta$$

$$\epsilon H_R(\hat{x}, \hat{p}, \hat{s}) \leqslant \epsilon H_R(x, p, s) + \epsilon$$

Then:

$$Q(\hat{x}, \hat{p}, \hat{s}) \leqslant Q(x, p, s) + \epsilon + \delta + \gamma - \beta$$

Since $\beta > \epsilon + \delta + \gamma$ we get $Q(\hat{x}, \hat{p}, \hat{s}) < Q(x, p, s)$, contradicting the minimality of $Q(x, p, s)$.

Case 4: Finally, we prove that the selected induced subgraph $G_1[A_x]$ is a cycle. For that we prove that $H_C(x, p, s) = H_R(x, p, s) = 0$. By contradiction, suppose that $H_C(x, p, s) > 0$, then, taking the trivial solution $\hat{x}_{u,i} = \hat{p}_u = \hat{s}_i = 0\ \forall u, i$.

$$Q(\hat{x}, \hat{p}, \hat{s}) = 0$$

$$Q(x, p, s) \geqslant -|A_x|\gamma + \delta + \epsilon H_R(x, p, s) > -2^n \gamma + \delta + \epsilon H_R(x, p, s)$$

Therefore, since $H_R(x, p, s) \geqslant 0$ and $\delta > 2^n \gamma$ we get that $Q(\hat{x}, \hat{p}, \hat{s}) < Q(x, p, s)$, leading to a contradiction.

Similarly, if $H_R(x, p, s) > 0$, we can take again the trivial solution $\hat{x}, \hat{p}, \hat{s}$:

$$Q(\hat{x}, \hat{p}, \hat{s}) = 0$$

$$Q(x, p, s) \geqslant -|A_x|\gamma + \epsilon + \delta H_C(x, p, s) > -2^n \gamma + \epsilon + \delta H_C(x, p, s)$$

Therefore, since $H_C(x, p, s) \geqslant 0$ and $\epsilon > 2^n \gamma$ we get that $Q(\hat{x}, \hat{p}, \hat{s}) < Q(x, p, s)$, leading to a contradiction. $\square$