Q-FW: A Hybrid Classical-Quantum Frank-Wolfe for Quadratic Binary Optimization

Alp Yurtsever[1], Tolga Birdal[2], and Vladislav Golyanik[3]

[1] Umeå University, Sweden
[2] Imperial College London, UK
[3] MPI for Informatics, SIC, Germany

Abstract. We present a hybrid classical-quantum framework based on the Frank-Wolfe algorithm, Q-FW, for solving quadratic, linearly-constrained, binary optimization problems on quantum annealers (QA). The computational premise of quantum computers has cultivated the re-design of various existing vision problems into quantum-friendly forms. Experimental QA realisations can solve a particular non-convex problem known as the quadratic unconstrained binary optimization (QUBO). Yet a naive-QUBO cannot take into account the restrictions on the parameters. To introduce additional structure in the parameter space, researchers have crafted ad-hoc solutions incorporating (linear) constraints in the form of regularizers. However, this comes at the expense of a hyper-parameter, balancing the impact of regularization. To date, a true constrained solver of quadratic binary optimization (QBO) problems has lacked. Q-FW first reformulates constrained-QBO as a copositive program (CP), then employs Frank-Wolfe iterations to solve CP while satisfying linear (in)equality constraints. This procedure unrolls the original constrained-QBO into a set of unconstrained QUBOs all of which are solved, in a sequel, on a QA. We use D-Wave Advantage QA to conduct synthetic and real experiments on two important computer vision problems, graph matching and permutation synchronization, which demonstrate that our approach is effective in alleviating the need for an explicit regularization coefficient.

1 Introduction

Combinatorial optimization is at the heart of computer vision (CV). In a variety of applications such as structure-from-motion (SfM) [66], SLAM [56], 3D reconstruction [21], camera re-localization [65], image retrieval [51] and 3D scan stitching [40,24], correspondences serve as a powerful proxy to visual perception. In many problems, correspondences are defined over two or multiple point sets and can be encoded as *permutation matrices* that are binary assignment operators. Recovering permutations from observations involve solving NP-hard combinatorial problems. As a remedy, scholars have opted to *relax* those problems to arrive at tractable albeit suboptimal solutions [70,8,14,43]. However, recent advances in computer hardware urges us to re-visit our approaches.

Quantum computers (QCs) harness the collective properties of quantum states, such as superposition, interference and entanglement to perform calculations [60]. Thanks to the use of a more advanced physics, QCs can offer theoretical improvements in the face of complexity classes that are challenging to handle today [68]. With the experimental realization of quantum supremacy [4], we are now more confident that practical quantum computing is right around the corner, *i.e.* the numer of usable Qubits now reach 5000 (DWave Advantage) and are expected to exceed beyond 7000 (DWave Advantage 2) [69].

A particular quantum computational model, known as *Adiabatic Quantum Computing* (AQC), is based on the adiabatic theorem of quantum mechanics [12]. Closely related to it is *Quantum Annealing* (QA), which is a quantum optimization method (AQC-type) that implements a qubit-based quantum system described by the Ising model [42]. Albeit restricted, experimental realisations of QA, such as DWave [22], can solve non-convex, *quadratic unconstrained binary optimization* (QUBO) problems, without resorting to continuous relaxations. This premise of AQC and QA has led to the emergence of *quantum computer vision* (QCV), where researchers started to port existing computer vision problems into forms amenable to quantum computation [36,50,32,67,75,7].

Even though employing QA to solve CV problems has shown benefit[4], a large body of computer vision algorithms rely on some form of (in)equality constraints to be incorporated. For example, estimating correspondences require solving QBOs for permutations and not for arbitrary binary vectors. To this end, the state-of-the-art QCV methods either use a regularization with cherry-picked coefficients [7] or resort to heuristics for auto-controlling the impact of the constraints [67,75]. Unfortunately, none of these approaches are optimal and jeopardize the solution quality guarantees of quantum computers.

In this paper, we address the above issue of incorporating (in)equality constraints and introduce Quantum-Frank Wolfe (Q-FW), a Frank-Wolfe framework for satisfying linear (in)equality constraints in a QBO problem. Q-FW is based on an equivalent *copositive programming* formulation of constrained QBO and involves iteratively solving a sequence of classical, unconstrained QUBOs. At its core, on a classical computer, Q-FW employs one of the two variants of FW tailored for solving CP problems; FW with augmented Lagrangian (FWAL) [72] or FW with quadratic penalty (FWQP) [73]. At every iteration, these methods identify an update direction by minimizing a linear approximation of a penalized proxy of the objective function. Q-FW formulates this linear minimization as a QUBO and obtains the update direction via QA. We then take a small step in this update direction. In addition, FWAL maintains a dual variable, updated by a small gradient step for improved numerical performance.

Thanks to the convexity inherent in CP, Q-FW converges to the global minimum regardless the choice of its algorithm parameters. By virtue of the

[4] Quantum computers are still in early stages. However, a diverse set of CV experiments present optimistic predictions regarding the future.

exact copositive-reformulation, our solutions are oftentimes near the true global minimum, obtained via an exhaustive search in small problems. We deploy Q-FW on multiple computer vision tasks of permutation synchronization and graph matching, which both have wide applicability. Our contributions are:

- We introduce Q-FW, an adaptation of the classical FW algorithm for solving copositive programs on a hybrid classical-quantum computing system.
- We solve the challenging QUBO sub-problems using an actual experimental realisation of a quantum annealer (QA), DWave Advantage 4.1 [62,22].
- We tackle both graph matching and permutation synchronization problems and obtain excellent results on both synthetic and real benchmarks.

Our evaluations confirm the theoretical advantages of Q-FW: Q-FW is robust, can solve larger problems than brute-force search, can exactly satisfy (in)equality constraints and enjoys a tight copositive relaxation. Our MATLAB implementation as well as scripts required to run D-Wave are available under: github.com/QuantumComputerVision/QuantumFrankWolfe.

2 Related Work

Our approach relates to different methods both in classical optimization and quantum computer vision. In this section, we review the most related works in QCV, copositive programming and FW.

Quantum computer vision (QCV). QCV encompasses hybrid classical-quantum methods with parts solved on a gate-based quantum computer or a quantum annealer. This young field seeks to identify how challenging problems can be formulated for and benefit from quantum hardware. While it remained predominantly theoretical at early stages [58,17], QCV methods from various domains were evaluated on real quantum hardware during the recent few years, including image classification [57,59,16], object detection [50], graph matching [67], mesh alignment [5], robust fitting [25] and permutation synchronisation [7].

Some of the proposed algorithms require additional constraints formulated as weighted linear terms (Lagrange multipliers) [67,7,75]. Such conditions rectify the original unconstrained objective and preserve the QUBO form consumable by modern QA. However, since the linear constraints modify the problem's energy landscape, the corresponding weights have to be chosen with care; too high or too low weights can significantly decrease the probability of measuring optimal solutions after the sampling. Birdal *et al.* [7] select the weights with a time-consuming grid search (for small problem instances). Benkner *et al.* [67] derive lower bounds on the rectification weights for the quadratic assignment problem. Both policies have a common limitation: The determined weights are problem-specific and do not generalise to other problems. Moreover, even problems of the same type and size can demand new multipliers.

In contrast to existing methods, our unified policy does not require selecting the weights of linear terms in advance. Similar to Q-Match [5], our method is iterative: a sequence of optimisation tasks are solved on QPU; in each iteration, the control is returned to CPU to define a follow-up QUBO until convergence. Q-Match [5] update its solutions via a series of permutation-ness-preserving directions (collections of 2-cycles). Its policy does not generalise to other problems, arbitrary solution encodings and weighted linear constraints, as our method does.

Copositive programming (CP). CP is a subfield of convex optimization concerned with optimizing a linear objective under affine constraints over the cone of copositive matrices, or its dual cone, the cone of completely positive matrices. By definition, a matrix $\mathbf{X} \in \mathbb{R}^{n\times n}$ is said to be copositive if its quadratic form is nonnegative on the first orthant (*i.e.*, $\mathbf{z}^\top \mathbf{X}\mathbf{z} \geq 0$ for all $\mathbf{z} \in \mathbb{R}^n_+$) and completely positive if $\mathbf{X} \in \mathrm{conv}\{\mathbf{x}\mathbf{x}^\top : \mathbf{x} \in \mathrm{R}^n_+\}$. Compared to semidefinite programming, CP provides a tighter relaxation of quadratic problems [63]. However, despite its convexity, solving a CP problem is NP-Hard [10]. Several NP-Hard problems in quadratic and combinatorial optimization are subsets of CP, including the binary quadratic problems [13], problems of finding stability and chromatic numbers of a graph [23,26], quadratic assignment problem [61], and training of vector-output RELU networks [64]. We refer to the excellent surveys [27,28] and references therein for more details.

Frank Wolfe (FW). FW (also known as conditional gradient method or CGM) is a classical method in convex optimization dating back to 1956 [30]. Initially, the method is proposed for minimizing a convex quadratic loss function over a polytope. The analysis is extended in [49] to minimize a generic smooth and convex objective over an arbitrary convex and compact set. The eccentric feature of FW is that it does not require a projection step, which is in stark contrast with most other methods for constrained optimization, and it makes FW efficacious for problems where projection is computationally prohibitive. FW is demonstrated as an effective method for optimization over simplex [19] or spactrahedron domains [35]. We refer to [41] for convergence analysis of FW and a detailed discussion on its applications, and to [11] for a review on recent advances in FW.

The original form of FW is not suitable to tackle affine equality constraints present in our CP formulation. Instead, we consider two design variants of FW: FWQP [73], which equips FW with a quadratic penalty strategy for affine constraints; and FWAL [72], which extends FWQP for an augmented Lagrangian penalty. Our choice is inspired by [74] using FWAL for solving semidefinite programs. We adopt a similar approach for solving CPs.

In what follows, we first formulate QBO as an instance of the more general copositive program class §3. We then provide our Q-FW framework for solving copositive programs in a generic way (§4). Finally, we cast graph matching (§5.1) and permutation synchronization (§5.2) tasks as instances of QBOs with equality constraints, which Q-FW could solve effectively.

3 Problem Formulation

This section presents our model problem, a quadratic binary optimization (QBO) with affine (in)equality constraints[5], and an equivalent copositive program outlined in [13].

We assume that the problems are presented in the following form:

$$\min_{\mathbf{x}\in\mathbb{Z}_2^n} \quad \mathbf{x}^\top\mathbf{Q}\mathbf{x} + 2\,\mathbf{s}^\top\mathbf{x} \quad \text{subject to} \quad \mathbf{a}_i^\top\mathbf{x} = b_i, \quad i = 1, 2, \ldots, m, \tag{1}$$

where $\mathbf{x} \in \mathbb{Z}_2^n$ is the binary valued decision variable, $\mathbf{Q} \in \mathbb{R}^{n\times n}$ and $\mathbf{s} \in \mathbb{R}^n$ are the quadratic and linear cost coefficients, and $\{(\mathbf{a}_i, b_i) \in \mathbb{R}^n \times \mathbb{R}\}$ are the constraint coefficients. We assume $b_i \geq 0$ without loss of generality. Throughout, we treat $\mathbf{Q}$ as a symmetric matrix since (1) is invariant under symmetrization of $\mathbf{Q}$:

$$\mathbf{x}^\top\mathbf{Q}\mathbf{x} = \tfrac{1}{2}\mathbf{x}^\top\mathbf{Q}\mathbf{x} + \tfrac{1}{2}(\mathbf{x}^\top\mathbf{Q}\mathbf{x})^\top = \mathbf{x}^\top(\tfrac{1}{2}\mathbf{Q} + \tfrac{1}{2}\mathbf{Q}^\top)\mathbf{x}. \tag{2}$$

One can also drop the linear term $\mathbf{s}^\top\mathbf{x}$ from the objective, because we can translate it into the quadratic term: Given that $\mathbf{x}$ is binary valued, $\mathbf{s}^\top\mathbf{x} = \mathbf{x}^\top\mathrm{Diag}(\mathbf{s})\,\mathbf{x}$.

To reformulate this problem, consider the rank-one completely positive matrix $\mathbf{X} = \mathbf{x}\mathbf{x}^\top \in \mathbb{Z}_2^{n\times n}$. Since $\mathbf{x}$ is binary valued, we have $\mathrm{diag}(\mathbf{X}) = \mathbf{x}$. Then, the quadratic objective in (1) can be cast as a linear function of $\mathbf{X}$:

$$\mathbf{x}^\top\mathbf{Q}\mathbf{x} = \mathrm{Tr}(\mathbf{x}^\top\mathbf{Q}\mathbf{x}) = \mathrm{Tr}(\mathbf{Q}\,\mathbf{x}\mathbf{x}^\top) = \mathrm{Tr}(\mathbf{Q}\mathbf{X}). \tag{3}$$

Similarly, we rewrite affine constraints from problem (1) by using

$$\begin{aligned} \mathbf{a}_i^\top\mathbf{x} = b_i \;&\Longleftrightarrow\; (\mathbf{a}_i^\top\mathbf{x})^2 = b_i^2 \\ &\Longleftrightarrow\; \mathrm{Tr}(\mathbf{A}_i\mathbf{X}) = b_i^2, \quad \text{where} \quad \mathbf{A}_i := \mathbf{a}_i\mathbf{a}_i^\top, \end{aligned} \tag{4}$$

which holds true since $(\mathbf{a}_i^\top\mathbf{x})^2 = \mathbf{x}^\top\mathbf{a}_i\mathbf{a}_i^\top\mathbf{x} = \mathrm{Tr}(\mathbf{x}^\top\mathbf{a}_i\mathbf{a}_i^\top\mathbf{x}) = \mathrm{Tr}(\mathbf{a}_i\mathbf{a}_i^\top\mathbf{x}\mathbf{x}^\top)$.

Now, we reformulate problem (1) as follows:

$$\begin{aligned} \min_{\mathbf{x},\mathbf{X}} \quad \mathrm{Tr}(\mathbf{Q}\mathbf{X}) + 2\,\mathbf{s}^\top\mathbf{x} \quad \text{subject to} \quad & \mathbf{a}_i^\top\mathbf{x} = b_i, \quad i = 1, 2, \ldots, m, \\ & \mathrm{Tr}(\mathbf{A}_i\mathbf{X}) = b_i^2, \quad i = 1, 2, \ldots, m, \\ & \mathbf{X} = \mathbf{x}\mathbf{x}^\top, \quad \text{and} \quad \mathbf{x} \in \mathbb{Z}_2^n. \end{aligned} \tag{5}$$

By replacing the nonconvex nonlinear constraint $\{\mathbf{X} = \mathbf{x}\mathbf{x}^\top, \mathbf{x} \in \mathbb{Z}_2^n\}$ with

$$\mathrm{diag}(\mathbf{X}) = \mathbf{x}, \quad \text{and} \quad \begin{bmatrix} 1 & \mathbf{x}^\top \\ \mathbf{x} & \mathbf{X} \end{bmatrix} \in \Delta^{n+1} \text{ where } \Delta^n := \mathrm{conv}\{\mathbf{x}\mathbf{x}^\top : \mathbf{x} \in \mathbb{Z}_2^n\}, \tag{6}$$

[5] Throughout the paper we concentrate on the equality constraints and provide a simple modification to satisfy inequality constraints in our supplementary material.

we get a CP problem:

$$\begin{aligned} \min_{\mathbf{x},\mathbf{X}} \ \operatorname{Tr}(\mathbf{Q}\mathbf{X}) \quad \text{subject to} \quad & \mathbf{a}_i^\top \mathbf{x} = b_i, \quad i = 1, 2, \dots, m, \\ & \operatorname{Tr}(\mathbf{A}_i \mathbf{X}) = b_i^2, \quad i = 1, 2, \dots, m, \\ & \operatorname{diag}(\mathbf{X}) = \mathbf{x}, \text{ and } \begin{bmatrix} 1 & \mathbf{x}^\top \\ \mathbf{x} & \mathbf{X} \end{bmatrix} \in \mathcal{\Delta}^{n+1}. \end{aligned} \tag{7}$$

This reformulation is tight, see Theorem 2.6 in [13] for the technical derivation. Our numerical experiments demonstrate the tightness of this reformulation empirically for the graph matching and permutation synchronization problems.

Compact notation. We introduce a compact notation for problem (7) for convenience. Let $p = n + 1$, denote the new decision variable by $\mathbf{W} \in \mathcal{\Delta}^p$, and introduce a new cost matrix $\mathbf{C} = \left[\begin{smallmatrix} 0 & \mathbf{s}^\top \\ \mathbf{s} & \mathbf{Q} \end{smallmatrix}\right]$. Further, let $d = 2m + n + 1$ and introduce a linear map $\mathcal{A} : \mathbb{R}^{p \times p} \to \mathbb{R}^d$ and vector $\mathbf{v} \in \mathbb{R}^d$ combining all affine constraints in problem (7), including $\{\mathbf{a}_i^\top \mathbf{x} = b_i\}$, $\{\operatorname{Tr}(\mathbf{A}_i \mathbf{X}) = b_i^2\}$, $\operatorname{diag}(\mathbf{X}) = \mathbf{x}$, and $W_{1,1} = 1$.

In this notation, problem (7) becomes

$$\min_{\mathbf{W} \in \mathcal{\Delta}^p} \ \operatorname{Tr}(\mathbf{C}\mathbf{W}) \quad \text{subject to} \quad \mathcal{A}\mathbf{W} = \mathbf{v}. \tag{8}$$

This is a convex optimization problem, but it is NP-Hard because of the complete positivity constraint.

4 Quantum Frank-Wolfe (Q-FW)

In the light of the copositive reformulation above, we now develop the main algorithm for solving a constrained-QBO. We describe the algorithm with FWAL. FWQP is covered as a special case by removing the dual steps of FWAL.

First, we construct the augmented Lagrangian of problem (8) by introducing a dual variable $\mathbf{y} \in \mathbb{R}^d$ and a penalty parameter $\beta > 0$:

$$L_\beta(\mathbf{W}; \mathbf{y}) = \operatorname{Tr}(\mathbf{C}\mathbf{W}) + \mathbf{y}^\top (\mathcal{A}\mathbf{W} - \mathbf{v}) + \frac{\beta}{2} \|\mathcal{A}\mathbf{W} - \mathbf{v}\|^2 \quad \text{for } \mathbf{W} \in \mathcal{\Delta}^p. \tag{9}$$

The goal is to minimize $L_\beta(\mathbf{W}; \mathbf{y})$ with respect to the primal variable $\mathbf{W}$ and maximize with respect to the dual variable $\mathbf{y}$:

$$\min_{\mathbf{W} \in \mathcal{\Delta}^p} \ \max_{\mathbf{y} \in \mathbb{R}^d} \ \operatorname{Tr}(\mathbf{C}\mathbf{W}) + \mathbf{y}^\top (\mathcal{A}\mathbf{W} - \mathbf{v}) + \frac{\beta}{2} \|\mathcal{A}\mathbf{W} - \mathbf{v}\|^2. \tag{10}$$

Note, the inner maximization gives an indicator function for $\mathcal{A}\mathbf{W} = \mathbf{v}$:

$$\max_{y \in \mathbb{R}^m} \ \mathbf{y}^\top (\mathcal{A}\mathbf{W} - \mathbf{v}) = \begin{cases} 0 & \text{if } \mathcal{A}\mathbf{W} = \mathbf{v} \\ +\infty & \text{otherwise} \end{cases} \tag{11}$$

Hence, the saddle point problem (10) is equivalent to our model problem (8).

The FWAL iteration employs a simple optimization strategy with two main steps, performed on the augmented Lagrangian loss function $L_\beta(\mathbf{W};\mathbf{y})$:
(1) A primal step to update $\mathbf{W}$, inspired by the FW algorithm, and
(2) A dual gradient ascent step to update $\mathbf{y}$.
The penalty parameter, β, is increased at a specific rate to ensure convergence of $\mathbf{W}$ to a feasible solution. Next, we describe the algorithm steps in detail.

Initialization. Choose an initial penalty parameter $\beta_0 > 0$, and initial primal and dual estimates $\mathbf{W}_0 \in \Delta^p$ and $\mathbf{y}_0 \in \mathbb{R}^d$. In practice, we let $\beta_0 = 1$, and we choose $\mathbf{W}$ and $\mathbf{y}$ as the matrix/vector of zeros.

At iteration $t = 1, 2, \ldots$, we increase the penalty parameter $\beta_t = \beta_0\sqrt{t+1}$ and perform the following updates:

Primal step. For primal step, we fix the dual variable $\mathbf{y}_t$ and take a FW step on the primal variable $\mathbf{W}_t$ with respect to the augmented Lagrangian loss (9). First, we compute the partial derivative of L_{β_t} with respect to $\mathbf{W}$:

$$\mathbf{G}_t = \mathbf{C} + \mathcal{A}^\top \mathbf{y}_t + \beta_t \mathcal{A}^\top(\mathcal{A}\mathbf{W}_t - \mathbf{v}). \tag{12}$$

Then, we find an update direction $\mathbf{H}_t \in \Delta^p$ by minimizing the first-order Taylor expansion of L_{β_t}:

$$\mathbf{H}_t \in \operatorname*{arg\,min}_{\mathbf{W}\in\Delta^p} L_{\beta_t}(\mathbf{W}_t;\mathbf{y}_t) + \operatorname{Tr}(\mathbf{G}_t(\mathbf{W}-\mathbf{W}_t)) \equiv \operatorname*{arg\,min}_{\mathbf{W}\in\Delta^p} \operatorname{Tr}(\mathbf{G}_t\mathbf{W}). \tag{13}$$

This step can be written as a standard, unconstrained QUBO. Specifically,

$$\text{if} \quad \mathbf{w}_t \in \operatorname*{arg\,min}_{\mathbf{w}\in\mathbb{Z}_2^p} \mathbf{w}^\top \mathbf{G}_t \mathbf{w}, \quad \text{then} \quad \mathbf{H}_t := \mathbf{w}_t\mathbf{w}_t^\top \in \operatorname*{arg\,min}_{\mathbf{W}\in\Delta^p} \operatorname{Tr}(\mathbf{G}_t\mathbf{W}). \tag{14}$$

Therefore, we can **implement and solve this step effectively on an AQC**. This is a key observation for our framework.

Finally, we update the primal variable $\mathbf{W}_t$ by taking a step towards $\mathbf{H}_t$:

$$\mathbf{W}_{t+1} = (1-\eta_t)\mathbf{W}_t + \eta_t\mathbf{H}_t, \quad \text{with step-size } \eta_t = \frac{2}{t+1}. \tag{15}$$

Dual step. For dual step, we fix $\mathbf{W}_{t+1}$ and take a gradient ascent step on the dual variable with respect to the augmented Lagrangian loss (9). The partial derivative of L_{β_t} with respect to $\mathbf{y}$ is

$$\mathbf{g}_t = \mathcal{A}\mathbf{W}_{t+1} - \mathbf{v}. \tag{16}$$

Then we take a gradient step in this direction

$$\mathbf{y}_{t+1} = \mathbf{y}_t + \gamma_t\mathbf{g}_t, \quad \text{with step-size } \gamma_t \geq 0. \tag{17}$$

There are two different strategies for the dual step-size γ_t, for more details we refer to Section 3.1 in [72]. In practice, we choose a constant step-size $\gamma_t = \beta_0$.

This completes one FWAL iteration. The following proposition, a simple adaptation from [72, Theorem 3.1], establishes the convergence rate of FWAL for our model problem (8).

Proposition 1. *Consider FWAL for problem (8). Choose an initial penalty parameter $\beta_0 > 0$. Assume that the solution set is nonempty, strong duality holds*[6]*, and the effective dual domain is bounded (i.e., there exists $D < +\infty$ such that $\|y_t\| \leq D$ at every iteration). Then, the primal sequence $\mathbf{W}_t \in \Delta^p$ converges to a solution $\mathbf{W}_\star$ with the following bounds on the error:*

$$\mathrm{Tr}(\mathbf{C}\mathbf{W}_t) - \mathrm{Tr}(\mathbf{C}\mathbf{W}_\star) \leq \frac{1}{\sqrt{t}}\left(6\beta_0 p^2 \|\mathcal{A}\|^2 + \frac{D^2}{2\beta_0}\right) \quad \textit{(objective suboptimality)} \tag{18}$$

$$\|\mathcal{A}\mathbf{W}_t - \mathbf{v}\| \leq \frac{1}{\sqrt{t}}\left(2\sqrt{3}p\|\mathcal{A}\| + \frac{4D}{\beta_0}\right) \quad \textit{(infeasibility error)} \tag{19}$$

where $\|\mathcal{A}\| := \sup\{\|\mathcal{A}\mathbf{X}\| : \|\mathbf{X}\|_F \leq 1\}$ is the operator norm of $\mathcal{A}$.

Remark 1. We recover FWQP from FWAL by choosing $\mathbf{y}_0 = \mathbf{0}$ and $\gamma_t = 0$, in other words, by removing the dual steps. These two methods have similar guarantees with the same rate of convergence up to a constant factor, but FWAL is reported to perform better for most instances in practice [72].

Rounding. We can immediately extract a solution for the original QBO problem (1) from a solution $\mathbf{W}_\star$ of CP reformulation (8). However, in practice, with finite time and computation, we get only an approximate solution $\hat{\mathbf{W}}$. A naive estimate that we extract from $\hat{\mathbf{W}}$ can be infeasible for (1). To this end, we implement the following rounding procedure: First, we get $\hat{\mathbf{X}}$ by removing the first row and first column of $\hat{\mathbf{W}}$. Next, we compute the best rank-one approximation $\hat{\mathbf{x}}\hat{\mathbf{x}}^\top$ of $\hat{\mathbf{X}}$ with respect to the Frobenius norm.[7] Finally, as an optional step, we project $\hat{\mathbf{x}}$ onto the feasible set of (1). The set of permutation matrices is the feasible set in our numerical experiments. We use Hungarian algorithm [46] for projection.

Quantum Annealing (QA). QA converts a QUBO objective to the equivalent Ising problem that is then solved by a meta-heuristic governed by quantum fluctuations [29]. Since this analogue optimisation process is prone to different physical disturbances (e.g., state decoherence and cosmic radiation)—and is, hence, non-deterministic—multiple repetitions are required to obtain an optimal solution with high probability. Furthermore, the current experimental QA realisations do not easily allow defining high-level constraints; the latter must be integrated so that the QUBO structure is preserved. In practice, constraints are

[6] Strong duality is a standard assumption for primal-dual methods in optimization.

[7] This amounts to computing the top singular vector of $\hat{\mathbf{X}}$ [54].

formulated as weighted linear terms adjusting qubit couplings and biases [67,7]. Finding optimal weights (e.g., by a grid search) is a tedious procedure that does not guarantee the generalisation of the selected multipliers across the problems. We provide further details on quantum annealing in our supplementary material.

On computational complexity. The convergence of FW is sub-linear and hence may require significant number of iterations, *e.g.* 200-1000. At each iteration, Q-FW attempts to solve an NP-Hard QUBO problem whose computational complexity class is FP^{NP}-complete[8] [71]. Thanks to the exploitation of quantum phenomena, QA can bring a quadratic improvement reducing the theoretical complexity from $O(e^N)$ to $O(e^{\sqrt{N}})$, in a similar vein to Grover algorithm [1,34]. Though, it is not straightforward to get a problem-specific, realistic estimate of the time complexity of the QA process. Nevertheless, fixing a constant annealing time and a constant number of repetitions, as we do for our small problems, can lead to an optimistic, polynomial time algorithm [3].

5 Experimental Evaluation

The proposed approach (Q-FWAL) is general and not tailored towards a specific problem. Hence, we assess its validity in realizing quantum versions two different problems, *graph matching* and *permutation synchronization*, both requiring equality constraints to be accounted for.[9] We use problem-specific synthetic and real datasets to showcase the effectiveness of our approach.

Implementation details. In both of the experiments we use the DWave Advantage 4.1 system [53] which has at least 5,000 qubits and ∼35,000 couplers. Except the ablation studies, we use 50 or 250 annealing cycles of $20\mu s$ in each iteration with an annealing schedule of $100\mu s$ breaks. We set the chain strength ξ according to the *maximum chain strength criterion*: We inspect the minor embedding calculated by Cai *et al.* [15] and set $\xi = s_{\max} + \omega$, with $s_{\max}$ being the maximum chain length in the minor embedding and $\omega = 0.5$ is the strengthening weight. If we observe frequent chain breaks for larger problems, we increase ω to 3.0. We access DWave at each iteration through the *Leap2* API [20]. We investigate three modes of Q-FW: (i) with intermediary exhaustive solution instead of DWave (FWAL), (ii) without Hungarian rounding (Q-FWAL relaxed) and (iii) the full configuration (Q-FWAL). Note that vanilla FWAL (i) cannot be applied to large problems due to the combinatorial explosion. In all of our problems, we are interested in linear permutation constraints, as those are the most common in CV problems. Hence, we use *Hungarian algorithm* [46] as the projector onto the constraint set (*cf. Rounding* in §4) and formulate permutation-ness into linear constraints as in [7,67] (*cf.*supplementary material).

[8] A binary relation $P(x, y)$, is in FP^{NP} if and only if there is a deterministic polynomial time algorithm that can determine whether $P(x, y)$ holds given both x and y.

[9] While still providing a way to handle inequalities in our supplementary material, we leave it as a future work to study problems with inequality constraints.

Table 1: Evaluations of graph matching on random problem instances with different sizes [67]. We report mean normalized energies over ten instances (the lower the better). Last five columns correspond to the variants of our method.

N	[67] ins.	[67] row.	[6] DS*	[44] SA	FWAL	Q-FWAL relaxed (50)	Q-FWAL (50)	Q-FWAL relaxed (250)	Q-FWAL (250)
3	1.49	2.12	0.85	0.82	**7e-4**	1.72	0.093	1.72	**7e-4**
4	5.68	7.37	0.43	2.43	**1.3e-3**	3.41	1.82	0.23	**1.43e-3**

5.1 Quantum Graph Matching (QGM)

In general, 3D vision problems relate two abstract shape/image manifolds $\mathcal{M}_1$ and $\mathcal{M}_2$. In many applications, these manifolds can be sampled by two point clouds (e.g. keypoints) $\mathcal{X}_1 \in \mathbb{R}^{\mathtt{N}_1 \times \mathtt{n}}$ and $\mathcal{X}_2 \in \mathbb{R}^{\mathtt{N}_2 \times \mathtt{n}}$ where $\mathtt{n}$ is the dimensionality of the problem domain, *e.g.* two for images, three for meshes and etc. We further assume a distance function $\phi(.)$ defined over the points of these point clouds. The quadratic assignment problem (QAP) then takes the form:

$$\max_{\mathbf{\Pi}} \ \mathrm{vec}(\mathbf{\Pi})^\top \mathbf{Q}_{\mathrm{QGM}} \, \mathrm{vec}(\mathbf{\Pi}) \quad \text{subject to} \quad \mathbf{\Pi} \in \mathcal{P} \tag{20}$$

where $\mathcal{P}$ denotes the set of (partial) permutations and $\mathrm{vec}(\cdot)$ acts as a vectorizer. Assuming $\mathtt{N} := \mathtt{N}_1 = \mathtt{N}_2$, *i.e. total* permutations (TP), $\mathbf{Q}_{\mathrm{QGM}} \in \mathbb{R}^{\mathtt{N}^2 \times \mathtt{N}^2}$ denotes a *ground cost* matrix or the quadratic energy measuring the gain of matching $\mathcal{M}_1$ and $\mathcal{M}_2$ by a sub-permutation $\mathbf{\Pi}$, computed using the distance $\phi(\cdot)$.

Baselines & dataset. We benchmark QGM against the exhaustive solution, obtained by searching over all possible permutations, as well as against the first AQC approach which was proposed by Benkner *et al.* [67] who used multiple strategies (*e.g. inserted*, *row-wise*) to inject soft-permutation constraints into QUBO. This required tuning of a parameter $\lambda \in \mathbb{R}$, whose large values are found to cause problems [7,67,75]. As a heuristic, [67] suggested a *spectral-gap*[10] analysis to bound the regularization coefficient λ. We also include: (i) the result obtained by running *simulated annealing* (SA) [44] on a CPU (the implementation from the Ocean tools [20]); (ii) a state of the art classical graph matching algorithm [6].

To assess, we use two sets of ten random problem instances with $\mathtt{N} = 3$ and $\mathtt{N} = 4$ as in §5.1 of [67]. The ground-truth permutations are calculated by brute force and compared qualitatively with the expected outcomes on real data. The number or qubits in the minor embeddings equals to 14 ($\mathtt{N} = 3$) and 40 ($\mathtt{N} = 4$).

Results. We report the *mean normalized energies over ten instances* in Tab. 1. This quantity is obtained by first shifting all energies by the minimum energy (of the ground-truth solution) and then averaging them. Clearly, FWAL and Q-FWAL perform the best on this experiment. However, FWAL cannot be scaled

[10] the difference between the lowest and second-lowest energy state / eigen-value

to large problems, and as we will see later in §5.2, Q-FWAL is able to handle much larger problems thanks to the advances in AQCs. DS* is a powerful classical algorithm, yet it cannot match the errors we achieve. SA is good for small problems, but its solution quality quickly drops with the problem size. Finally, it is visible that 50 cycles might be insufficient to get high quality results. Note, in practical scenarios where the number of keypoints across two images/scenes are different (*e.g.* $N_1 \neq N_2$), we will have to adopt *partial* permutations as non-square matrices yielding *inequality* constraints. We provided how inequality constraints can be factored into our framework, theoretically (*cf.* supp. §1.5).

Fig. 1: Willow Dataset [18]. (**left**) Manual annotations of keypoints. (**right**) Ground truth multi-image matches.

5.2 Quantum Permutation Synchronization (QPS)

Many multi-shape/view/instance computer vision problems can be solved by synchronization, including shape (point set) alignment [31,37], structure from motion [9,33], multi-view matching [52,8] and motion segmentation [38,2].

A specific branch, *permutation synchronization* seeks to find globally consistent image/shape matches from a set of relative matches over a collection. In particular, consider a *collection* of $\mathtt{K}$ point sets $\mathcal{X}_1, \ldots, \mathcal{X}_{\mathtt{K}}$ [11] of $\mathtt{N}$ points each such that there exists a bijective map for each pair $(\mathcal{X}_i, \mathcal{X}_j)$. We assume the availability of a set of noisy *relative* permutations $\{\mathbf{P}_{ij} : \mathcal{X}_i \to \mathcal{X}_j\}_{ij}$ estimated in *isolation*, *i.e.* independently. Our goal is then to solve this multi-graph matching problem even when a significant fraction of the pairwise matches are incorrect. To this end, a large body works minimize a *cycle-consistency loss*, that is shown to be equivalent to a QUBO (*cf.* [7] for a proof):

$$\underset{\{\mathbf{X}_i \in \mathcal{P}_{\mathtt{n}}\}}{\arg\min} \sum_{(i,j)\in\mathcal{E}} \|\mathbf{P}_{ij} - \mathbf{X}_i \mathbf{X}_j^\top\|_{\mathrm{F}}^2 = \underset{\{\mathbf{X}_i \in \mathcal{P}_{\mathtt{n}}\}}{\arg\min}\, \mathbf{x}^\top \mathbf{Q}_{QPS} \mathbf{x}. \tag{21}$$

Here, $\mathbf{x} = [\cdots \mathbf{x}_i^\top \cdots]^\top$ and $\mathbf{x}_i = \mathrm{vec}(\mathbf{X}_i)$ depict the *canonical ordering* of points. The first AQC approach to this problem is proposed by Birdal *et al.* [7], who, similar to [67], regularize $\mathbf{Q}_{QPS}$ to incorporate permutation-ness as a soft constraint. Note that, this whole problem has a gauge freedom, where we can freely choose $\mathbf{X}_0$ *e.g.* , as an identity matrix (*cf.* supplementary material).

[11] Such sets are easy to obtain by keypoint detection or sampling either on images or on shapes, *e.g.* by detecting N landmarks per image, in a M-view image collection.

Table 2: Evaluations on Willow Dataset.

	Car	Duck	Motorbike	Winebottle	Average
MatchEIG [52]	0.81 ± 0.083	0.86 ± 0.102	0.77 ± 0.059	0.87 ± 0.107	0.83 ± 0.088
MatchALSS [76]	0.84 ± 0.095	0.90 ± 0.102	0.81 ± 0.078	0.94 ± 0.092	0.87 ± 0.092
MatchLIFT [39]	0.84 ± 0.102	0.90 ± 0.103	0.81 ± 0.078	0.94 ± 0.092	0.87 ± 0.094
MatchBirkhoff [8]	0.84 ± 0.094	0.90 ± 0.107	0.81 ± 0.079	0.94 ± 0.093	0.87 ± 0.093
QuantumSync [7]	0.84 ± 0.104	0.90 ± 0.104	0.81 ± 0.080	0.93 ± 0.095	0.87 ± 0.096
[7]-search	0.84 ± 0.104	0.91 ± 0.115	0.82 ± 0.10	0.95 ± 0.096	0.88 ± 0.104
Q-FWAL (ours)	**0.92 ± 0.094**	**0.97 ± 0.072**	**0.89 ± 0.093**	**0.99 ± 0.044**	**0.94 ± 0.076**

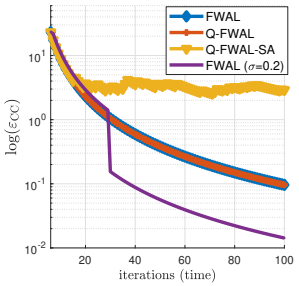
Datasets. As a real dataset, we follow [7] and use the kindly provided subset of the Willow Object Classes [18] composed of four categories (*duck*, *car*, *winebottle*, *motorbike*) with 40 RGB images each, acquired *in the wild* (*cf.* Fig. 1). This subset contains multiple sets of four points sampled out of ten annotations. This leads to 35 small problems per category each of which is a fully connected graph of all four consecutive frames. Initial permutations are obtained via a Hungarian algorithm [55] applied to matching costs obtained by Alexnet [45] features. As the data is manually annotated, the ground-truth relative maps are known.

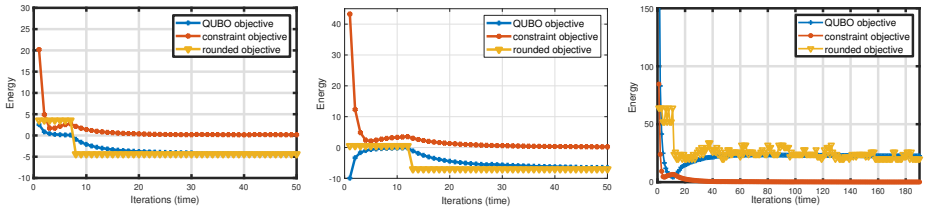
Baselines. We compare Q-FWAL against the classical algorithms of MatchEIG [52], MatchALS [76], MatchLift [39], MatchBirkhoff [8] as well as the first Quantum approach, QuantumSync [7]. QuantumSync uses $\lambda = 2.5$ in all experiments. The *exhaustive* solution is obtained by enumerating all possible permutations. Note that due to the limitations in the available DWave time, we had to implement an *early-stopping* heuristic whose details are provided in the supplementary document. The number or qubits in the minor embeddings in this experiment (for $\mathbf{Q} \in \mathbb{R}^{64\times 64}$) was ≈270, and the chain length did not exceed eight.

Results. We follow the protocol of Birdal *et al.* [7] and report in Tab. 2, the portion of correct bits *i.e. accuracy*. Our approach consistently and significantly outperforms both the classical algorithms and the state-of-the-art quantum approach, QuantumSync [7]. *[7]-search* denotes the softly-constrained search detailed in [7]. Overall, Q-FW is more applicable to problems of growing size.

5.3 Ablation Studies

Tightness of the copositive relaxation. To assess the tightness of our algorithm, we randomly generate fully connected, synthetic synchronization problems with $\mathtt{N} = 3$ and $\mathtt{K} = 3$ with different noise levels $\sigma \in \{0, 0.2\}$. For this small problem we could use an exact QUBO solver and monitor the convergence of the relaxed problem to the ground truth (GT): $\varepsilon_{\mathrm{CC}} = |\mathrm{Tr}(\mathbf{Q}\mathbf{X}_t) - \mathrm{Tr}(\mathbf{Q}\mathbf{X}_t^{\mathrm{gt}})|$ where $\mathbf{X}^{\mathrm{gt}}$ is obtained by lifting the GT permutations.





(a) Graph matching ($\mathbf{Q} \in \mathbb{R}^{9\times 9}$)
(b) Graph matching ($\mathbf{Q} \in \mathbb{R}^{16\times 16}$)
(c) Synchronization ($\mathbf{Q} \in \mathbb{R}^{64\times 64}$).

Fig. 2: Solving two graph matching and one synchronization problem using Q-FWAL. The problem gets more complex from left to right. Thus, the required number of iterations to converge increases.

As shown on the right, ε_{CC} decreases monotonically for all methods, even in the case of noise. Moreover, our D-Wave implementation strictly matches FWAL.

Monitoring convergence. As heuristic fearly stopping criteria are harmful for the convergence guarantees we provide, it is of interest to see how our algorithm behaves as iterations progress. In Fig. 2 we plot minimization curves for different problems we consider: two graph matching (a,b) and one synchronization (c). For each problem, we plot the QUBO objective, infeasibility eror (constraint objective) and the error attained after Hungarian rounding. It is visible that simplicity of the problem has a positive impact on finding good solutions early on. For larger problems, settling on a good solution can take > 200 iterations, when early stopping is not used. We also note that the QUBO ojective converges to the rounded objective, indicating the tightness of our relaxation.

On the evolution of sub-problems & sparsity. We now visually compare the sub-problems emerging in solving the noiseless, synthetic synchronization problem (detailed in the previous experiment and in our supplementary material), for our exact method and for the D-Wave implementation. As seen in Fig. 3, there is no noticeable difference between the two evolutions, confirming that D-Wave could solve the sub-QUBO-problems reliably. Moreover, over iterations the sparsity pattern of $\mathbf{W}_t$ is fixed, which means that we could compute the minor embedding[12], and re-use it throughout Q-FW. This ability of avoiding repetitive minor embeddings makes Q-FW a practically feasible algorithm.

6 Discussions and Conclusion

We have proposed Q-FW, a quantum computation backed, hybrid Frank Wolfe Augmented Lagrangian method. Thanks to the tight copositive relaxation and the QUBO formulation, our algorithm has successfully satisfied linear (in)equality

[12] requires solving a combinatorial optimization problem with heuristics.

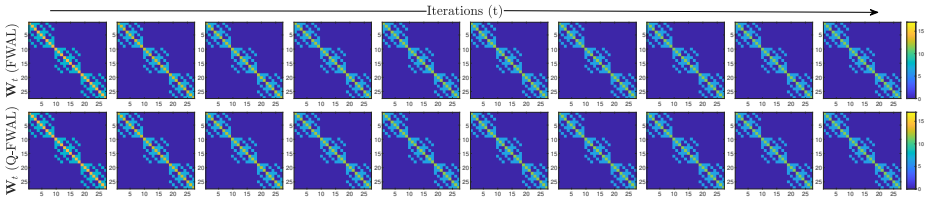


Fig. 3: Evolution of the gradient $\mathbf{W}_t$ for $0 < t < 100$ sampled in steps of 10: FWAL (top) and Q-FWAL (bottom).

constraints, such as permutation-ness, arising in many computer vision applications. We have solved the intermediary QUBO problems on a quantum computer to obtain high quality update directions and demonstrated the validity of Q-FW both on graph matching and permutation synchronization.

Limitations. The most obvious concern is the sub-linear convergence of our algorithm, which could sometimes require a large number of iterations particularly if high accuracy is needed. However,the rates of FWAL (hence, Q-FWAL) are optimal, *i.e.* they match the worst-case computational lower bounds for a generic class of linear optimization based convex programming algorithms [48]. Certain design variants of FW with stronger oracles [47] achieve faster rates. Such variants and their implications for Q-FWAL pose valuable questions for future study. We observed in practice a maximum of 300-400 iterations can be sufficient thanks to the high quality of the DWave solver. We are also limited by the small problem sizes, just likes the previous studies [7,67]. Yet, quantum computers evolve steadily and we are hopeful that the problems we could solve will only grow with time.

Looking forward. Q-FWAL leaves ample room for future works. First, a plethora of QCV methods concerned with constraint satisfaction can benefit our approach. Using our algorithm to ensure constraints other than permutations (especially inequalities like partial permutations) is a future study. We would also like to deploy our algorithm in training vector-output RELU networks [64].

Acknowledgements

A.Y. received support from the Wallenberg AI, Autonomous Systems and Software Program (WASP) funded by the Knut and Alice Wallenberg Foundation.

References

1. Albash, T., Lidar, D.A.: Adiabatic quantum computation. Reviews of Modern Physics **90**(1), 015002 (2018) 9
2. Arrigoni, F., Ricci, E., Pajdla, T.: Multi-frame motion segmentation by combining two-frame results. International Journal of Computer Vision pp. 1–33 (2022) 11
3. Arthur, D., Pusey-Nazzaro, L., et al.: Qubo formulations for training machine learning models. Scientific Reports **11**(1), 1–10 (2021) 9
4. Arute, F., Arya, K., Babbush, R., Bacon, D., Bardin, J.C., Barends, R., Biswas, R., Boixo, S., Brandao, F.G.S.L., Buell, D.A., et al.: Quantum supremacy using a programmable superconducting processor. Nature **574**(7779), 505–510 (2019) 2
5. Benkner, M.S., Lähner, Z., Golyanik, V., Wunderlich, C., Theobalt, C., Moeller, M.: Q-match: Iterative shape matching via quantum annealing. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 7586–7596 (2021) 3, 4
6. Bernard, F., Theobalt, C., Moeller, M.: Ds*: Tighter lifting-free convex relaxations for quadratic matching problems. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (June 2018) 10
7. Birdal, T., Golyanik, V., Theobalt, C., Guibas, L.J.: Quantum permutation synchronization. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 13122–13133 (2021) 2, 3, 9, 10, 11, 12, 14
8. Birdal, T., Simsekli, U.: Probabilistic permutation synchronization using the riemannian structure of the birkhoff polytope. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 11105–11116 (2019) 1, 11, 12
9. Birdal, T., Simsekli, U., Eken, M.O., Ilic, S.: Bayesian pose graph optimization via bingham distributions and tempered geodesic mcmc. Advances in Neural Information Processing Systems **31** (2018) 11
10. Bomze, I.M., Dür, M., De Klerk, E., Roos, C., Quist, A.J., Terlaky, T.: On copositive programming and standard quadratic optimization problems. Journal of Global Optimization **18**(4), 301–320 (2000) 4
11. Bomze, I.M., Rinaldi, F., Zeffiro, D.: Frank–wolfe and friends: a journey into projection-free first-order optimization methods. 4OR **19**(3), 313–345 (2021) 4
12. Born, M., Fock, V.: Beweis des adiabatensatzes. Zeitschrift für Physik **51**(3), 165–180 (1928) 2
13. Burer, S.: On the copositive representation of binary and continuous nonconvex quadratic programs. Mathematical Programming **120**(2), 479–495 (2009) 4, 5, 6
14. Caetano, T.S., McAuley, J.J., Cheng, L., Le, Q.V., Smola, A.J.: Learning graph matching. IEEE transactions on pattern analysis and machine intelligence **31**(6), 1048–1058 (2009) 1
15. Cai, J., Macready, W.G., Roy, A.: A practical heuristic for finding graph minors. arXiv e-prints (2014) 9
16. Cavallaro, G., Willsch, D., Willsch, M., Michielsen, K., Riedel, M.: Approaching remote sensing image classification with ensembles of support vector machines on the d-wave quantum annealer. In: IEEE International Geoscience and Remote Sensing Symposium (IGARSS) (2020) 3
17. Chin, T.J., Suter, D., Chng, S.F., Quach, J.: Quantum robust fitting. arXiv preprint arXiv:2006.06986 (2020) 3
18. Cho, M., Alahari, K., Ponce, J.: Learning graphs to match. In: International Conference on Computer Vision (ICCV). pp. 25–32 (2013) 11, 12

19. Clarkson, K.L.: Coresets, sparse greedy approximation, and the Frank-Wolfe algorithm. ACM Transactions on Algorithms (TALG) **6**(4) (2010) 4
20. D-Wave Systems, Inc.: dwave-system documentation. https://docs.ocean.dwavesys.com/_/downloads/system/en/latest/pdf/ (2022), online; latest access on 05.03.2022 9, 10
21. Dai, A., Nießner, M., Zollhöfer, M., Izadi, S., Theobalt, C.: Bundlefusion: Real-time globally consistent 3d reconstruction using on-the-fly surface reintegration. ACM Transactions on Graphics (TOG) **36**(4), 76a (2017) 1
22. Dattani, N., Szalay, S., Chancellor, N.: Pegasus: The second connectivity graph for large-scale quantum annealing hardware. arXiv e-prints (2019) 2, 3
23. De Klerk, E., Pasechnik, D.V.: Approximation of the stability number of a graph via copositive programming. SIAM Journal on Optimization **12**(4), 875–892 (2002) 4
24. Deng, H., Birdal, T., Ilic, S.: Ppf-foldnet: Unsupervised learning of rotation invariant 3d local descriptors. In: Proceedings of the European Conference on Computer Vision (ECCV). pp. 602–618 (2018) 1
25. Doan, A.D., Sasdelli, M., Chin, T.J., Suter, D.: A hybrid quantum-classical algorithm for robust fitting. arXiv preprint arXiv:2201.10110 (2022) 3
26. Dukanovic, I., Rendl, F.: Copositive programming motivated bounds on the stability and the chromatic numbers. Mathematical programming **121**(2), 249–268 (2010) 4
27. Dür, M.: Copositive programming–a survey. In: Recent advances in optimization and its applications in engineering, pp. 3–20. Springer (2010) 4
28. Dür, M., Rendl, F.: Conic optimization: A survey with special focus on copositive optimization and binary quadratic problems. EURO Journal on Computational Optimization **9**, 100021 (2021) 4
29. Farhi, E., Goldstone, J., Gutmann, S., Lapan, J., Lundgren, A., Preda, D.: A quantum adiabatic evolution algorithm applied to random instances of an np-complete problem. Science **292**(5516), 472–475 (2001) 8
30. Frank, M., Wolfe, P., et al.: An algorithm for quadratic programming. Naval research logistics quarterly (1956) 4
31. Gao, M., Lahner, Z., Thunberg, J., Cremers, D., Bernard, F.: Isometric multi-shape matching. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 14183–14193 (2021) 11
32. Golyanik, V., Theobalt, C.: A quantum computational approach to correspondence problems on point sets. In: Computer Vision and Pattern Recognition (CVPR) (2020) 2
33. Govindu, V.M.: Lie-algebraic averaging for globally consistent motion estimation. In: Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004. vol. 1, pp. I–I. IEEE (2004) 11
34. Grover, L.K.: A fast quantum mechanical algorithm for database search. In: Proceedings of the twenty-eighth annual ACM symposium on Theory of computing. pp. 212–219 (1996) 9
35. Hazan, E.: Sparse approximate solutions to semidefinite programs. In: Latin American symposium on theoretical informatics. pp. 306–316. Springer (2008) 4
36. Hu, F., Wang, B.N., Wang, N., Wang, C.: Quantum machine learning with d-wave quantum computer. Quantum Engineering **1**(2), e12 (2019) 2
37. Huang, J., Birdal, T., Gojcic, Z., Guibas, L.J., Hu, S.M.: Multiway non-rigid point cloud registration via learned functional map synchronization. arXiv preprint arXiv:2111.12878 (2021) 11

38. Huang, J., Wang, H., Birdal, T., Sung, M., Arrigoni, F., Hu, S.M., Guibas, L.J.: Multibodysync: Multi-body segmentation and motion estimation via 3d scan synchronization. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 7108–7118 (2021) 11
39. Huang, Q.X., Guibas, L.: Consistent shape maps via semidefinite programming. In: Eurographics/ACM SIGGRAPH Symposium on Geometry Processing (SGP). Eurographics Association (2013) 12
40. Huber, D.F., Hebert, M.: Fully automatic registration of multiple 3d data sets. Image and Vision Computing **21**(7), 637–650 (2003) 1
41. Jaggi, M.: Revisiting frank-wolfe: Projection-free sparse convex optimization. In: International Conference on Machine Learning. pp. 427–435. PMLR (2013) 4
42. Kadowaki, T., Nishimori, H.: Quantum annealing in the transverse ising model. Phys. Rev. E **58**, 5355–5363 (1998) 2
43. Kezurer, I., Kovalsky, S.Z., Basri, R., Lipman, Y.: Tight relaxation of quadratic matching. In: Computer Graphics Forum. vol. 34, pp. 115–128. Wiley Online Library (2015) 1
44. Kirkpatrick, S., Gelatt, C.D., Vecchi, M.P.: Optimization by simulated annealing. Science **220**(4598), 671–680 (1983) 10
45. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. In: Advances in neural information processing systems (NeurIPS). pp. 1097–1105 (2012) 12
46. Kuhn, H.W.: The hungarian method for the assignment problem. Naval research logistics quarterly **2**(1-2), 83–97 (1955) 8, 9
47. Lacoste-Julien, S., Jaggi, M.: On the global linear convergence of frank-wolfe optimization variants. Advances in neural information processing systems **28** (2015) 14
48. Lan, G.: The complexity of large-scale convex programming under a linear optimization oracle. arXiv preprint arXiv:1309.5550 14
49. Levitin, E., Polyak, B.: Constrained minimization methods. USSR Comput. Math. & Math. Phys. **6**(5), 1–50 (1966) 4
50. Li, J., Ghosh, S.: Quantum-soft qubo suppression for accurate object detection. In: European Conference on Computer Vision (ECCV) (2020) 2, 3
51. Li, X., Larson, M., Hanjalic, A.: Pairwise geometric matching for large-scale object retrieval. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 5153–5161 (2015) 1
52. Maset, E., Arrigoni, F., Fusiello, A.: Practical and efficient multi-view matching. In: International Conference on Computer Vision (ICCV) (2017) 11, 12
53. McGeoch, C., Farré, P.: The advantage system: Performance update. Tech. rep., Technical Report 14-1054A-A, D-Wave Systems Inc., Burnaby, Canada (2021) 9
54. Mirsky, L.: Symmetric gauge functions and unitarily invariant norms. The quarterly journal of mathematics **11**(1), 50–59 (1960) 8
55. Munkres, J.: Algorithms for the assignment and transportation problems. Journal of the society for industrial and applied mathematics **5**(1), 32–38 (1957) 12
56. Mur-Artal, R., Tardós, J.D.: ORB-SLAM2: an open-source SLAM system for monocular, stereo and RGB-D cameras. IEEE Transactions on Robotics **33**(5), 1255–1262 (2017). https://doi.org/10.1109/TRO.2017.2705103 1
57. Neven, H., Denchev, V.S., Rose, G., Macready, W.G.: Qboost: Large scale classifier training with adiabatic quantum optimization. In: Asian Conference on Machine Learning (ACML) (2012) 3

58. Neven, H., Rose, G., Macready, W.G.: Image recognition with an adiabatic quantum computer i. mapping to quadratic unconstrained binary optimization. arXiv e-prints (2008) 3
59. Nguyen, N.T.T., Kenyon, G.T.: Image classification using quantum inference on the d-wave 2x. arXiv e-prints (2019) 3
60. Nielsen, M.A., Chuang, I.: Quantum computation and quantum information (2002) 2
61. Povh, J., Rendl, F.: Copositive and semidefinite relaxations of the quadratic assignment problem. Discrete Optimization **6**(3), 231–241 (2009) 4
62. Technical description of the d-wave quantum processing unit. https://docs.dwavesys.com/docs/latest/doc_qpu.html, accessed on the 05.03.2022 3
63. Quist, A.J., de Klerk, E., Roos, C., Terlaky, T.: Copositive realxation for general quadratic programming. Optimization methods and software **9**(1-3), 185–208 (1998) 4
64. Sahiner, A., Ergen, T., Pauly, J.M., Pilanci, M.: Vector-output relu neural network problems are copositive programs: Convex analysis of two layer networks and polynomial-time algorithms. In: International Conference on Learning Representations (2020) 4, 14
65. Sattler, T., Weyand, T., Leibe, B., Kobbelt, L.: Image retrieval for image-based localization revisited (2012) 1
66. Schonberger, J.L., Frahm, J.M.: Structure-from-motion revisited. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 4104–4113 (2016) 1
67. Seelbach Benkner, M., Golyanik, V., Theobalt, C., Moeller, M.: Adiabatic quantum graph matching with permutation matrix constraints. In: International Conference on 3D Vision (3DV) (2020) 2, 3, 9, 10, 11, 14
68. Shor, P.W.: Algorithms for quantum computation: discrete logarithms and factoring. In: Annual Symposium on Foundations of Computer Science (1994) 2
69. Systems, D.: D-wave details product expansion & cross platform roadmap (2021), https://tinyurl.com/yeyknn7y 2
70. Xiang, R., Lai, R., Zhao, H.: Efficient and robust shape correspondence via sparsity-enforced quadratic assignment. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 9513–9522 (2020) 1
71. Yasuoka, H.: Computational complexity of quadratic unconstrained binary optimization. arXiv preprint arXiv:2109.10048 (2021) 9
72. Yurtsever, A., Fercoq, O., Cevher, V.: A conditional-gradient-based augmented lagrangian framework. In: International Conference on Machine Learning. pp. 7272–7281. PMLR (2019) 2, 4, 8
73. Yurtsever, A., Fercoq, O., Locatello, F., Cevher, V.: A conditional gradient framework for composite convex minimization with applications to semidefinite programming. In: International Conference on Machine Learning. pp. 5727–5736. PMLR (2018) 2, 4
74. Yurtsever, A., Tropp, J.A., Fercoq, O., Udell, M., Cevher, V.: Scalable semidefinite programming. SIAM Journal on Mathematics of Data Science **3**(1), 171–200 (2021) 4
75. Zaech, J.N., Liniger, A., Danelljan, M., Dai, D., Van Gool, L.: Adiabatic quantum computing for multi object tracking. arXiv preprint arXiv:2202.08837 (2022) 2, 3, 10
76. Zhou, X., Zhu, M., Daniilidis, K.: Multi-image matching via fast alternating minimization. In: International Conference on Computer Vision (ICCV). pp. 4032–4040 (2015) 12