A Quantum Computational Approach to Correspondence Problems on Point Sets

Vladislav Golyanik
Max Planck Institute for Informatics, Saarland Informatics Campus

Christian Theobalt

Abstract

Modern adiabatic quantum computers (AQC) are already used to solve difficult combinatorial optimisation problems in various domains of science. Currently, only a few applications of AQC in computer vision have been demonstrated. We review AQC and derive a new algorithm for correspondence problems on point sets suitable for execution on AQC. Our algorithm has a subquadratic computational complexity of the state preparation. Examples of successful transformation estimation and point set alignment by simulated sampling are shown in the systematic experimental evaluation. Finally, we analyse the differences in the solutions and the corresponding energy values.

1. Introduction

Since their proposal in the early eighties [8, 43, 27], quantum computers have attracted much attention of physicists and computer scientists. Impressive advances both in quantum computing hardware and algorithms have been demonstrated over the last thirty years [40, 30, 61, 58, 42, 19, 25, 49, 65, 48]. Quantum computers are not universally faster than conventional machines, but they can natively execute algorithms relying on quantum parallelism, i.e., the ability to perform operations on exponentially many superimposed memory states simultaneously [59].

To harness the advantages, carefully designed algorithms are required. Nowadays, the motivation to take advantage of quantum effects in computing is also facilitated by the classical computing paradigm approaching its limits, since the quantum effects are becoming non-neglectable while manufacturing and using conventional CPUs. As a result, alternative paradigms such as massively parallel computing devices have been brought into being.

While universal gate quantum computer technology has not yet reached the maturity, modern adiabatic quantum annealers (AQA) are already capable of solving difficult real-world combinatorial optimisation problems [15, 14, 23, 49]. The primary difference of universal gate quantum computing and AQA is that the latter can address objectives formulated as quadratic unconstrained binary optimisation problems (QUBOP) defined as

$$
    \arg \min_{q \in \mathbb{B}^n} q^T P q,
$$

where $q$ is a set of $n$ binary variables, and $P$ is a symmetric matrix of weights between the variables. The operational principle of AQA is grounded on the adiabatic theorem of quantum mechanics [17] which states that

if a quantum-mechanical system is in the ground state of a time-dependent Hamiltonian and parameters of this Hamiltonian are changing gradually enough, the system will continue to remain in the ground state during the evolution (see Table 1 for quantum notions). In their seminal paper, Farhi et al. [26] have shown that the adiabatic principle (2) can be used for solving $\mathcal{NP}$-complete optimisation problems and laid the foundation for adiabatic quantum computing. Several years later, Aharonov et al. [2] theoretically showed the equivalence between classical quantum computing and quantum annealing models. As of 2019-2020, general-purpose quantum computers accessible for research purposes and applications contain up to 20 qubits [19]. In contrast, the latest quantum annealers support up to $2^{10}$ qubits [21]. Nevertheless, due

the amount of logical qubits which are available on this system is an order of magnitude lower since most qubits are used for the error correction.
to design and practical restrictions, quantum algorithms for the gates model such as Shor’s prime number factorisation [61] or Grover’s search algorithms [30] cannot be implemented on current quantum annealers.

Motivation and Contributions. Considering recent successful applications of AQA in several fields of computational science [42, 49, 65], we are motivated to investigate how useful AQA can be for computer vision and which problems can be potentially solved on the new hardware. The vast majority of available materials about quantum annealers are either oriented to physicists or lack technical details and clarity. Our goal is to fill this gap, introduce the reader into the modern AQA and provide all notions and the background to understand, analyse, simulate and design quantum algorithms for computer vision which can potentially run on modern AQA, as well as interpret the results.

We consider correspondence problems on point sets which have various applications in computer vision. They consist in finding an optimal rigid transformation between inputs [34, 13, 47, 66]. While transformation estimation assumes known matches, point set alignment is more general and targets, in addition, the recovery of correspondences. We consider two inputs, i.e., a fixed reference point set and a template undergoing a rigid transformation. Thus, our goal is to design a quantum approach for point set alignment which can potentially run on AQA and show that it offers advantages compared to the classical counterparts.

Therefore, we adapt the recent progress in rigid point set alignment and formulate a globally multiply-linked energy functional which does not require any intermediate correspondence updates [29]. In the gravitational approach (GA) [29], the optimal alignment is achieved when the gravitational potential energy (GPE) of the system with two interacting particle swarms is locally minimal. Proceeding from GA, we build the weight matrix $P$ for the associated QUBOP (1) which is unalterably valid in the course of the optimisation process. Along with that, we are targeting at a method QUBOP (1) which is unalterably valid in the course of the optimisation process. The vast majority of available materials about quantum annealing [39] or Grover’s search algorithms [30] cannot be interpreted in the classical optimisation theory for computer vision.

Qubit. Quantum computing encompasses tasks which can be performed on quantum-mechanical systems [53]. Quantum superposition and entanglement are two forms of parallelism evidenced in quantum computers. A qubit is a quantum-mechanical equivalent of a classical bit. A qubit $|\phi\rangle$ — written in the Dirac notation — can be in the state $|0\rangle, |1\rangle$ or an arbitrary superposition of both states denoted by $|\phi\rangle = \alpha|0\rangle + \beta|1\rangle$, where $\alpha$ and $\beta$ are the (generally, complex) probability amplitudes satisfying $|\alpha|^2 + |\beta|^2 = 1$. In quantum computing, the state $|0\rangle$ denoted by $|\psi\rangle$ is often used for initialisation of a qubit register. The state of a qubit remains hidden during the entire computation and reveals when measured. If qubits are entangled, measuring one of them influences the measurement outcome of the other one [59]. During the measurement, the qubit’s state irreversibly collapses to one of the basis states $|0\rangle$ or $|1\rangle$. Efficient physical realisation of a qubit demand very low temperatures. Otherwise, thermal fluctuations will destroy it and lead to arbitrary changes of the measured qubit state.

One possible physical implementation of a qubit is an electron which possesses a spin, i.e., its intrinsic magnetic moment [53, 65]. The spin of an electron can be manipulated and brought to the state spin down, spin up, or a superposition of both. A concrete experimentally realised scheme that uses this property is represented by an atom of phosphorus $^{31}$P embedded into a $^{28}$Si silicon lattice attached to a transistor [37, 46, 67]. The nucleus of $^{31}$P has a positive charge compensated by electrons. The bundle of electrons in the transistor is filled up to the energetic level between the energy of spin-down and spin-up state of $^{31}$P. To change a state of a $^{31}$P–$^{28}$Si qubit, a microwave pulse of the frequency — which is equal to the resonance frequency of the atom — is applied to it. The new state $|\psi\rangle$ depends on the duration of the exposure. A transistor is used to measure a state of the $^{31}$P–$^{28}$Si qubit. If the extra electron of $^{31}$P tunnels into the electron bundle, a positive charge is measured in the transistor indicating the spin-up state (e.g., $|1\rangle$).

Fig. 2-(a) visualises a qubit with a so-called Bloch sphere. Every qubit can be both in a superposition and en-

<table>
<thead>
<tr>
<th>quantum notion</th>
<th>classical counterpart</th>
</tr>
</thead>
<tbody>
<tr>
<td>qubit (states 0 and 1)</td>
<td>bit (states 0 and 1)</td>
</tr>
<tr>
<td>(time-dependent) Hamiltonian eigenstate</td>
<td>energy functional</td>
</tr>
<tr>
<td>ground state</td>
<td>some energy state</td>
</tr>
<tr>
<td>quantum system evolution</td>
<td>globally optimal energy state optimisation process</td>
</tr>
<tr>
<td>quantum annealing [26]</td>
<td>simulated annealing [39]</td>
</tr>
</tbody>
</table>

Table 1: Quantum notions and their counterparts in computer vision.

2. Preliminaries, Definitions and Notations

In this section, we introduce the reader into the basics of quantum computing. See Table 1 for a lookup of notions specific to AQA which have counterparts and interpretation in the classical optimisation theory for computer vision.

Qubit. Quantum computing encompasses tasks which can be performed on quantum-mechanical systems [53]. Quantum superposition and entanglement are two forms of parallelism evidenced in quantum computers. A qubit is a quantum-mechanical equivalent of a classical bit. A qubit $|\phi\rangle$ — written in the Dirac notation — can be in the state $|0\rangle, |1\rangle$ or an arbitrary superposition of both states denoted by $|\phi\rangle = \alpha|0\rangle + \beta|1\rangle$, where $\alpha$ and $\beta$ are the (generally, complex) probability amplitudes satisfying $|\alpha|^2 + |\beta|^2 = 1$. In quantum computing, the state $|0\rangle$ denoted by $|\psi\rangle$ is often used for initialisation of a qubit register. The state of a qubit remains hidden during the entire computation and reveals when measured. If qubits are entangled, measuring one of them influences the measurement outcome of the other one [59]. During the measurement, the qubit’s state irreversibly collapses to one of the basis states $|0\rangle$ or $|1\rangle$. Efficient physical realisation of a qubit demand very low temperatures. Otherwise, thermal fluctuations will destroy it and lead to arbitrary changes of the measured qubit state.

One possible physical implementation of a qubit is an electron which possesses a spin, i.e., its intrinsic magnetic moment [53, 65]. The spin of an electron can be manipulated and brought to the state spin down, spin up, or a superposition of both. A concrete experimentally realised scheme that uses this property is represented by an atom of phosphorus $^{31}$P embedded into a $^{28}$Si silicon lattice attached to a transistor [37, 46, 67]. The nucleus of $^{31}$P has a positive charge compensated by electrons. The bundle of electrons in the transistor is filled up to the energetic level between the energy of spin-down and spin-up state of $^{31}$P. To change a state of a $^{31}$P–$^{28}$Si qubit, a microwave pulse of the frequency — which is equal to the resonance frequency of the atom — is applied to it. The new state $|\psi\rangle$ depends on the duration of the exposure. A transistor is used to measure a state of the $^{31}$P–$^{28}$Si qubit. If the extra electron of $^{31}$P tunnels into the electron bundle, a positive charge is measured in the transistor indicating the spin-up state (e.g., $|1\rangle$).

Fig. 2-(a) visualises a qubit with a so-called Bloch sphere. Every qubit can be both in a superposition and en-
An arbitrary Hamiltonian of a and evaluation in quantum annealing. The spectral gap influences for the simplicity, we denote here by \(|\psi\rangle\). For the measurement, the states of variables are classical. Superposition and tunnelling) \([11, 36]\). In particular, it is used to find a ground state of an Ising Hamiltonian \([35, 57]\), which encodes the target computational problem, see Fig. 2-(b). Quantum annealing is the quantum counterpart of simulated annealing \([45, 39]\). Starting from the superposition state \(|+\rangle^{\otimes n}\) (this is a shorthand for \(n\) qubits in the state \(|+\rangle\), cf. \((9)\)), the system evolves according to \((3)\) under an external time-dependent magnetic field (a transverse field). When the external field is faded away, the system reaches the ground state of an Ising model \([35]\). According to \((2)\), if an external magnetic field is changing gradually enough, the system remains near the ground state with high probability throughout the optimisation. Quantum annealing systems taking advantage of \((2)\) are called adiabatic quantum computers (AQC). QUBOP is the most common problem form which can be mapped to current realisations of AQC.

3. Modern Adiabatic Quantum Computation

Adiabatic quantum computation is a form of quantum annealing which relies on the adiabatic theorem of quantum mechanics \((2)\) \([17]\). Starting from a ground state of an initial default Hamiltonian \(\mathbf{H}_I\), an AQC system adiabatically evolves into the ground state of a problem Hamiltonian \(\mathbf{H}_P\) which encodes a solution to a problem \([26]\). In the case of adiabatic quantum annealing (AQA), the problem Hamiltonian \(\mathbf{H}_P\) is given by the Ising model \([35]\):

\[
\mathbf{H}_P = \sum_{j \in V} h_j \sigma_j^z + \sum_{(j, k) \in E_P} J_{j,k} \sigma_j^z \otimes \sigma_k^z, \tag{5}
\]

with the Kronecker product \(\otimes\), \(h_j\) denoting exterior local magnetic fields and \(J_{j,k}\) standing for the pairwise connections between the particles. \(V\) is a set of particles, and \(E_P\) is a set of edges (intra-particle links) of the graph. Eq. \((5)\) is written in a notation common in physics. The first term of \((5)\) on the right side in the explicit notation reads

\[
\mathbf{H}_{P}\in V = \left[ \begin{array}{c}
\sigma^z \otimes I \otimes \cdots \otimes I \\
I \otimes \sigma^z \otimes \cdots \otimes I \\
\vdots \\
I \otimes I \otimes \cdots \otimes \sigma^z
\end{array} \right]_{2^n \times 2^n \times 2^n \times 2^n \times \cdots \times 2^n} + \left[ \begin{array}{c}
h_1 I_{2^n \times 2^n} \\
h_2 I_{2^n \times 2^n} \\
\vdots \\
h_n I_{2^n \times 2^n}
\end{array} \right]_{2^n \times 2^n \times 2^n \times \cdots \times 2^n}.
\]
where I without a subscript is a $2 \times 2$ identity matrix. The second term $\hat{H}_{P}^{(j,k) \in E_{P}}$ of (5) can be expressed in a similar manner, involving pairs of $\sigma^{z}$ in the tensor product depending on the connectivity of the lattice.

Theoretically, each particle can interact with any other particle from the whole set of qubits. In practice, the couplings are restricted to local neighbourhoods (see Sec. 3.2). Thus, (5) describes a system of $N$ interacting spin-$\frac{1}{2}$ particles under the influence of distributed magnetic forces, and in the expanded form, $\hat{H}_{P}$ is a $2^{n} \times 2^{n}$ matrix. Finding a ground state of an Ising model is an $NP$-hard problem [7]. In the ground state, the spin configuration of all particles which minimises Ising energy $E_{\text{Ising}}$ is given by:

$$E_{\text{Ising}} = \sum_{i} h_{i} s_{i} + \sum_{i,j} J_{i,j} s_{i}s_{j},$$  

(7)

where $s_{i} \in \{1, -1\}$ denotes two possible spin measurement outcomes of a spin-$\frac{1}{2}$ particle.

### 3.1. Quantum System Evolution

Solving $NP$-hard problems such as QUBOP on a classical computer requires exponential time in the size of the input. The main idea of the AQC is that a QUBOP (1) can be mapped to the Ising model (5) and optimised by allowing the system to evolve according to the adiabatic principle (2). Once annealing is finished, the qubit register will represent the system to evolve according to the adiabatic principle (2).

The initial Hamiltonian of the system is always initialised with a certain state. The ground state of (8) is achieved when

$$\hat{H}_{I} = - \sum_{j \in V} B_{x} \sigma_{j}^{x},$$  

(8)

where $B_{x} > 0$ stands for a magnetic field pointing in the $x$ direction. The ground state of (8) is a symmetrised superposition with equal normalised probability amplitudes for the states $|0\rangle$ and $|1\rangle$ for all qubits, i.e.,

$$\left[|+\rangle\otimes n\right] = \left(\frac{|0\rangle + |1\rangle}{\sqrt{2}}\right)^{\otimes n}.$$  

(9)

This initial state (9) is comparably easy to construct by radiating a microwave of the same duration and wavelength to all qubits. In mathematical terms, (9) is obtained by applying a Hadamard transform $H = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ to $n$ $|0\rangle$ qubits.

The lowest energy $E_{\text{GS}} = -nB_{x}$ of (8) is achieved when all qubits in the system point in the anti-parallel direction of the magnetic field, so that $\sigma_{j}^{x}(s_{j}) = s_{j}$. During AQC, the initial Hamiltonian $\hat{H}_{I}$ is evolving into the problem Hamiltonian $\hat{H}_{P}$, with a high probability of reaching the ground state of $\hat{H}_{P}$ [26]. The interpolation between the Hamiltonians can be written as

$$\hat{H} = [1 - s] \hat{H}_{I} + s \hat{H}_{P},$$  

(10)

with $s \in [0; 1]$ being the time in relative units from the start of annealing at $s = 0$ until reaching the ground state of $\hat{H}_{P}$ at $s = 1$. The problem Hamiltonian and the final state of the system depend on the objective function $f(x)$ or the matrix of weights between the qubits $P$ in (1). After the annealing is accomplished, the state of each qubit is measured, and the result corresponds to the solution of the programmed problem with a high probability. At this stage, the states of all binary variables are classical, and not quantum anymore.

To remain in the ground state during the system evolution, the annealing rate has to be carefully chosen. The condition of adiabaticity (2) is derived from the time-dependent perturbation theory of quantum systems. It is achieved when the average energy pumped into the system per time interval $T$ is smaller than the minimal energy difference between the ground state and the first excited state. This statement was quantified in [4] which generalises the original adiabatic theorem [17] for periodic driving, see Appendix A for further details.

### 3.2. Quantum Annealer D-WAVE

D-WAVE relies on the adiabatic criterion in its specified form and currently supports up to $\approx 2000$ qubits [22]. It reflects the state of the art in physical realisation of quantum processors. It is relatively inexpensive to bring the system in the superposition state, and every computation on D-WAVE starts with the problem-independent $\hat{H}_{I}$ (8). Qubits can interact with a restricted number of other qubits, and it is possible to define qubit equality and entanglement constraints [22]. Possible interactions can be seen from the chimera graph which schematically depicts the layout of the quantum processor [22, 16]. At the same time, the physically realised connectivity can model QUBOP with arbitrary connectivities through an internal conversion [16]. The drawback is that in the worst case, a quadratic increase in the number of variables is required. A fully connected graph of layers with $N$ qubits would require $N^{2}$ qubits for processing. Some QUBOP cannot be mapped to the chimera graph, and some problems can be mapped in multiple ways [55].

### 4. Previous and Related Work

#### Universal Quantum Computers.

The paradigm of the universal quantum computer originates in the attempts to gain control over individual quantum systems in the early eighties [62, 53]. Later, extending the control to multiple quantum systems has attracted the interest of physicists, promising to facilitate discoveries in quantum physics [53]. By that time, it was noticed that simulating a quantum-mechanical system on a classical computer requires exponential time in the number of simulated elements [43, 27]. “Can you do it² with a new kind of computer – a quantum com-

---

²to simulate quantum-mechanical effects
puter?” [27] is a famous quote by R. Feynman which has triggered research on quantum computers in the subsequent years. The so-called no-cloning theorem [56, 64] belongs to the first discoveries strongly influenced quantum information theory and quantum computations. Nowadays, quantum computers can be used not only to fulfil their primary goal, i.e., to simulate quantum-mechanical systems for different branches of science, but also to solve other computational problems — such as balanced function decision problem [24], quantum Turing machines for complexity analysis [12], prime number factorisation and discrete logarithms [61], database search [30], and graph matching [3], data classification [58] and principal component analysis [42] — faster than on classical machines. The related field of quantum communication and quantum key distribution has already found broad practical use nowadays [10, 9, 60].

Classical Methods using Quantum Analogies. Quantum-mechanical effects inspired multiple techniques for conventional computers including variants of genetic and evolutionary algorithms [31, 32], non-rigid mesh analysis [5] and image segmentation [6], among others.

Quantum Annealers in Computer Vision. Only a few theoretical results and applications of AQC to image processing, machine learning and computer vision are known. Neven et al. [51] have shown how image recognition can be formulated as QUBOP. Image classification on 12 × 12 images with AQC was addressed in [52]. The approach of O’Malley et al. can learn facial features and reproduce facial image collections [54]. Boyda et al. [18] propose an AQC method to detect areas with trees from aerial images. Several methods target classification, dimensionality reduction and training of deep neural networks [50, 38, 1]. Not all theoretical findings of these works are possible to test on the real AQC hardware yet. Nonetheless, we believe that it is essential to explore the theory and highlight the advantages of the upcoming hardware for computer vision tasks.

5. Quantum Transformation Estimation

In this section, we introduce our QA to transformation estimation. The inputs are a reference point set \( \mathbf{x}_n \in \mathbb{R}^{D \times N} \) and a template point set \( \mathbf{y}_n \in \mathbb{R}^{D \times N} \), \( n \in \{1, \ldots, N\} \). \( N \) is the number of points in both point sets and \( D \) is the dimensionality of the points. We assume that translation is resolved, the centroids of the point sets coincide, and points are in correspondence.

5.1. Transformation Estimation in 2D

To obtain an advantage in solving transformation estimation on a quantum annealer we should avoid uniform sampling of rotations applied to \( \mathbf{Y} \). Elements of the rotation group are non-commutative, and it is not possible to formulate multiplication of basis rotations as QUBOP. Instead, we propose to represent the transformation matrix as a linear combination of basis elements. Recall that for any rotation matrix, \( \mathbf{R}^{-1} = \mathbf{R}^T \). Rotation in 2D consists of four elements, i.e., \( \mathbf{R} = \begin{pmatrix} r_{1.2} & r_{2.2} \\ r_{2.1} & r_{2.2} \end{pmatrix} \). Additively, we can create a basis for all possible values of \( \mathbf{R} \) and encode the influence of the additive elements as binary variables. Consider instead the power series of \( \mathbf{R} \) in 2D. Every such matrix has a corresponding skew-symmetric matrix of the form

\[
\mathbf{S} = \theta \mathbf{M}, \quad \mathbf{M} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},
\]

with a real number \( \theta \). According to the Cayley-Hamilton theorem, \( \mathbf{S}^2 + \theta^2 \mathbf{I} = 0 \) which leads to the following exponential map for \( \mathbf{R} \) with power series:

\[
\mathbf{R} = \exp(M) = \cos(\theta) \mathbf{I} + \sin(\theta) \mathbf{M}.
\]

From (12) we see that \( \mathbf{R} \) is composed of an identity weighted by \( \cos(\theta) \) and \( \mathbf{M} \) weighted by \( \sin(\theta) \). If the basis would resemble additive elements \( \mathbf{I} \) and \( \mathbf{M} \) of the exponential map, we can stronger constrain the resulting \( \mathbf{R} \). We see that \( r_{1.1} \) is entangled with \( r_{2.2} \), and \( r_{1.2} \) is entangled with \( r_{2.1} \). Eventually, we need fewer basis elements, the optimisation will finish faster and the method can be also implemented and tested on a classical computer. Thus, our basis \( \mathbf{Q} = \{ \mathbf{Q}_k \} \) for \( \mathbf{R} \) is a compound of 20 elements:

\[
\mathbf{Q}_k = \mathbf{\omega} \mathbf{C} \in \mathbb{R}^{2 \times 2}, \forall \mathbf{\omega} \in \{0.5, 0.2, 0.1, 0.1, 0.05\}, \\
\forall \mathbf{C} \in \{ \mathbf{I}, \mathbf{M}, -\mathbf{I}, -\mathbf{M} \}.
\]

Next, we set the weight matrix in (1) as

\[
\mathbf{P} = \mathbf{\Phi} \mathbf{\Phi}^T,
\]

and the final QUBOP reads

\[
\arg \min_{\mathbf{q} \in \mathbb{B}^{2^N}} \mathbf{q}^T \mathbf{\Phi} \mathbf{\Phi}^T \mathbf{q}.
\]

In total, 21 qubits are required to resolve the transformation on AQC in 2D, with the first qubit of \( \mathbf{q} \) being fixed to \( |1\rangle \). After solving (16) with quantum annealing and measuring \( \mathbf{q} \), we obtain a classical bitstring \( \hat{\mathbf{q}} \). The resulting (perhaps approximate) \( \mathbf{R} \) is then obtained by unembedding as

\[
\mathbf{R} = \sum_{k=1}^{K} \hat{q}_{k+1} \mathbf{Q}_k.
\]
5.2. Transformation Estimation in 3D

In 3D, a skew-symmetric matrix can be represented as

\[
S = \theta \mathbf{M}, \quad \mathbf{M} = \begin{bmatrix}
  m_{1,1} & m_{1,2} & m_{1,3} \\
  m_{2,1} & m_{2,2} & m_{2,3} \\
  m_{3,1} & m_{3,2} & m_{3,3}
\end{bmatrix} = \begin{bmatrix}
  0 & a & b \\
  -a & 0 & c \\
  -b & -c & 0
\end{bmatrix},
\]

where \( \theta, a, b \) and \( c \) are real numbers, and \( a^2 + b^2 + c^2 = 1 \). In the 3D case, the Cayley-Hamilton theorem states that \(-S^3 - \theta^2 S = 0\). The exponential map for \( R \) in 3D with power series reads

\[
R = \exp(S) = \mathbf{I} + \left( \frac{\sin \theta}{\theta} \right) S + \left( \frac{1 - \cos \theta}{\theta^2} \right) S^2 = \mathbf{I} + \sin \theta \mathbf{M} + (1 - \cos \theta) \mathbf{M}^2.
\]

Next, \( \mathbf{M} \) can be decomposed as follows:

\[
\mathbf{M} = a \begin{bmatrix}
  0 & 1 & 0 \\
  -1 & 0 & 0 \\
  0 & 0 & 0
\end{bmatrix} + b \begin{bmatrix}
  0 & 0 & 1 \\
  0 & 0 & 0 \\
  -1 & 0 & 0
\end{bmatrix} + c \begin{bmatrix}
  0 & 0 & 0 \\
  0 & 0 & 1 \\
  0 & -1 & 0
\end{bmatrix} .
\]

Regarding \( \mathbf{M} \), we see that

- \( \{m_{1,2}; m_{2,1}\}, \{m_{1,3}; m_{3,1}\} \) and \( \{m_{2,3}; m_{3,2}\} \) are mutually dependent or entangled,
- \( m_{i,j} \in [-1, 1] \),
- \( \mathbf{M} = -\mathbf{M}^T \), \( \mathbf{M} = -\mathbf{M}^T \), \( \mathbf{M}_b = -\mathbf{M}_b^T \) and \( \mathbf{M}_c = -\mathbf{M}_c^T \), i.e., they are anti-symmetric, and
- \( \mathbf{M}^2 = \begin{bmatrix}
  v_1^2 & d & e \\
  d & v_2^2 & f \\
  e & f & v_3^2
\end{bmatrix} \) is symmetric negative semidefinite, with \( \{v_1, v_2, v_3\} \in \mathbb{R}^+ \), and \( \{d, e, f\} \in \mathbb{R} \).

The basis for rotation in 3D is comprised of the identity matrix \( \mathbf{I} \), \( \mathbf{M}_a \), \( \mathbf{M}_b \), \( \mathbf{M}_c \) as well as the basis for \( \mathbf{M}^2 \):

\[
\mathbf{M}_d = \begin{bmatrix}
  0 & 1 & 0 \\
  1 & 0 & 0 \\
  0 & 0 & 0
\end{bmatrix}, \quad \mathbf{M}_e = \begin{bmatrix}
  0 & 0 & 1 \\
  0 & 0 & 0 \\
  0 & 0 & 1
\end{bmatrix}, \quad \mathbf{M}_f = \begin{bmatrix}
  0 & 0 & 0 \\
  0 & 1 & 0 \\
  0 & 0 & 1
\end{bmatrix}.
\]

Thus, our basis \( \mathbf{Q}^{3D} = \{\mathbf{Q}_k^{3D}\} \) for \( \mathbf{R} \) in 3D is a compound of \( K = 80 \) elements:

\[
\{\mathbf{Q}_k^{3D} = \omega \mathbf{C}^{3D} \in \mathbb{R}^{3 \times 3}, \forall \omega \in \{0.5, 0.2, 0.1, 0.05\}\}, \quad \forall \mathbf{C}^{3D} \in \{\mathbf{I}, -\mathbf{I}, \mathbf{M}_a, -\mathbf{M}_a, \mathbf{M}_b, -\mathbf{M}_b, \mathbf{M}_c, -\mathbf{M}_c, \mathbf{M}_d, -\mathbf{M}_d, \mathbf{M}_e, -\mathbf{M}_e, \mathbf{M}_f, -\mathbf{M}_f\}\}.
\]

The final QUBOP and the unembedding (i.e., decoding the solution to QUBOP) after quantum annealing for the 3D case are obtained similarly to (14)–(17) with \( \mathbf{q} \in \mathbb{B}^{81} \) (\( \mathbf{q}_0 \) remains fixed to \( |1\rangle \) and \( \mathbf{Q}_k \) are replaced by \( \mathbf{Q}_k^{3D} \) in (17)).

6. Quantum Point Set Registration

In point set registration, the input point sets are of different cardinalities, and correspondences between points are, generally, not known, i.e., \( \{x_n\} \in \mathbf{X} \in \mathbb{R}^{D \times N} \) and \( \{y_m\} \in \mathbf{Y} \in \mathbb{R}^{D \times M} \), \( m \in \{1, \ldots, M\} \). \( N \) and \( M \) are the numbers of points in the reference and template, respectively, while \( D \) is the point dimensionality. The objective of point set alignment is to recover rotation \( \mathbf{R} (\mathbf{R}^{-1} = \mathbf{R}^T) \), \( \det(\mathbf{R}) = 1 \) and translation \( \mathbf{t} \) aligning \( \mathbf{Y} \) to \( \mathbf{X} \). We assume that the translation is resolved in the pre-processing step by bringing the point set centroids into coincidence.

Point set alignment can be alternatingly solved on AQC by finding some point matches and estimating the transformation with the given correspondences in the ICP fashion [13]. This would result in a sequence of QUBOP of the form (16). To express alignment as a single QUBOP, we have to find an energy functional which is correspondence-free and which, when minimised in one shot on AQC, would result in an optimal alignment. The desired form of the energy functional has been recently shown in the literature [29].

6.1. Particle Dynamics Based Alignment

Barnes-Hut Rigid Gravitational Approach (BHRGA) [29] is a recent point set alignment method with a single energy functional which remains unchanged during the entire optimisation. BHRGA is a globally multiply-linked approach, i.e., all \( y_m \) interact with all \( x_n \). In [29], point sets are aligned by minimising the mutual gravitational potential energy (GPE) \( E \) of the corresponding system of particles in the force field induced by \( \mathbf{X} \):

\[
E(\mathbf{R}, \mathbf{t}) = \sum_m \sum_n \mu_{y_m} \mu_{x_n} \| \mathbf{R} y_m + \mathbf{t} - \mathbf{x}_n \|_2, \quad (23)
\]

where \( \mu_{y_m} \) and \( \mu_{x_n} \) denote masses of \( y_m \) and \( x_n \), respectively. With no imposed boundary conditions, particles are initialised with unit masses. In [29], (23) is optimised with the Levenberg-Marquardt algorithm [41, 44], and the optimum is achieved when the system’s GPE is locally minimal. Without acceleration by a 2\(^D\)-tree, the method has quadratic complexity and (23) involves all possible interactions between the template and reference points.

We can now derive a QUBOP in the similar fashion as in Sec. 5 for the transformation estimation. Note, however, that the bases (13) and (22) allow for affine transformations and scaling. Thus, implicitly, we would optimise

\[
E(\mathbf{R}, \mathbf{t}, s) = \sum_m \sum_n \mu_{y_m} \mu_{x_n} \| \mathbf{R} y_m s + \mathbf{t} - \mathbf{x}_n \|_2, \quad (24)
\]

where the scalar \( s \) is the scaling of the template. As proven in [28], allowing for scale in globally multiply-linked point set alignment results in the shrinkage of the template to a single point with a very high probability. To remedy the
problem, either prior correspondences can be used, or point interactions can be restricted to local vicinities [28]. In our QA, we opt for the second solution which allows to use the rotational bases (13) and (22) elaborated in Sec. 5. Eventually, the \( \Phi \in \mathbb{R}^{(K+1)\times (L(1)+L(2)+\ldots+L(N))} \) matrix encoding point interactions for point set alignment reads

\[
\Phi = \begin{bmatrix} \Phi_1 \Phi_2 \ldots \Phi_N \end{bmatrix},
\]

with \( \Phi_n, n \in \{1, \ldots, N\} \), of the form

\[
\begin{bmatrix}
-x_n^T & -x_n^T & \ldots & -x_n^T \\
-[Q_1y_1]^T & -[Q_1y_2]^T & \ldots & -[Q_1y_{L(n)}]^T \\
-[Q_2y_1]^T & -[Q_2y_2]^T & \ldots & -[Q_2y_{L(n)}]^T \\
\vdots & \vdots & \ddots & \vdots \\
-[Q_Ky_1]^T & -[Q_Ky_2]^T & \ldots & -[Q_Ky_{L(n)}]^T \\
\end{bmatrix},
\]

(26)

with \( Q_k \) being as in (13) or (22) for the 2D and 3D case, respectively. \( \Phi_1, \Phi_2, \ldots, \Phi_N \) encode point interactions between every \( x_n \) and corresponding \( L(n) \ll M \) points of the template denoted by superscripted \( \{y_1^T, y_2^T, \ldots, y_{L(n)}^T\} \). Note that the latter build \( N \) subsets of \( \{y_1, y_2, \ldots, y_M\} \) of different cardinalities \( L(n), L \) on average. If different \( x_n \) interact with the same \( y_m \), the corresponding sub-columns \( [[Q_1y_m]^T [Q_2y_m]^T \ldots [Q_Ky_m]^T] \) of \( \Phi_n \) can be computed only once and reused. The final QUOBOP for point set alignment with \( \Phi_n \) as in (26) reads

\[
\arg \min_{q \in \mathbb{B}^{K+1}} q^T \Phi q.
\]

(27)

In total, \( K + 1 = 21 \) and \( K + 1 = 81 \) qubits are required to align point sets on AQC in the 2D and 3D case, respectively. Both transformation estimation and point set alignment need the same number of qubits in the same dimensions, and the difference lies in the complexity to construct \( P \) (see Sec. 6.2). Note that if the same template has to be aligned to multiple references, the corresponding \( \Phi \) can be obtained by reusing \( [[Q_1y_m]^T [Q_2y_m]^T \ldots [Q_Ky_m]^T] \) (which has to be computed only once). The first qubit of \( q \) has to be fixed to \( |1\rangle \), since the first element of every column contains a reference point which has to be active during the entire optimisation. The unembedding is performed similarly to the case of transformation estimation, see Sec. 5.

6.2. Complexity to Prepare \( P = \Phi \Phi^T \)

To prepare \( \Phi \), \( O(KDN\xi) \) and \( O(KDNL\xi) \) operations are required for the transformation estimation and point set alignment, respectively. \( \xi \) denotes the number of operations for multiplying \( y_m \) with one element of the additive basis \( Q_k \). To obtain the final \( P \), we need to transpose \( \Phi \) and multiply \( \Phi \) with \( \Phi \) which, in the worst case, takes \( O(K^2DN) \) operations for the transformation estimation and \( O(K^2DNL) \) operations for the point set alignment. There are also slightly faster algorithms for matrix multiplication compared to the naïve way [20].

| \( e_{2D} \) | 0.023 | 0.026 | 0.041 | 0.078 | 0.17 | 0.3 | 0.5 |
| \( e_{2D} \) | 0.012 | 0.013 | 0.012 | 0.012 | 0.012 | 0.013 |
| \( e_{R} \) | 0.058 | 0.062 | 0.083 | 0.22 | 0.47 | 0.764 |
| \( e_{R} \) | 0.041 | 0.044 | 0.041 | 0.036 | 0.031 | 0.03 |

Table 2: The accuracy of QA under random initial misalignments, for the transformation estimation ("TE") and point set alignment (\( K \geq 1 \)).

7. Experimental Evaluation

The current generation of D-WAVE annealers does not support the precision of weights in \( P \) necessary for our method [22]. It is foreseeable that future generations will enable a higher accuracy for couplings. We thus implement and test QA with an AQC sampler on a conventional computer (Intel i7-6700K CPU with 32GB RAM). All quantitative tests are performed with 21 binary variables corresponding to the size of the \( Q \) basis in 2D.

We report two error metrics, i.e., the alignment error \( e_{2D} \) and the transformation discrepancy \( e_R \), together with their standard deviations denoted by \( \sigma_{2D} \) and \( \sigma_R \), respectively. The alignment error \( e_{2D} = \frac{\|Ry_1 - x_1\|_{HS}}{\|x_1\|_{HS}} \) (\( \| \| \) denotes the Hilbert-Schmidt norm) measures how accurately the aligned shape coincides with the reference and requires ground truth correspondences. The transformation discrepancy is defined as \( e_R = \|I - RR^T\|_{HS} \), where \( R \) is the recovered rotation. It measures how closely the recovered transformation resembles a valid rigid transformation. The usage of two complementary metrics is necessary because a low \( e_R \) does not automatically imply an accurate registration. On the other hand, a low \( e_{2D} \) does not quantify how rigid the recovered transformation is.

Datasets and Proof of Concept. We use four 2D datasets, i.e., fish [47], qubit, kanji and composer with cardinalities varying from 91 (fish) to 7676 (composer), see Fig. 1 for qualitative registration results. For point sets with up to a few thousand points, the simulation time \( \tau_P < 1 \) sec. For \( \sim 7.7k \), \( \tau_P \) grows to 20.178 sec (by a factor of \( \sim 10^3 \)). Simulation with \( n = 30 \) takes already \( \sim 2.5 \) days. More binary variables allow for more elements in the basis \( Q \) resulting in more accurate alignment. Note that even with 80 qubits, i.e., for problems with \( n = 80 \), annealing on AQC takes around 100 ms. A simulation with \( n = 80 \) is not possible even on a conventional supercomputer in a reasonable time.

Initial Misalignment and Point Linking. We test how accurately our method recovers the transformation under the random angle of initial misalignment \( \theta \) and the different size of the point linking region. We generate 500 random transformations in the range \( \theta \in [0; 2\pi] \) of the fish dataset and resolve them with QA, for each \( K \in \{1, 10, 20, 30\} \). The results are summarised in Table 2. We see that \( e_{2D} \) corre-
lates with $\epsilon_R$ for all tested $K$. For $K = 30$ — which corresponds to one third of the template points — both metrics are still comparably low. We also study how the choice of the point interaction region or $K$ affects the accuracy of the transformation recovery and plot $e_{2D}$ and $e_R$ as the functions of $K$ for several angles of initial misalignment $\theta$ in Fig. 3-A. Interacting points are determined with the $K$ nearest neighbour rule for each $x_n$. Recall that according to the singularity theorem [28], the globally multiply-linked alignment (here, $K = 91$) results in a shrinkage of the template to a single point, which is observed experimentally.

Next, we systematically vary the angle of initial misalignment $\theta$ in the range $[0; 2\pi]$ with the angular step $\frac{\pi}{10}$ and report $e_{2D}$ and $e_R$ as the functions of $\theta$, for $K \in \{1, 10, 20, 30, 40, 50\}$. This test reveals the differences in the transformations caused by $\theta$, which arise due to the composition and the expressiveness of the chosen basis $M$, see Fig. 3-B. QA is almost agnostic to $\theta$, which is a desirable property of every point set alignment method.

**Sensitivity to Noise.** We systematically add uniformly distributed noise to the template and test the robustness of the proposed QA to outliers in the data, since real data often contains outliers. The highest template noise ratio amounts to $50\%$. Each metric for every noise ratio and every $K$ is averaged over 50 runs, see Fig. 3-C. $\sigma_R$ and $\sigma_{2D}$ do not exceed 0.057 and 0.03, respectively. We observe both the increasing alignment error and the discrepancy in the obtained transformations with the increasing noise level. For small $K$, nonetheless, even large noise ratios seem not to influence the metrics significantly.

**Spectral Gap Analysis.** Spectral gap $\Delta(\hat{\mathcal{H}})$ is the difference between the energy of the ground state and the second-lowest eigenstate. Each problem has an intrinsic and unique $\Delta(\hat{\mathcal{H}})$. Even though a rigorous analysis of the spectral gap is out of the scope of this paper, we make several qualitative observations about the energy landscape of QA, the difference in the energy values and the corresponding registrations for one exemplary problem. In Fig. 4, we plot the sequences of energy-decreasing transitions together with the energy values in the experiment with fish, for three $\theta$ values.

We notice that some solutions have very small differences in the energies and are qualitatively indistinguishable from each other. This is accounted for by the choice of the additive basis, i.e., that the same alignment can be encoded in different ways. In contrast, we see significant differences in the energy values of the qualitatively different solutions (orders of magnitudes larger in the analysed experiment).

We conclude that even though $\Delta(\hat{\mathcal{H}})$ is small, the alignments corresponding to several few lowest eigenstates are qualitatively similar. This suggests that our selection of the basis leads to problems with sufficient spectral gaps.

### 8. Conclusions

This paper introduces AQC for the computer vision community and shows that fundamental low-level problems can be brought to a representation suitable for solving on AQC. In simulations on a classical computer and in a wide range of scenarios, our QA is shown to successfully recover 2D transformations which are close approximations of globally optimal transformations. With the chosen basis of 20 elements, the solutions result in low transformation discrepancy and alignment errors. Observations on how to avoid singularities as well as the noise sensitivity and spectral gap analysis complement the experimental section.

In future work, our technique can be extended to affine transformations and other related computer vision problems. We hope to see more research on computer vision methods with quantum hardware in the next decades.

**Acknowledgements.** This work was supported by the ERC Consolidator Grant 770784. VG is grateful to Polina Matveeva for many enlightening discussions on the physical foundations of adiabatic quantum computing. The authors thank Bertram Taetz and Hanno Ackermann for reviewing an earlier version of this paper.
A. Appendix

In this additional section, we provide details on the selection of the annealing rate, analyse the structure of $P$ and formalise the unembedding, i.e., conversion of the solution to QUBOP (16) to the solution of the original alignment problem on point sets. We preserve the notations referring to the sections and equations from the main matter. The equations and the figure introduced in this supplement are equipped with Roman numerals.

**Annealing Rate.** Suppose $E_n(s)$ is the ground state of instantaneous Hamiltonian, $E_n(0)$ is the initial state (ground state) of the system and $E_m(s)$ is any other excited state of the instantaneous Hamiltonian. Let $s = \frac{T}{r} \in [0; 1]$, where $T$ is the overall time of interpolation and $t$ is physical time. Then, according to [4], $T$ has to be chosen so that

$$T \gg \frac{|\langle E_{m}(s) | dH/ds | E_{n}(s) \rangle|}{E_{nm}(s)^2}, \forall m \neq n,$$

where $dH/ds$ is the rate of change of Hamiltonian with respect to $s$ and $E_{nm}$ is the difference in the corresponding instantaneous energies.

**Analysis of $P$.** Fig. 1 visualises several exemplary weight matrices $P$ from the experiments with clean and noisy data (see Sec. 7). There are several observations. First, $P = \Phi \Phi^T$ is symmetric upon algorithm design. We also see that the columns of $\Phi$ can be arbitrarily reshuffled as long as the correspondences are preserved\(^4\). Second, $P$ contains regularly arranged zero submatrices, due to our choice of the basis. As soon as a row of $\Phi$ induced by $qC_1$, where $C_1 \in \{I, -I\}$, is multiplied by a column of $\Phi^T$ induced by $qC_M$, where $C_M \in \{M, -M\}$, and vice versa, we obtain a zero entry in $P$. The reason is that

$$\begin{cases}
[I \sum_i y_i]^T [M \sum_j y_j] = 0 \\
[-I \sum_i y_i]^T [M \sum_j y_j] = 0
\end{cases}$$

if $\sum_i y_i = \sum_j y_j$, which holds in our case since each row of $\Phi$ except the first row includes all points of $Y$ multiplied by a single basis element $Q_k$ (see Fig. 1-(top left) for $C$ pairs resulting in zero matrices). Third, the structure of $P$ reflects that its diagonal entries encode biases, and non-diagonal elements represent couplings between the qubits.

With the increasing $K$, the span of the absolute energy values increases, due to the higher number of point interactions. As expected, $P$ depends on data and the angle of initial misalignment between the point sets. For all possible inputs and initial conditions — point sets of different cardinalities, $K$ and $\theta$ — the structure of $P$ is the same for the chosen basis. From $P$, we also recognise that the considered alignment problem is not purely combinatorial and requires high-precision weights $J_{j,k}$ in (5).

**Unembedding.** Unembedding is the decoding of the solution to QUBOP (16) to the solution of the original alignment problem. Upon the design, our QA method assembles the entries of the transformation matrix in the additive basis $Q_k$ (see Secs. (5.1)–(6.1)). Suppose $\hat{q}$ is the measurement result of $q$, i.e., it is a classical bitstring with $K + 1$ elements. Recall that $q_1$ is reserved for reference points and does not contribute to the assembly of the transformation. Once $\hat{q}$ is measured and returned, we obtain the corresponding transformation $R$ by summing up $Q_k$ multiplied by $q_{k+1}$:

$$R = \sum_k \hat{q}_{k+1} Q_k.$$

The obtained $R$ is an affine transformation. If the solution has to represent a valid rotation matrix $R_\theta$, $R$ can be projected to the rotation group by solving the closest orthogonal approximation problem with constraints:

$$\min \|R_t - R\|_{HS}^2,$$

s. t. $R_t^{-1} = R_t^T$ and $\det(R_t) = 1$.

For a solution to (iv) by singular value decomposition, see [33].

\(^4\) a reshuffling of rows requires changing the order of elements in $Q$.
References


[28] Vladislav Golyanik and Christian Theobalt. Optimising for scale in globally multiply-linked gravitational point set registration leads to singularities. In International Conference on 3D Vision (3DV), 2019. 6, 7, 8


[47] Andriy Myronenko and Xubo Song. Point-set registration: Coherent point drift. Transactions on Pattern Analysis and Machine Intelligence (TPAMI), 2010. 1, 2, 7