

Article

A Quantum Adiabatic Algorithm for Multiobjective Combinatorial Optimization [†]

Benjamín Barán and Marcos Villagra * 

Núcleo de Investigación y Desarrollo Tecnológico, Universidad Nacional de Asunción,
San Lorenzo C.P. 2619, Paraguay; bbaran@cba.com.py

* Correspondence: mvillagra@pol.una.py

[†] This Paper is an Extended Version of Our Paper Pulished in Proceedings of the 42nd Latin American Conference on Informatics (CLEI), Valparaíso, Chile, 10–14 October 2016.

Received: 14 November 2018; Accepted: 1 March 2019; Published: 9 March 2019



Abstract: In this work we show how to use a quantum adiabatic algorithm to solve multiobjective optimization problems. For the first time, we demonstrate a theorem proving that the quantum adiabatic algorithm can find Pareto-optimal solutions in finite-time, provided some restrictions to the problem are met. A numerical example illustrates an application of the theorem to a well-known problem in multiobjective optimization. This result opens the door to solve multiobjective optimization problems using current technology based on quantum annealing.

Keywords: quantum computation; multiobjective optimization; quantum adiabatic evolution

1. Introduction

Currently, quantum computation has many practical applications in engineering and computer science like machine learning, bioinformatics and artificial intelligence [1]. At the core of all these applications, there is an optimization procedure, and the quantum adiabatic computing paradigm of Farhi et al. [2] is the best method known thus far for optimization problems.

In this work, we show how to use a quantum adiabatic algorithm in multiobjective combinatorial optimization problems or MCO. An optimization problem is said to be “multiobjective” or “multicriteria” if there are two or more objective functions involved [3]. When these objective functions are required to be optimized at the same time, the optimality of solutions must be revised and, thus, an MCO can have a set of so-called “Pareto-optimal solutions” where no solution is better than any other. Here we show that the quantum adiabatic algorithm can find Pareto-optimal solutions in finite time, provided certain restrictions are met. In Theorem 2 we identify two structural features that a given MCO must satisfy in order to make an effective use of the quantum adiabatic algorithm presented in this work.

Even though most known quantum adiabatic optimization algorithms consider single-objective optimization problems (see [1,4]), only a few works discuss quantum algorithms for MCOs. Those few algorithms make use of Grover’s search method [5,6] as a subroutine that is invoked inside a classical algorithm. Alanis et al. [7] presented a quantum optimization algorithm in the context of routing problems. In [8], a general algorithm for MCO was presented and experimentally compared against a state-of-the-art metaheuristic. Both papers [7,8] use Grover’s search algorithm to solve an MCO; however, Grover’s algorithm is not naturally constructed for optimization problems. Having Grover’s algorithm as the main subroutine for optimization gives place to an “ad hoc” heuristic method whose finite time convergence has not yet been proved. Hence, these previous works [7,8] relied on numerical experiments instead of rigorous proofs.

This work presents the first quantum algorithm for MCO that guarantees finite time converge to a Pareto-optimal solution. Furthermore, since our method is constructed on the quantum adiabatic paradigm, it can be implemented in current technologies based on quantum annealing [1,9]. An extended abstract of this work appeared in [10].

The outline of this paper is as follows. In Sections 2 and 3 we present the main concepts of MCO and adiabatic quantum computing that are relevant for this paper. In Section 4 we formally state our main result and present a full proof. In Section 5 we present a small numerical example of an MCO. Finally, in Section 6 we conclude this paper.

2. Preliminaries on Multiobjective Combinatorial Optimization

We present here the notation used throughout this paper. We also briefly review the main concepts of multiobjective optimization, which are standard definitions and can be found in several papers in the literature, for example [3]. Definitions 3, 4, 5 and 9 and Lemmas 1 and 3 are originals of this work.

The set of natural numbers (including 0) is denoted \mathbb{N} , the set of integers is \mathbb{Z} , the set of real numbers is denoted \mathbb{R} and the set of positive real numbers is \mathbb{R}^+ . For any $i, j \in \mathbb{Z}$, with $i < j$, we let $[i, j]_{\mathbb{Z}}$ denote the discrete interval $\{i, i+1, \dots, j-1, j\}$. The set of binary words of length n is denoted $\{0, 1\}^n$. We also let $\text{poly}(n) = O(n^c)$ be a polynomial in n , i.e., $c \in \mathbb{N}$.

A multiobjective combinatorial optimization problem (MCO) is an optimization problem involving multiple objectives over a finite set of feasible solutions. These objectives typically present trade-offs among solutions and in general, there is no single optimal solution but a set of compromise solutions known as the Pareto set [3]. In this work, we follow the definition of Kung et al. [11]. Furthermore, with no loss of generality, all optimization problems considered in this work are minimization problems.

Let S_1, \dots, S_d be totally ordered sets and let \leq_i be an order on set S_i for each $i \in [1, d]_{\mathbb{Z}}$. We also let n_i be the cardinality of S_i . Define the natural partial order relation \prec over the cartesian product $S_1 \times \dots \times S_d$ in the following way. For any $u = (u_1, \dots, u_d)$ and $v = (v_1, \dots, v_d)$ in $S_1 \times \dots \times S_d$, we write $u \prec v$ if and only if for any $i \in [1, d]_{\mathbb{Z}}$ it holds that $u_i \leq_i v_i$; otherwise we write $u \not\prec v$. An element $u \in S$ is a minimal element if there is no $v \in S$ such that $v \prec u$ and $v \neq u$. Moreover, we say that u is non-comparable with v if $u \not\prec v$ and $v \not\prec u$ and succinctly write $u \sim v$. In the context of multiobjective optimization, the relation \prec as defined here is often referred to as the Pareto-order relation [11].

Definition 1. An MCO is defined as a tuple $\Pi = (D, R, d, \mathcal{F}, \prec)$ where D is a finite set called domain, $R \subseteq \mathbb{R}^+ \cup \{0\}$ is a set of real values, d is a positive integer, \mathcal{F} is a finite collection of functions $\{f_i\}_{i \in [1, d]_{\mathbb{Z}}}$ where each f_i maps from D to R , and \prec is the Pareto-order relation on R^d (here R^d is the d -fold cartesian product on R). We also define a function f that maps D to R^d as $f(x) = (f_1(x), \dots, f_d(x))$ referred to as the objective vector of Π . If $f(x)$ is a minimal element of R^d , we say that x is a Pareto-optimal solution of Π . The set of all Pareto-optimal solutions of Π is denoted $P(\Pi)$.

Definition 2. For any two elements $x, y \in D$, if $f(x) \prec f(y)$ we write $x \prec y$; similarly, if $f(x) \sim f(y)$ we write $x \sim y$. For any $x, y \in D$, if $x \prec y$ and $y \prec x$ we say that x and y are equivalent and it is denoted as $x \equiv y$.

A typical example of a multiobjective optimization problem is the two-parabolas problem of Figure 1. In this problem we have two continuous objective functions defined by two parabolas that intersect in a single point.

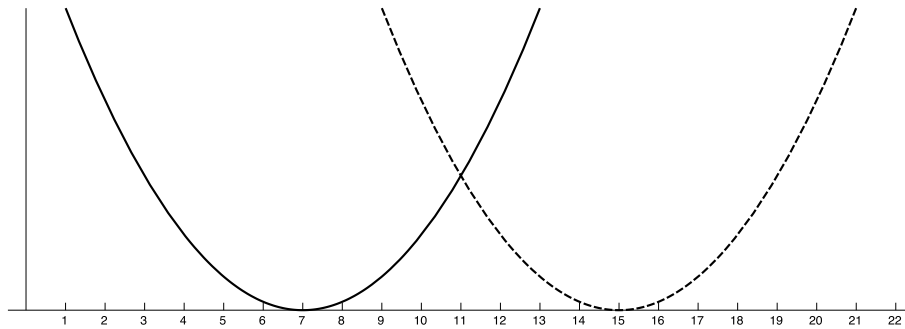


Figure 1. The two-parabolas Problem. The first objective function $f_1(x) = (x - 7)^2$ is represented by the bold line and the second objective function $f_2(x) = (x - 15)^2$ by the dashed line. Note that there are no equivalent elements in the domain. In this particular example, all the solutions between seven and 15 are Pareto-optimal.

The set of Pareto-optimal solutions can be very large, and therefore, most methods for MCOs are concerned with finding a subset of the Pareto-optimal solutions, or an approximation. Optimal query algorithms were discovered by Kung et al. [11] where all Pareto-optimal solutions can be found for $d = 2, 3$ and proved almost tight upper and lower bounds for any $d \geq 4$ up to polylogarithmic factors. Papadimitriou and Yannakakis [12] initiated the field of approximation algorithms for MCOs where an approximation to all Pareto-optimal solutions can be found in polynomial time.

For the remainder of this work, \prec will always be the Pareto-order relation and will be omitted from the definition of any MCO. Furthermore, for convenience, we will often write $\Pi_d = (D, R, \mathcal{F})$ as a short-hand for $\Pi = (D, R, d, \mathcal{F})$. In addition, we will assume that each function $f_i \in \mathcal{F}$ is polynomial-time computable and each $f_i(x)$ is bounded by a polynomial in the number of bits of x . This is a typical assumption in the theory of optimization algorithms in order to have the computational complexity of any optimization algorithm to depend only on the number of “black-box” accesses to the objective function.

Definition 3. Given any MCO Π_d we say that Π_d is well-formed if and only if for each $f_i \in \mathcal{F}$ there is a unique $x \in D$ such that $f_i(x) = 0$.

Definition 4. An MCO Π_d is normal if and only if Π_d is well-formed and $f_i(x) = 0$ and $f_j(y) = 0$, for $i \neq j$, implies $x \neq y$.

In a normal MCO, the value of an optimal solution in each f_i is 0, and all optimal solutions are different. In Figure 1, solutions $x = 7$ and $x = 15$ are optimal solutions of f_1 and f_2 with value 0, respectively; hence, the two-parabolas problem of Figure 1 is normal.

Definition 5. An MCO Π_d is collision-free if given $\lambda = (\lambda_1, \dots, \lambda_d)$, with each $\lambda_i \in \mathbb{R}^+$, for any $i \in [1, d]_{\mathbb{Z}}$ and any pair $x, y \in D$, with $x \neq y$, it holds that $|f_i(x) - f_i(y)| \geq \lambda_i$. If Π_d is collision-free we write succinctly as Π_d^λ .

The two-parabolas problem of Figure 1 is not collision-free; for example, for solutions $x = 5$ and $x = 9$ we have that $f_1(5) = f_1(9)$.

Definition 6. A Pareto-optimal solution x is trivial if x is an optimal solution of some $f_i \in \mathcal{F}$.

In Figure 1, solutions $x = 7$ and $x = 15$ are trivial Pareto-optimal solutions, whereas any x between seven and 15 is non-trivial.

Lemma 1. For any normal MCO Π_d , if x and y are trivial Pareto-optimal solutions of Π_d , then x and y are not equivalent.

Proof. Let x, y be two trivial Pareto-optimal solutions of Π_d . There exists i, j such that $f_i(x) = 0$ and $f_j(y) = 0$. Since Π_d is normal we have that $x \neq y$ and $f_i(y) > 0$ and $f_j(x) > 0$, hence, $x \sim y$ and they are not equivalent. \square

Let W_d be a set of normalized vectors in $[0, 1]^d$, the continuous interval between zero and less than one, defined as

$$W_d = \left\{ w = (w_1, \dots, w_d) \in [0, 1]^d \mid \sum_{i=1}^d w_i = 1 \right\}. \quad (1)$$

For any $w \in W_d$, define $\langle f(x), w \rangle = \langle f(x), w \rangle = w_1 f_1(x) + \dots + w_d f_d(x)$. In the literature of multiobjective optimization, if each f_i is an objective function of an MCO Π_d , it is said that $\langle f(x), w \rangle$ is a linearization or scalarization of Π_d [13].

The following fact is a well-known property of MCOs.

Lemma 2. Let $\Pi_d = (D, R, \mathcal{F})$. For any $w \in W_d$ there exists $x \in D$ such that if $\langle f(x), w \rangle = \min_{y \in D} \{ \langle f(y), w \rangle \}$, then x is a Pareto-optimal solution of Π_d .

Proof. Fix $w \in W_d$ and let $x \in D$ be such that $\langle f(x), w \rangle$ is minimum among all elements of D . For any $y \in D$, with $y \neq x$, we need to consider two cases: (1) $\langle f(y), w \rangle = \langle f(x), w \rangle$ and (2) $\langle f(y), w \rangle > \langle f(x), w \rangle$.

Case (1). Here we have another two subcases, either $f_i(y) = f_i(x)$ for all i or there exists at least one pair $i, j \in \{1, \dots, d\}$ such that $w_i f_i(x) < w_i f_i(y)$ and $w_j f_j(y) < w_j f_j(x)$. When $f_i(x) = f_i(y)$ for each $i = 1, \dots, d$ we have that x and y are equivalent. On the contrary, if $w_i f_i(x) < w_i f_i(y)$ and $w_j f_j(y) < w_j f_j(x)$, we have that $f_i(x) < f_i(y)$ and $f_j(y) < f_j(x)$, and hence, $x \sim y$.

Case (2). In this case, there exists $i \in \{1, \dots, d\}$ such that $w_i f_i(x) < w_i f_i(y)$, and hence, $f_i(x) < f_i(y)$. Thus, $f(y) \not\preceq f(x)$ and $y \not\preceq x$ for any $y \neq x$.

We conclude from Case (1) that $x \equiv y$ or $x \sim y$, and from Case (2) that $y \not\preceq x$. Therefore, x is Pareto-optimal. \square

Given any linearization of an MCO, by Lemma 2, an optimal solution of a linearized MCO corresponds to a Pareto-optimal solution; however, it does not hold in general that each Pareto-optimal solution has a corresponding linearization, i.e., no all Pareto-optimal solutions can always be found using linearizations.

Lemma 3. Given $\Pi_d = (D, R, \mathcal{F})$, any two elements $x, y \in D$ are equivalent if and only if for all $w \in W_d$ it holds that $\langle f(x), w \rangle = \langle f(y), w \rangle$.

Proof. Assume that $x \equiv y$. Hence $f(x) = f(y)$. If we pick any $w \in W_d$ we have that

$$\langle f(x), w \rangle = w_1 f_1(x) + \dots + w_d f_d(x) = w_1 f_1(y) + \dots + w_d f_d(y) = \langle f(y), w \rangle.$$

Now suppose that for all $w \in W_d$ it holds $\langle f(x), w \rangle = \langle f(y), w \rangle$. By contradiction, assume that $x \not\equiv y$. With no loss of generality, assume further that there is exactly one $i \in [1, d]_{\mathbb{Z}}$ such that $f_i(x) \neq f_i(y)$. Hence

$$w_i(f_i(x) - f_i(y)) = \sum_{j \neq i} w_j(f_j(y) - f_j(x)). \quad (2)$$

The right hand of Equation (2) is 0 because for all $j \neq i$ we have that $f_j(x) = f_j(y)$. The left hand of Equation (2), however, is not 0 by our assumption, hence, a contradiction. Therefore, x is equivalent to y . \square

In this work, we are only interested in finding non-trivial Pareto-optimal solutions. Finding trivial elements can be done by optimizing each f_i independently; consequently, in Equation (1) we do not allow for any w_i to be 1.

Definition 7. The set of supported Pareto-optimal solutions denoted $S(\Pi)$ is defined as the set of Pareto-optimal solutions x where $\langle f(x), w \rangle$ is optimal for some $w \in W_d$.

From Lemma 2, we know that some Pareto-optimal solutions cannot be found using any linearization $w \in W_d$.

Definition 8. The set of non-supported Pareto-optimal solutions is the set $N(\Pi) = P(\Pi) \setminus S(\Pi)$.

Note that there may be Pareto-optimal solutions x and y that are non-comparable and $\langle f(x), w \rangle = \langle f(y), w \rangle$ for some $w \in W_d$. That is equivalent to saying that the objective function obtained from a linearization of an MCO is not one-to-one (injective).

Definition 9. Any two elements $x, y \in D$ are weakly-equivalent if and only if there exists $w \in W_d$ such that $\langle f(x), w \rangle = \langle f(y), w \rangle$.

By Lemma 3, any two equivalent solutions x, y are also weakly-equivalent; on the other hand, if x and y are weakly-equivalent it does not imply that they are equivalent. For example, consider two objective vectors $f(x) = (1, 2, 3)$ and $f(y) = (1, 3, 2)$. Clearly, x and y are not equivalent; however, if $w = (1/3, 1/3, 1/3)$ we can see that x and y are indeed weakly-equivalent. In Figure 1, points $x = 10$ and $x = 12$ are weakly-equivalent.

3. The Quantum Adiabatic Algorithm

The quantum adiabatic computing paradigm was discovered by Farhi et al. [2] and the quantum adiabatic algorithm was designed to solve single-objective optimization problems. The main idea behind quantum adiabatic algorithms is that optimization problems are somehow encoded into time-dependent Hamiltonians. Then, we start the algorithm with an easy-to-prepare quantum state and we let the system evolve according to the adiabatic theorem. After some time, we measure the system and obtain an optimal solution to our optimization problem. In this work, we follow the definition of McGeoch [1].

Given an objective function f where each element in its domain can be represented with $\text{poly}(n)$ bits, a quantum adiabatic algorithm for f is constructed from three components:

1. an initial Hamiltonian H_0 chosen in such a way that its ground state is easy to prepare;
2. a final Hamiltonian H_1 encodes the function f in such a way that the minimum eigenvalue of H_1 corresponds to $f(x)$ and its ground-state corresponds to x ;
3. an adiabatic evolution path, that is, a function $s(t)$ that decreases from 1 to 0 as the time t goes from 0 to a given time T . In this work, we will always use a linear path $s(t) = 1 - t/T$.

The time-dependent Hamiltonian H for the algorithm is thus defined as

$$H(t) = s(t)H_0 + (1 - s(t))H_1. \quad (3)$$

If $|\psi(t)\rangle$ is an eigenvector of $H(t)$ that corresponds to the minimum eigenvalue, the quantum adiabatic algorithm works as follows. Prepare the system in state $|\psi(0)\rangle$ and let it evolve according to the Schrödinger's equation. After some time T , measure the state $|\psi(T)\rangle$ with respect to some

well-defined basis. The adiabatic theorem says that if T is large enough and H is non-degenerate in its minimum eigenvalue, the quantum state that is observed from the measurement is very close to $|\psi(T)\rangle$.

One of the most recent versions of the adiabatic algorithm is due to Ambainis and Regev [14] where a lower bound in the value of T can be estimated. Below we transcribe the complete statement of their theorem since it is one of the main pieces in our proof. Let $H(s)$ be a time-dependent Hamiltonian with a linear path $s = 1 - t/T$ for some given T . Let $\|H\| = \max_{s \in [0,1]} \|H(s)\|$ where $\|\cdot\|$ is the operator norm with respect to the ℓ_2 -norm, and let H' and H'' be the first and second derivatives of H , respectively.

Theorem 1 (Ambainis and Regev [14]). *Let $H(s)$, $0 \leq s \leq 1$, be a time-dependent Hamiltonian, let $|\psi(s)\rangle$ be one of its eigenstates, and let $\gamma(s)$ be the corresponding eigenvalue. Assume that for any $s \in [0, 1]$, all other eigenvalues of $H(s)$ are either smaller than $\gamma(s) - \lambda$ or larger than $\gamma(s) + \lambda$. Consider the adiabatic evolution given by H and $|\psi(s)\rangle$ applied for time T . Then, the following condition is enough to guarantee that the final state is at distance at most δ from $|\psi(1)\rangle$:*

$$T \geq \frac{10^5}{\delta^2} \cdot \max \left\{ \frac{\|H\|^3}{\lambda^4}, \frac{\|H'\| \cdot \|H''\|}{\lambda^3} \right\}.$$

One of the main design decisions for adiabatic algorithms is choosing an adequate initial Hamiltonian. To make use of the adiabatic theorem, we need to construct an initial Hamiltonian that does not commute with the final Hamiltonian; otherwise, the eigenvalue gap will disappear [2].

4. Main Result of This Work

With all the technical definitions established, now we are ready to formally state our main result. For any vector w in Euclidian space we define the ℓ_1 -norm of w as $\|w\|_1 = |w_1| + \dots + |w_d|$.

Theorem 2. *Let Π_d^λ be any normal and collision-free MCO. If there are no equivalent Pareto-optimal solutions, then for any $w \in W_d$ there exists $w' \in W_d$ and a Hamiltonian $H_{w'}$, satisfying $\|w - w'\|_1 \leq 1/\text{poly}(n)$, such that the quantum adiabatic algorithm, using $H_{w'}$ as final Hamiltonian, can find a Pareto-optimal solution x corresponding to w in finite time.*

In the following subsections we present a proof of Theorem 2. In Section 4.1 we show how to construct the initial and final Hamiltonians, and in Section 4.2 we show that our construction is correct.

4.1. The Initial and Final Hamiltonians for MCOs

In this section we show how to construct the initial and final Hamiltonians. Given any normal and collision-free MCO $\Pi_d^\lambda = (D, R, \mathcal{F})$, assume with no loss of generality that the domain of the objective functions is $D = \{0, 1\}^n$, that is, the set of all bit strings of length n .

For each $i \in [1, d]_{\mathbb{Z}}$ we define a Hamiltonian $H_{f_i} = \sum_{x \in \{0,1\}^n} f(x) |x\rangle \langle x|$. Since Π_d^λ is collision-free and normal, each H_{f_i} is nondegenerate in all its eigenvalues and its minimum eigenvalue is 0. Given a linearization w of Π_d^λ , we construct the final Hamiltonian H_w for our quantum algorithm as

$$\begin{aligned} H_w &= w_1 H_{f_1} + \dots + w_d H_{f_d} \\ &= \sum_{x \in \{0,1\}^n} \langle f(x), w \rangle |x\rangle \langle x|. \end{aligned} \quad (4)$$

Now we construct the initial Hamiltonian, which cannot commute with the final Hamiltonian recently defined. In this work we make use of the Hamiltonian defined in [2]. Let $|\hat{0}\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$ and $|\hat{1}\rangle = (|0\rangle - |1\rangle)/\sqrt{2}$. The operation F that makes $|\hat{0}\rangle = F|0\rangle$ and $|\hat{1}\rangle = F|1\rangle$ is called the

Walsh-Hadamard transform. Thus, for any $x \in \{0,1\}^n$ we can make $|\hat{x}\rangle = F^{\otimes n}|x\rangle$, where F is the n -fold Walsh-Hadamard transform. The initial Hamiltonian is

$$H_0 = \sum_{x \in \{0,1\}^n} h(x) |\hat{x}\rangle \langle \hat{x}|, \quad (5)$$

where $h(0^n) = 0$ and $h(x) = 1$ for any $x \neq 0^n$.

Now that we have defined our initial and final Hamiltonians we need to show that the interpolating hamiltonian $H(t)$ of Equation (3) is indeed nondegenerate in all its eigenvalues and that it fulfils the requirements of Theorem 2.

4.2. Analysis of the Final Hamiltonian

Note that if the initial Hamiltonian does not commute with the final Hamiltonian, it suffices to prove that the final Hamiltonian is nondegenerate in its minimum eigenvalue [2]. For the remaining of this work, we let σ_w and α_w be the smallest and second smallest eigenvalues of H_w corresponding to a normal and collision-free MCO $\Pi_d^\lambda = (D, R, \mathcal{F})$.

Lemma 4. Let x be a non-trivial Pareto-optimal solution of Π_d^λ . For any $w \in W_d$ it holds that $\sigma_w > \langle \lambda, w \rangle$.

Proof. Let $\sigma_w = w_1 f_1(x) + \dots + w_d f_d(x)$ and let x be a non-trivial Pareto-optimal element. For each $w_i \in N$ we have that

$$\sigma_w = \sum_i w_i f_i(x) > \sum_i w_i \lambda_i = \langle \lambda, w \rangle.$$

□

Lemma 5. For any $w \in W_d$, let H_w be a Hamiltonian with a nondegenerate minimum eigenvalue. The eigenvalue gap between the smallest and second smallest eigenvalues of H_w is at least $\langle \lambda, w \rangle$.

Proof. Let σ_w be the unique minimum eigenvalue of H_w . We have that $\sigma_w = \langle f(x), w \rangle$ for some $x \in \{0,1\}^n$. Now let $\alpha_w = \langle f(y), w \rangle$ be a second smallest eigenvalue of H_w for some $y \in \{0,1\}^n$ where $y \neq x$. Hence,

$$\begin{aligned} \alpha_w - \sigma_w &= \langle f(y), w \rangle - \langle f(x), w \rangle \\ &= w_1 f_1(y) - w_1 f_1(x) + w_2 f_2(y) - w_2 f_2(x) \\ &\geq w_1 \lambda_1 + w_2 \lambda_2 \\ &= \langle \lambda, w \rangle. \end{aligned}$$

□

Lemma 6. If there are no weakly-equivalent Pareto-optimal solutions in Π_d^λ , then the Hamiltonian H_w is non-degenerate in its minimum eigenvalue.

Proof. By the contrapositive, suppose H_w is degenerate in its minimum eigenvalue σ_w . Take any two degenerate minimal eigenstates $|x\rangle$ and $|y\rangle$, with $x \neq y$, such that

$$w_1 f_1(x) + \dots + w_d f_d(x) = w_1 f_1(y) + \dots + w_d f_d(y) = \sigma_w.$$

Then it holds that x and y are weakly-equivalent. □

We further show that even if Π_d^λ has weakly-equivalent Pareto-optimal solutions, we can find a nondegenerate Hamiltonian. Let $m = \max_{x,i} \{f_i(x)\}$.

Lemma 7. For any Π_d^λ , let $x_1, \dots, x_\ell \in D$ be Pareto-optimal solutions that are not pairwise equivalent. If there exists $w \in W_d$ and $\sigma_w \in \mathbb{R}^+$ such that $\langle f(x_1), w \rangle = \dots = \langle f(x_\ell), w \rangle = \sigma_w$ is minimum among all $y \in D$, then there exists $w' \in W_d$ and $i \in [1, \ell]_{\mathbb{Z}}$ such that for all $j \in [1, \ell]_{\mathbb{Z}}$, with $j \neq i$, it holds $\langle f(x_i), w' \rangle < \langle f(x_j), w' \rangle$. Additionally, if the linearization w' satisfies $\|w - w'\|_1 \leq \frac{\langle \lambda, w \rangle}{md}$, then $\langle f(x_i), w' \rangle$ is unique and minimum among all $\langle f(y), w' \rangle$ for $y \in D$.

Proof. We prove the lemma by induction on ℓ . Let $\ell = 2$, then $\langle f(x_1), w \rangle = \langle f(x_2), w \rangle$, and hence,

$$\begin{aligned} w_1 f_1(x_1) + \dots + w_d f_d(x_1) &= \sigma_w \\ w_1 f_1(x_2) + \dots + w_d f_d(x_2) &= \sigma_w. \end{aligned} \quad (6)$$

for some $\sigma_w \in \mathbb{R}^+$. From linear algebra we know that there is an infinite number of elements of W_d that simultaneously satisfy Equation (6). With no loss of generality choose adequately f_1 and f_2 , fix w_3, \dots, w_d and set $b_1 = w_3 f_3(x_1) + \dots + w_d f_d(x_1)$ and $b_2 = w_3 f_3(x_2) + \dots + w_d f_d(x_2)$. We have that

$$\begin{aligned} w_1 f_1(x_1) + w_2 f_2(x_1) &= \sigma_w - b_1 \\ w_1 f_1(x_2) + w_2 f_2(x_2) &= \sigma_w - b_2. \end{aligned} \quad (7)$$

Again, by linear algebra, we know that Equation (7) has a unique solution w_1 and w_2 ; it suffices to note that the determinant of the coefficient matrix of Equation (7) is not 0, as it is proved in Appendix A.

Choose any $w'_1 \neq w_1$ and $w'_2 \neq w_2$ satisfying $w'_1 + w'_2 + w_3 + \dots + w_d = 1$ and let $w' = (w'_1, w'_2, w_3, \dots, w_d)$. Then we have that $\langle f(x_1), w' \rangle \neq \langle f(x_2), w' \rangle$ because w'_1 and w'_2 are not solutions to Equation (7). Hence, either $\langle f(x_1), w' \rangle$ or $\langle f(x_2), w' \rangle$ must be smaller than the other.

Suppose that $\langle f(x_1), w' \rangle < \langle f(x_2), w' \rangle$. We now claim that $\langle f(x_1), w' \rangle$ is minimum and unique among all $y \in D$. In addition to the constraint of the preceding paragraph that w' must satisfy, in order for $\langle f(x_1), w' \rangle$ to be minimum, we must choose w' such that $\|w - w'\|_1 \leq \frac{\langle \lambda, w \rangle}{md}$.

Assume for the sake of contradiction the existence of $y \in D$ such that $\langle f(y), w' \rangle \leq \langle f(x_1), w' \rangle$. Hence,

$$\langle f(y), w' \rangle \leq \langle f(x_1), w' \rangle < \langle f(y), w \rangle.$$

From Lemma 4, we know that $|\langle f(x_1), w \rangle - \langle f(y), w \rangle| > \langle \lambda, w \rangle$, and thus,

$$|\langle f(y), w' \rangle - \langle f(y), w \rangle| > \langle \lambda, w \rangle. \quad (8)$$

Using the Cauchy–Schwarz inequality we have that

$$\begin{aligned} |\langle f(y), w' \rangle - \langle f(y), w \rangle| &= |\langle f(y), w' - w \rangle| \\ &\leq \|f(y)\|_1 \cdot \|w' - w\|_1 \\ &\leq \langle \lambda, w \rangle, \end{aligned}$$

where the last line follows from $\|f(y)\|_1 \leq md$ and $\|w - w'\|_1 \leq \frac{\langle \lambda, w \rangle}{md}$; from Equation (8), however, we have that $|\langle f(y), w' - w \rangle| > \langle \lambda, w \rangle$, which is a contradiction. Therefore, we conclude that $\langle f(y), w' \rangle > \langle f(x_1), w' \rangle$ for any $y \in D$; the case for $\langle f(x_1), w' \rangle > \langle f(x_2), w' \rangle$ can be proved similarly. The base case of the induction is thus proved.

Now suppose the statement holds for ℓ . Let $x_1, \dots, x_\ell, x_{\ell+1}$ be Pareto-optimal solutions that are not pairwise equivalent. Let $w \in W_d$ be such that $\langle f(x_1), w \rangle = \dots = \langle f(x_{\ell+1}), w \rangle$ holds. By our induction hypothesis, there exists $w' \in W_d$ and $i \in [1, \ell]_{\mathbb{Z}}$ such that $\langle f(x_i), w' \rangle < \langle f(y), w' \rangle$ for any other $y \in D$.

If $\langle f(x_i), w' \rangle \neq \langle f(x_{\ell+1}), w' \rangle$ then we are done, because either one must be smaller. Suppose, however, that $\langle f(x_{\ell+1}), w' \rangle = \langle f(x_i), w' \rangle = \sigma_{w'}$ for some $\sigma_{w'} \in \mathbb{R}^+$. From the base case of the

induction we know there exists $w'' \neq w'$ that makes $\langle f(x_i), w'' \rangle < \langle f(x_{\ell+1}), w'' \rangle$, and hence, $\langle f(x_i), w'' \rangle < \langle f(y), w'' \rangle$ for any $y \in D$. Therefore, the lemma is proved. \square

The premise in Lemma 7, that each x_1, \dots, x_ℓ must be Pareto-optimal solutions, is a sufficient condition because if one solution is not Pareto-optimal, then the statement will contradict Lemma 2.

We now apply Lemma 7 to find a Hamiltonian with a nondegenerate minimum eigenvalue.

Lemma 8. Let Π_d^λ be a MCO with no equivalent Pareto-optimal solutions and let H_w be a degenerate Hamiltonian in its minimum eigenvalue with corresponding minimum eigenstates $|x_1\rangle, \dots, |x_\ell\rangle$. There exists $w' \in W_d$, satisfying $\|w - w'\|_1 \leq \frac{\langle \lambda, w \rangle}{md}$, and $i \in [1, \ell]_{\mathbb{Z}}$ such that $H_{w'}$ is nondegenerate in its smallest eigenvalue with corresponding eigenvector $|x_i\rangle$.

Proof. From Lemma 6, we know that if Π_d^λ has no weakly-equivalent Pareto-optimal solutions, then for any w the Hamiltonian H_w is nondegenerate.

We consider now the case when the minimum eigenvalue of H_w is degenerate with ℓ Pareto-optimal solutions that are weakly-equivalent. Let x_1, \dots, x_ℓ be such weakly-equivalent Pareto-optimal solutions that are non-trivial and $x_i \neq x_j$ for all $i \neq j$. By Lemma 7 there exists $w' \in W_d$, where $w \neq w'$, such that $\langle f(x_i), w' \rangle$ is minimum among all $y \in D$. \square

If we consider our assumption from Section 2 that $m = \max_{x,i} \{f_i(x)\}$ is bounded by $\text{poly}(n)$, where n is the maximum number of bits of any element in $x \in D$, by Lemma 8 we have that any w' must satisfy $\|w - w'\|_1 \leq 1/\text{poly}(n)$. Then Theorem 2 follows immediately from Lemmas 2 and 8.

To see that the adiabatic evolution takes finite-time let $\Delta_{\max} = \max_s \|\frac{d}{ds} H(s)\|$ and $g_{\min} = \min_s g(s)$, where $g(s)$ is the eigenvalue gap of $H(s)$. Letting $T = O(\frac{\Delta_{\max}}{g_{\min}^2})$ suffices to find a supported solution corresponding to w . Since $g_{\min} > 0$ and $\|\frac{d}{ds} H(s)\| = \text{poly}(n)$, we conclude that T is finite.

5. An Application to the Two-Parabolas Problem

In this section we present a numerical example on how to construct a quantum adiabatic algorithm for the two-parabolas problem. Let TP_2^λ denote the two-parabolas problem with two objective functions and gap vector $\lambda = (0.2, 0.4)$ whose description is in Appendix B. In order to use the adiabatic algorithm of Section 3 we need to consider a collision-free version of the problem.

Let $n = 7$ be the number of bits that are needed to encode the entire domain of each objective function. Thus, a feasible solution $x \in \{0, \dots, 127\}$. Since the gap vector is $\lambda = (0.2, 0.4)$, we construct objective functions f_1 and f_2 that resemble two parabolas in such a way that for each pair of feasible solutions x, y it holds that $|f_1(x) - f_1(y)| \geq 0.2$ and $|f_2(x) - f_2(y)| \geq 0.4$. See Table A1 for a complete definition of TP_2^λ and Figure 2 shows a plot of all points.

We define the final and initial Hamiltonians following Equations (4) and (5), respectively. In particular, the initial Hamiltonian is defined as

$$H_0 = \sum_{x \in \{0,1\}^n \setminus \{0^n\}} 8|\hat{x}\rangle\langle\hat{x}|. \quad (9)$$

The number 8 in Equation (9) is there only to enhance the visual presentation of the plots in this paper. The Hamiltonian of the entire system for TP_2^λ is

$$H(s) = (1-s)H_0 + sH_w. \quad (10)$$

From the previous section we know that $T = O(\frac{\Delta_{max}}{g_{min}^2})$ suffices to find a supported solution corresponding to w [15]. The quantity Δ_{max} is usually easy to estimate [2]. The eigenvalue gap g_{min} is, however, very difficult to compute; indeed, determining for any Hamiltonian if $g_{min} > 0$ is undecidable [16].

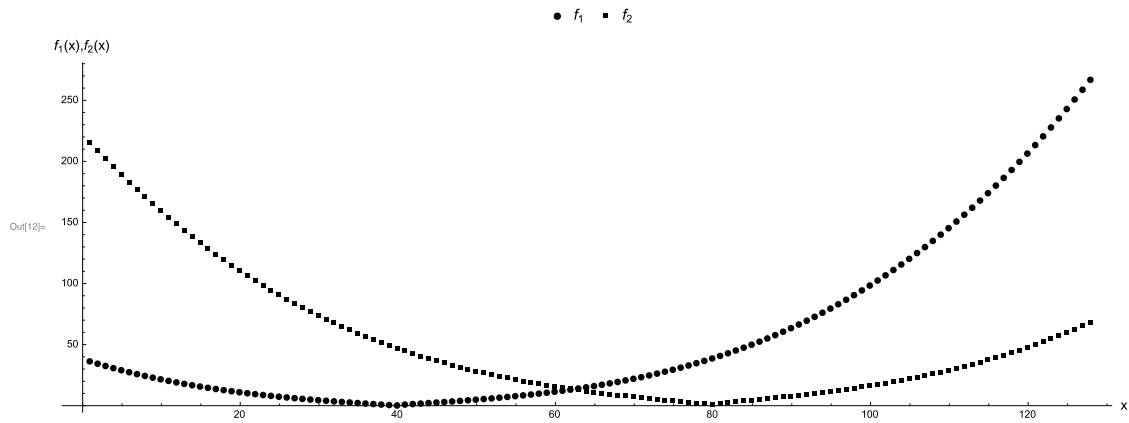


Figure 2. A discrete two-parabolas problem on seven qubits. Each objective function f_1 and f_2 is represented by the rounded points and the squared points, respectively. The gap vector is $\lambda = (0.2, 0.4)$. The trivial Pareto-optimal points are 40 and 80.

In Figure 3 we present the eigenvalue gap of TP_2^λ for $w = 0.57$ where we let $w_1 = w$ and $w_2 = 1 - w_1$; for this particular value of w the final Hamiltonian H_w has a unique minimum eigenstate which corresponds to Pareto-optimal solution $x_0 = 59$. The two smallest eigenvalues never touch, and exactly at $s = 1$ the gap is $|\langle f(x_0), w \rangle - \langle f(x_1), w \rangle|$, where $x_0 = 59$ and $x_1 = 60$ are the smallest and second smallest solutions with respect to w , which agrees with Lemmas 4 and 5. Figure 4 shows the eigenvalue gap as a function of s in a logarithmic scale.

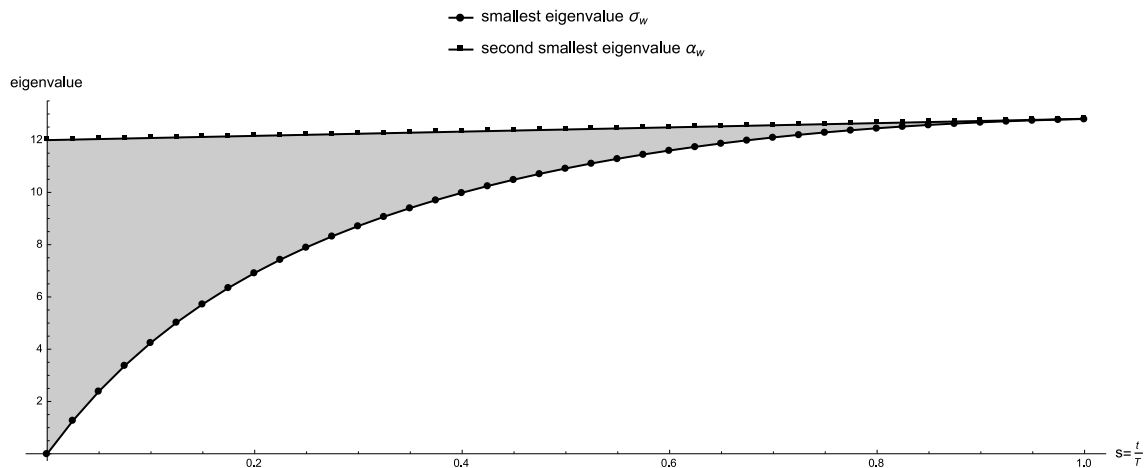


Figure 3. Eigenvalues of Equation (10) for the two-parabolas problem TP_2^λ of Figure 2 for $w = 0.57$. The eigenvalue gap $g(s)$ at $s = 1$ is exactly $|\langle w, f(x_0) \rangle - \langle w, f(x_1) \rangle|$, where $x_0 = 59$ and $x_1 = 60$ are the smallest and second smallest solutions with respect to this value of w .

Similar results can be observed for different values of w and a different number of qubits. Therefore, the experimental evidence lead us to conjecture that in the two-parabolas problem $g_{min} \geq |\langle f(x), w \rangle - \langle f(y), w \rangle|$, where x and y are the smallest and second smallest solutions with respect to w .

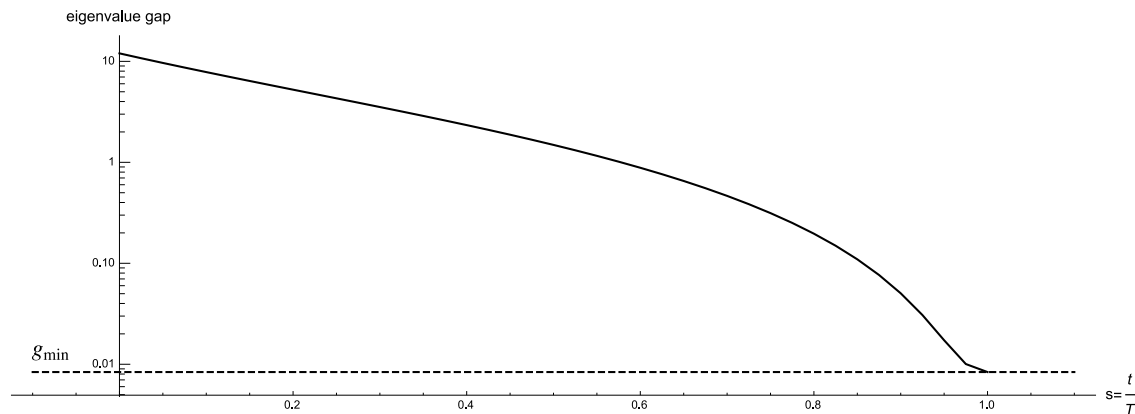


Figure 4. Logarithmic plot of the eigenvalue gap $g(s) = (\alpha_w - \sigma_w)$ as a function of s .

6. Concluding Remarks and Open Problems

In the last few years the field of quantum computation is finding new applications in artificial intelligence, machine learning and data analysis. These new discoveries were fueled by a deeper understanding in the foundations of quantum information and computation which resulted in new quantum algorithms for optimization problems (see for example [17]).

In this paper we addressed another side of optimization problems, namely, multiobjective optimization problems, where multiple objective functions must be optimized at the same time. This paper presented the first quantum multiobjective optimization algorithm with provable finite-time convergence. Other authors proposed quantum algorithms for multiobjective optimization [7,8] but these algorithms were ad-hoc heuristics with no theoretical guarantees for convergence. Furthermore, these aforementioned proposals utilized a hybrid approach of classical and quantum computation where Grover's search algorithm constitutes the only "quantum part" and the other parts of the algorithm are classical. The quantum algorithm of this work is based on the successful quantum adiabatic paradigm of Farhi et al. [2] and it is a full quantum algorithm, that is, all of its execution is performed by quantum operations with no classical parts except the initialization and read-out of the results.

The quantum multiobjective optimization algorithm of this work finds a single Pareto-optimal solution and in order to find other Pareto-optimal solutions it must be executed several times. Furthermore, the main result of this work, Theorem 2, requires that all multiobjective optimization problems be normal, collision-free and with no equivalent solutions. Even though this result constrains the class of multiobjective problems that can be solved, this work presents a first step forward to a general purpose quantum multiobjective optimization algorithm.

We end this paper by listing a few promising and challenging open problems.

1. We know from Lemma 2 that if we linearize a multiobjective optimization problem, some Pareto-optimal solutions (the non-supported solutions) may not be found. Considering that our quantum algorithm uses a linearization technique, a new mapping or embedding method of a multiobjective problem into a Hamiltonian is necessary in order to construct a quantum adiabatic algorithm that can also find supported Pareto-optimal solutions.
2. For a practical application of our quantum algorithm, the linearization w in Theorem 2 must be chosen so that the resulting total Hamiltonian is non-degenerate in its ground-state. Therefore, more research is necessary in order to develop a heuristic for choosing w before executing the algorithm.
3. As mentioned before, currently our algorithm is only good for multiobjective problems with no equivalent solutions. Natural multiobjective optimization problems appear in engineering and science with several equivalent solutions, and hence, in order to use our algorithm in a real-world situation we need to take into account equivalent solutions. This is a crucial point

- mainly because equivalent solutions yields degenerate ground-states in the total Hamiltonian, and hence, the quantum adiabatic theorem cannot be used.
- The time complexity of our quantum multiobjective algorithm depends on the spectral gap of the total Hamiltonian. Even though we presented some numerical results that suggest a polynomial execution time for the two-parabolas problem, a more thorough and rigorous approach must be done. This depends on the analysis of the spectral gap of Hamiltonians that can be constructed for specific multiobjective problems, for example, solving our conjecture for the two-parabolas problem of Section 5.

Author Contributions: B.B. and M.V.; Formal analysis, M.V.; Funding acquisition, B.B. and M.V.; Investigation, M.V.; Methodology, M.V.; Project administration, M.V.; Supervision, M.V.; Writing—original draft, B.B. and M.V.; Writing—review & editing, B.B. and M.V.

Funding: M.V. is supported by Conacyt research grant PINV15-208.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Proof of Non-Singularity of Equation (7)

We want to demonstrate that the determinant of the matrix

$$A = \begin{pmatrix} f_1(x_1) & f_2(x_1) \\ f_1(x_2) & f_2(x_2) \end{pmatrix}$$

is not 0.

Let $\det(A) = f_1(x_1)f_2(x_2) - f_2(x_1)f_1(x_2)$. Since x_1 and x_2 are two different Pareto-optimal solutions, we have that $x_1 \sim x_2$; therefore, there exists a function $f_i(x)$ that we will call $f_1(x)$ for which $f_1(x_1) < f_1(x_2)$. Since x_1 and x_2 are Pareto-optimal, it cannot happen that $f_i(x_1) < f_i(x_2)$ for all i ; otherwise, x_1 would dominate x_2 . Therefore, there exists at least one function $f_j(x)$ that we denote $f_2(x)$ such that $f_2(x_1) > f_2(x_2)$. Thus it holds that

$$\begin{aligned} f_1(x_1) &< f_1(x_2), \\ f_2(x_2) &< f_2(x_1). \end{aligned}$$

Then, it follows that $f_1(x_1)f_2(x_2) < f_1(x_2)f_2(x_1)$ and consequently $\det(A) \neq 0$.

Appendix B. Data for the Two-Parabolas Problem of Figure 2

Table A1. Complete definition of the two-parabolas example of Figure 2 for seven qubits.

x	$f_1(x)$	$f_2(x)$	x	$f_1(x)$	$f_2(x)$	x	$f_1(x)$	$f_2(x)$	x	$f_1(x)$	$f_2(x)$
0	36.14	214.879	1	34.219	208.038	2	32.375	201.354	3	30.606	194.825
4	28.91	188.449	5	27.285	182.224	6	25.729	176.148	7	24.24	170.219
8	22.816	164.435	9	21.455	158.794	10	20.155	153.294	11	18.914	147.933
12	17.73	142.709	13	16.601	137.62	14	15.525	132.664	15	14.5	127.839
16	13.524	123.143	17	12.595	118.574	18	11.711	114.13	19	10.87	109.809
20	10.07	105.609	21	9.309	101.528	22	8.585	97.564	23	7.896	93.715
24	7.24	89.979	25	6.615	86.354	26	6.019	82.838	27	5.45	79.429
28	4.906	76.125	29	4.385	72.924	30	3.885	69.824	31	3.404	66.823
32	2.94	63.919	33	2.491	61.11	34	2.055	58.394	35	1.63	55.769
36	1.214	53.233	37	0.805	50.784	38	0.401	48.42	39	0	46.139
40	0.801	43.939	41	1.205	41.818	42	1.614	39.774	43	2.03	37.805

Table A1. Cont.

x	$f_1(x)$	$f_2(x)$	x	$f_1(x)$	$f_2(x)$	x	$f_1(x)$	$f_2(x)$	x	$f_1(x)$	$f_2(x)$
44	2.455	35.909	45	2.891	34.084	46	3.34	32.328	47	3.804	30.639
48	4.285	29.015	49	4.785	27.454	50	5.306	25.954	51	5.85	24.513
52	6.419	23.129	53	7.015	21.8	54	7.64	20.524	55	8.296	19.299
56	8.985	18.123	57	9.709	16.994	58	10.47	15.91	59	11.27	14.869
60	12.111	13.869	61	12.995	12.908	62	13.924	11.984	63	14.9	11.095
64	15.925	10.239	65	17.001	9.414	66	18.13	8.618	67	19.314	7.849
68	20.555	7.105	69	21.855	6.384	70	23.216	5.684	71	24.64	5.003
72	26.129	4.339	73	27.685	3.69	74	29.31	3.054	75	31.006	2.429
76	32.775	1.813	77	34.619	1.204	78	36.54	0.6	79	38.54	0
80	40.621	1.2	81	42.785	1.804	82	45.034	2.413	83	47.37	3.029
84	49.795	3.654	85	52.311	4.29	86	54.92	4.939	87	57.624	5.603
88	60.425	6.284	89	63.325	6.984	90	66.326	7.705	91	69.43	8.449
92	72.639	9.218	93	75.955	10.014	94	79.38	10.839	95	82.916	11.695
96	86.565	12.584	97	90.329	13.508	98	94.21	14.469	99	98.21	15.469
100	102.331	16.51	101	106.575	17.594	102	110.944	18.723	103	115.44	19.899
104	120.065	21.124	105	124.821	22.4	106	129.71	23.729	107	134.734	25.113
108	139.895	26.554	109	145.195	28.054	110	150.636	29.615	111	156.22	31.239
112	161.949	32.928	113	167.825	34.684	114	173.85	36.509	115	180.026	38.405
116	186.355	40.374	117	192.839	42.418	118	199.48	44.539	119	206.28	46.739
120	213.241	49.02	121	220.365	51.384	122	227.654	53.833	123	235.11	56.369
124	242.735	58.994	125	250.531	61.71	126	258.5	64.519	127	266.644	67.423

References

- Catherine, C. McGeoch, C.C. *Adiabatic Quantum Computation and Quantum Annealing: Theory and Practice*; Morgan and Claypool: San Rafael, CA, USA, 2014.
- Farhi, E.; Goldstone, J.; Gutman, S.; Sipser, M. Quantum computation by adiabatic evolution. *arXiv* **2000**, arXiv:quant-ph/0001106.
- von Lücken, C.; Barán, B.; Brizuela, C. A survey on multi-objective evolutionary algorithms for many-objective problems. *Comput. Optim. Appl.* **2014**, *58*, 707–756. [[CrossRef](#)]
- Venegas-Andraca, S.; Cruz-Santos, W.; McGeoch, C.; Lanzagorta, M. A cross-disciplinary introduction to quantum annealing-based algorithms. *Contemp. Phys.* **2018**, *59*, 174–197. [[CrossRef](#)]
- Grover, L. A fast quantum mechanical algorithm for database search. In Proceedings of the 28th Annual ACM Symposium on the Theory of Computing (STOC), Philadelphia, PA, USA, 22–24 May 1996; pp. 212–219.
- Baritomp, W.P.; Bulger, D.W.; Wood, G.R. Grover’s quantum algorithm applied to global optimization. *SIAM J. Optim.* **2005**, *15*, 11701184. [[CrossRef](#)]
- Alanis, D.; Botsinis, P.; Xin Ng, S.; Lajos Hanzo, L. Quantum-Assisted Routing Optimization for Self-Organizing Networks. *IEEE Access* **2014**, *2*, 614–632. [[CrossRef](#)]
- Fogel, G.; Barán, B.; Villagra, M. Comparison of two types of Quantum Oracles based on Grover’s Adaptive Search Algorithm for Multiobjective Optimization Problems. In Proceedings of the 10th International Workshop on Computational Optimization (WCO), Federated Conference in Computer Science and Information Systems (FedCSIS), ACSIS, Prague, Czech Republic, 3–6 September 2017; Volume 11, pp. 421–428.
- Das, A.; Chakrabarti, B.K. Quantum annealing and quantum computation. *Rev. Mod. Phys.* **2008**, *80*, 1061. [[CrossRef](#)]
- Barán, B.; Villagra, M. Multiobjective optimization in a quantum adiabatic computer. *Electr. Notes Theor. Comput. Sci.* **2016**, *329*, 27–38. [[CrossRef](#)]
- Kung, H.T.; Luccio, F.; Preparata, F.P. On finding the maxima of a set of vectors. *J. ACM* **1975**, *22*, 469–476. [[CrossRef](#)]
- Papadimitriou, C.; Yannakakis, M. On the approximability of trade-offs and optimal access of web sources. In Proceedings of the 41st Annual Symposium on Foundations of Computer Science (FOCS), Washington, DC, USA, 12–14 November 2000; pp. 86–92.

13. Ehrgott, M.; Gandibleux, X. A survey and annotated bibliography of multiobjective combinatorial optimization. *OR Spektrum* **2000**, *22*, 425–460. [[CrossRef](#)]
14. Ambainis, A.; Regev, O. An elementary proof of the quantum adiabatic theorem. *arXiv* **2004**, arXiv:quant-ph/0411152.
15. Wim van Dam, W.; Mosca, M.; Vazirani, U. How powerful is adiabatic quantum computation? In Proceedings of the 42nd IEEE Symposium on Foundations of Computer Science (FOCS), Las Vegas, NV, USA, 14–17 October 2001; pp. 279–287.
16. Cubitt, T.; Perez-Garcia, D.; Wolf, M. Undecidability of the Spectral Gap. *Nature* **2015**, *528*, 207–211. [[CrossRef](#)] [[PubMed](#)]
17. Biamonte, J.; Witteck, P.; Pancotti, N.; Rebentrost, P.; Wiebe, N.; Lloyd, S. Quantum machine learning. *Nature* **2017**, *549*, 195–202. [[CrossRef](#)] [[PubMed](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).