Cross Entropy Optimization of Constrained Problem Hamiltonians for Quantum Annealing

Christoph Roch, Alexander Impertro and Claudia Linnhoff-Popien

LMU Munich, Germany
christoph.roch@ifi.lmu.de

Abstract. This paper proposes a Cross Entropy approach to shape constrained Hamiltonians by optimizing their energy penalty values. The results show a significantly improved solution quality when run on D-Wave's quantum annealing hardware and the numerical computation of the eigenspectrum reveals that the solution quality is correlated with a larger minimum spectral gap. The experiments were conducted based on the Knapsack-, Minimum Exact Cover- and Set Packing Problem. For all three constrained optimization problems we could show a remarkably better solution quality compared to the conventional approach, where the energy penalty values have to be guessed.

Keywords: Quantum annealing · Cross entropy method · Optimization · Minimum spectral gap · Constrained Hamiltonian · D-Wave Systems

1 Introduction

Experimental quantum computing has gained a lot of attention in the last decade, since quantum computers have been commercialized by their companies [2,3]. There exist different types of quantum computing hardware. The more known one is the quantum gate-model computer. It is the quantum pendant to our classical computers, which work with classical logical gates. The other type of quantum computers are quantum annealers, which are particularly designed for solving or finding good approximations to optimization problems. D-Wave Systems is the first company, which has made their quantum annealing hardware available for public and a lot of research has been done since then [1, 5, 13, 23].

D-Wave's Quantum Annealing (QA) algorithm, which is implemented in hardware, is based on the adiabatic quantum computing principle [20]. First, one has to map the corresponding optimization problem to a so called Ising Hamiltonian in order to execute it on the hardware. The fundamental process of QA then is to physically interpolate between an initial Hamiltonian, whose minimal energy configuration (or ground state) is easy to prepare, and a problem Hamiltonian, whose minimal energy configuration, that corresponds to the best solution of the defined Ising problem Hamiltonian, is sought. According to the adiabatic theorem, if this process is executed slowly enough, and the coherence domain is sufficiently large, the probability to stay in the ground state and thus in the minimal energy configuration of the problem Hamiltonian is close to

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77980-1_5

one [6]. However, due to thermal fluctuations or a non-adiabatic anneal process the system can leap from the ground to an excited state. The minimum distance between the ground state and the first excited state throughout any point in the anneal process is called the minimum spectral gap. Since it is physically hard to ensure long coherence times in quantum systems, one can not increase the anneal time arbitrarily in order to avoid computational errors by jumping to excited states.

That is why we address this problem experimentally, not by adjusting the anneal time, but by applying a Cross-Entropy (CE) method to optimize the hyperparameters of the solution landscape, which are represented by the energy penalty values of the constrained Hamiltonian of the optimization problem. We can show that by optimizing the penalty values, the minimum spectral gap of the problem Hamiltonian is scaled and shifted. In effect, D-Wave's QA algorithm has a higher chance of remaining in the ground state, which results in a significant increase in solution quality as compared to the conventional approach without CE optimization. For our experiments we used the Knapsack Problem (KP), Minimum Exact Cover (MEC) Problem and the Set Packing (SP) Problem to verify our approach. We experimentally reveal the linear correlation between the size of the minimum spectral gap and the corresponding approximation ratio of the D-Wave annealer to the best known solution (BKS) of the problem instance.

2 Background

2.1 Quantum Annealing Algorithm

Quantum annealing is a metaheuristic for solving complex optimization and decision problems [15]. D-Wave's quantum annealing algorithm is implemented in hardware, and it is designed to find the lowest energy state of a spin glass. Such a system can be described by an Ising Hamiltonian of the form

$$\mathcal{H}(s) = \sum_i h_i s_i + \sum_{i<j} J_{ij} s_i s_j \tag{1}$$

where h_i is the on-site energy of qubit i, J_{ij} are the interaction energies of two qubits i and j, and s_i represents the spin $(-1, +1)$ of the i-th qubit. The basic process of quantum annealing is to physically interpolate between an initial Hamiltonian $\mathcal{H}_I$ with an easy to prepare minimal energy configuration (or ground state), and a problem Hamiltonian $\mathcal{H}_P$, whose minimal energy configuration, that corresponds to the best solution of the defined problem, is sought (see Eq. (2)). The physical principle on which the D-Wave computation process is based on can be described by a time-dependent Hamiltonian as stated in Eq. (2).

$$\mathcal{H}(t) = A(t)\mathcal{H}_I + B(t)\mathcal{H}_P \tag{2}$$

The anneal functions $A(t)$ and $B(t)$ must satisfy $B(t = 0) = 0$ and $A(t = \tau) = 0$, with τ being the total evolution time. When the state evolution changes from

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77980-1_5

$t = 0$ to $t = \tau$, the annealing process, described by $\mathcal{H}(t)$, leads to the final form of the Hamiltonian corresponding to the objective Ising problem that needs to be minimized. Therefore, the ground state of the initial Hamiltonian $\mathcal{H}(0) = \mathcal{H}_I$ evolves to the ground state of the problem Hamiltonian $\mathcal{H}(\tau) = \mathcal{H}_P$. The measurements performed at time τ deliver low energy states of the Ising Hamiltonian as stated in Eq. (1). According to the adiabatic theorem, if this transition is executed sufficiently slowly (i.e. τ is large enough), and the coherence domain is large enough, the probability to stay in the ground state of the problem Hamiltonian is close to one [6].

However, due to a non-adiabatic anneal process the system can jump from the ground to an excited state. The minimum distance between the ground state and the first excited state — the one with the lowest energy apart from the ground state — throughout any point in the anneal process is called the minimum spectral gap $g_{\min}$ of $\mathcal{H}(t)$, and is defined as

$$g_{\min} = \min_{0 \le t \le T} \min_{j \neq 0} [E_j(t) - E_0(t)] \tag{3}$$

where $E_j(t)$ is any higher lying energy state and $E_0(t)$ the ground state [14]. The adiabatic theorem states that staying in the ground state is enforced by setting the change rate of the time-dependent Hamiltonian $\mathcal{H}(t)$ proportional to $1/g_{\min}^{\delta}$, with δ depending on the distribution of eigenvalues at higher energy levels. δ may range from one to even three in some circumstances [12, 17, 26]. To understand the efficiency of adiabatic quantum computing, we need to analyze $g_{\min}$, but in practice, this is a difficult task [7], which we try to approach experimentally in this work.

For the sake of completeness, note that there exists an alternative and often used formulation to the Ising spin glass system. The so called Quadratic Unconstrained Binary Optimization (QUBO) formulation [8], which is mathematically equivalent and uses 0 and 1 for the spin variables [27]. The quantum annealer is also able to minimize the functional form of the QUBO formulation $x^T Q x$, with x being a vector of binary variables $\{0, 1\}$ of size n, and Q being an $n \times n$ real-valued matrix describing the relationship between the variables. Given the matrix $Q : n \times n$, the annealing process tries to find binary variable assignments $x \in \{0, 1\}^n$ to minimize the objective function.

2.2 Knapsack Problem

In the NP-Complete Knapsack Problem [16], n items are given, each having a certain weight w_α and a certain value c_α. The items must be picked in a way that the total weight of the items is less than or equal to the knapsack capacity W, i.e. $\sum_{\alpha=1}^{n} w_\alpha x_\alpha \le W$, and the total sum of the item values $\sum_{\alpha=1}^{n} c_\alpha x_\alpha$ is maximized. Variable x_α is set 1 if the item is packed in the knapsack and 0 otherwise [19]. In order to implement the KP on D-Wave's quantum computer using QA, we need to encode the objective function of the KP into a Hamiltonian which is diagonal in the computational basis.

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77980-1_5

The weight constraint can be encoded in the following quadratic Hamiltonian, as stated in [18]:

$$\mathcal{H}_1 = A\left(1 - \sum_{n=1}^{W} y_n\right)^2 + A\left(\sum_{n=1}^{W} n y_n - \sum_{\alpha=1}^{N} w_\alpha x_\alpha\right)^2 \tag{4}$$

while the objective function is straightforwardly

$$\mathcal{H}_2 = -B\sum_{\alpha=1}^{N} c_\alpha x_\alpha. \tag{5}$$

Here, y_n for $1 \leq n \leq W$ is a binary variable, which is set to 1, if the final weight of the knapsack is n and 0 otherwise. $\mathcal{H}_1$ enforces that the weight can only take exactly one value and that the weight of the items in the knapsack equals the value we claimed it did. The parameters A, B are chosen according to

$$0 < B \cdot \max(c_\alpha) < A \tag{6}$$

in order to penalize violations of the weight constraint. Note that one can reduce the number of binary variables using the so called log trick to $N + \lfloor 1 + logW \rfloor$ [18].

2.3 Minimum Exact Cover Problem

The Minimum Exact Cover Problem is an NP-Hard constrained optimization problem, where a set $U = \{1, ..., n\}$, and subsets $V_i \subseteq U (i = 1, ..., N)$ are given, such that $U = \bigcup_i V_i$.

The task is to find the minimum number of sets V_i with the elements of those sets being disjoint, and the union of the sets is U. The Hamiltonians $\mathcal{H}_3 = \mathcal{H}_1 + \mathcal{H}_2$ are stated in [18]:

$$\mathcal{H}_1 = A\sum_{\alpha}^{n}\left(1 - \sum_{i:\alpha\in V_i} x_i\right)^2 \tag{7}$$

$$\mathcal{H}_2 = B\sum_{i} x_i \tag{8}$$

In Eq. (7) α denotes the elements of U, while i denotes the subsets V_i. $\mathcal{H}_1 = 0$, if every element is included exactly one time, which implies that the unions of the subsets are disjoint. With the additional Hamiltonian $\mathcal{H}_2$ the smallest number of subsets is sought. The ground state of this Hamiltonian will be $m * B$, where m is the smallest number of subsets required for the complete union. The ratio of the penalty values A and B can be determined by regarding the worst case scenario which is that there are a very small number of subsets with a single common element, whose union is U. To ensure this does not happen, one can set

$$A > n \cdot B. \tag{9}$$

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77980-1_5

2.4 Set Packing Problem

The Set Packing Problem is also an NP-hard constrained optimization problem. Given a setup as in section 2.3, its difficulty lies in finding the maximum number of subsets V_i which are all disjoint. In [18] the following Hamiltonians $\mathcal{H}_3 = \mathcal{H}_1 + \mathcal{H}_2$ are given:

$$\mathcal{H}_1 = A \sum_{i,j:V_i \cap V_j \neq \emptyset} x_i x_j \tag{10}$$

$$\mathcal{H}_2 = -B \sum_i x_i \tag{11}$$

$\mathcal{H}_1$ is minimized only when all subsets are disjoint, while $\mathcal{H}_2$ simply counts the number of included sets. Choosing the penalty values

$$B < A \tag{12}$$

ensures that it is never favorable to violate the constraint $\mathcal{H}_1$. Considering there will always be a penalty of at least A per extra set included. Just as the Minimum Exact Cover Problem the Set Packing Problem requires N spins.

2.5 Cross-Entropy Method

The Cross-Entropy method is a Monte Carlo method for importance sampling and optimization, and is known to perform well on combinatorial optimization problem with noisy objective functions [24, 25].

We experimentally implement the method using a common CE algorithm (see Alg. 1), as stated for example in [28]. In each step of the iterative optimization, a set of points $a_1 ... a_n$ from the distribution p is sampled, based on its current parameterization Φ_{g-1} (line 3). To each point $a_1 ... a_n$, the objective function f of the optimization problem assigns values $v_1 ... v_n$ (line 4). Finally, a fraction ρ of elite samples is chosen based on a selection routine (line 5 and 6) and used to compute a new parameterization of p, Φ_g.

Algorithm 1 Cross-Entropy

1: **function** OPTIMIZE(p, Φ_0, f, ρ, n, G)
2: **for** $g = 1 \rightarrow G$ **do**
3: $a_1 ... a_n \sim p(\cdot|\Phi_{g-1}), \mathbf{a} \leftarrow a_1 ... a_n$
4: $v_1 ... v_n \sim f(a_1) ... f(a_n), \mathbf{v} \leftarrow v_1 ... v_n$
5: sort $\mathbf{a}$ according to $\mathbf{v}$
6: $\Phi_g \leftarrow \text{argmax}_\Phi \prod_{i=1}^{\lceil n\rho \rceil} p(a_i|\Phi)$
return a_1

The convergence rate of the algorithm critically depends on the distribution p, as it determines the new sample points $\mathbf{a}$ for each generation. The initial

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77980-1_5

distribution $p(\cdot|\Phi_0)$ should be chosen such that it reproduces optimal samples as closely as possible. However, when this is not possible, a generally applicable approach is to choose a distribution which covers the entire sample space. This increases the probability for the algorithm to evolve towards a good solution already in early generations.

After each iteration, a maximum likelihood estimate of the currently chosen elite fraction is done to update the parameterization Φ_g according to the following rule $\mu_g = \frac{\sum_{i=1}^{\lceil n\rho \rceil} X_i}{\lceil n\rho \rceil}$, $\sigma_g^2 = \frac{\sum_{i=1}^{\lceil n\rho \rceil} (X_i - \mu_g)^T (X_i - \mu_g)}{\lceil n\rho \rceil}$ and $\Phi_g = \langle \mu_g, \sigma_g^2 \rangle$, which is valid for a multivariate Gaussian distribution. Other critical parameters are the selected fraction ρ of elite samples, the population size n and the number of generations G, which must be carefully adjusted for a given problem in order to maximize the likelihood of finding a good solution. In Section 4, we explain how the CE method is adapted for our specific task.

3 Related Work

In 2017, Mark W. Coffey studied the Knapsack Problem within an Adiabatic Quantum Computing (AQC) framework. He mapped the optimization problem to an Ising model and used small problem instances to evaluate his approach. He points the relevance of theoretical and numerical investigations regarding the minimum spectral gap and its location in the anneal path out, in order to improve AQC [11].

More insights brought Choi in 2019. She theoretically showed, that adjusting the energy penalty of the Ising Maximum weighted Independent Set Problem, one may change the quantum evolution from one that has an anti-crossing to one that does not have, or the other way around, and thus drastically change the minimum spectral gap [10]. Following this insight, we propose an adapted CE method to automatically adjust the penalty values of constrained Hamiltonians to influence the quantum evolution of D-Wave's QA hardware, or more precisely the size and location of the corresponding minimum spectral gap, so that D-Wave's quantum annealer has a higher chance of remaining in the ground state.

In previous work, we already applied CE in the field of gate based quantum computers [22]. Similarly to this work, we shaped the solution landscape of the Knapsack problem Hamiltonian, which allowed the classical optimizer of the hybrid Quantum Approximate Optimization Algorithm (QAOA) to find better gate parameter and hence resulted in an improved performance.

4 QA with Cross-Entropy

In our approach we use an adapted CE method to optimize the penalty values A and B of our problem Hamiltonians. Varying those values significantly changes the energy landscape of the corresponding constrained optimization problem, and therefore also influences the pathway of the Quantum Annealing algorithm. By adjusting those penalty values it is possible to scale and shift the minimum

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77980-1_5

spectral gap $g_{\min}$ and hence improve the probability of staying in the ground state throughout the anneal process.

An example can be seen in Fig. 1. In Fig. 1a the time-dependent eigenspectrum of a Minimum Exact Cover Hamiltonian with randomly selected penalty values is visualized. The same setup, but with optimized penalty values, found by CE, can be seen in Fig. 1b. The histograms in Fig. 1c and 1d show the corresponding solution qualities w.r.t the best known solution (BKS). The solution quality is associated with the approximation ratio, which is calculated as follows: $\textit{Approx. ratio} = \frac{\#BKS}{\#Measurements}$ with #BKS being the number of measuring the BKS and #Measurements being the total number of measurements (default 100).

In Fig. 1a the minimum spectral gap $g_{\min}$ is ≈ 1 and located in the middle of the anneal process, while in Fig. 1b the gap $g_{\min}$ is ≈ 3 and shifted to the beginning of the anneal process. The impact of scaling and shifting $g_{\min}$ also reflects in an improved quantum annealing solution quality in Fig. 1d. Therefore, optimizing the penalty values of the problem Hamiltonian, in a way that $g_{\min}$ increases, decreases the likelihood of (thermal) excitations out of the instantaneous ground state, and consequently allows the quantum annealer to reach a better solution quality.

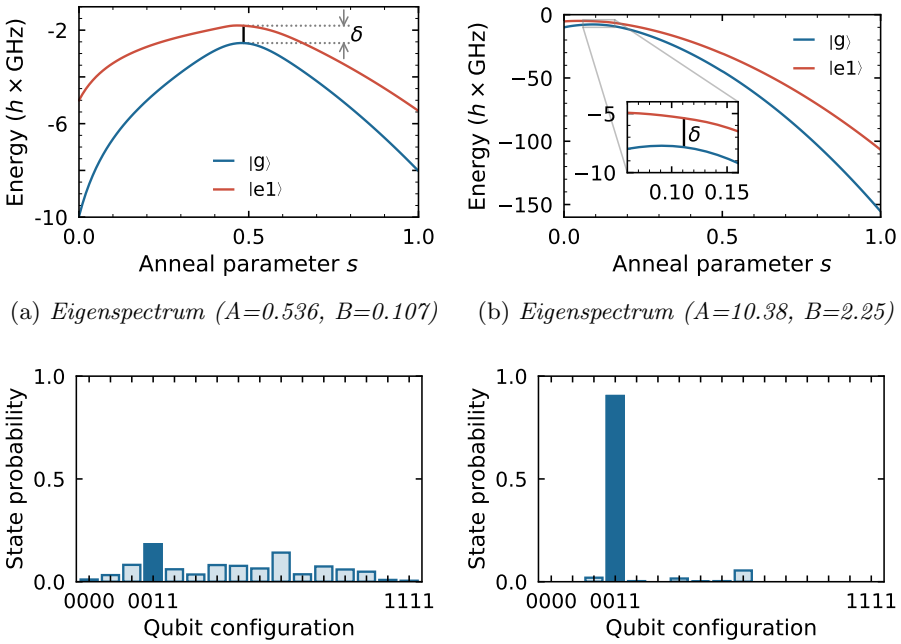
Regarding the problem Hamiltonians, the choice of penalty values are restricted by Eq. (6),(9) and (12), respectively for the Knapsack, Minimum Exact Cover and Set Packing Problem [18]. To satisfy those constraints, we use a modified CE optimization scheme (see Alg. 2), in which the penalty values A, B are sampled from truncated normal distributions p. Since the allowed values for A depend on the choice of B, we first draw a value for B (line 3) with an appropriately chosen sampling range Γ_B. Afterwards, the value for A is drawn over a sampling range $\Gamma_A(B)$, such that the penalization constraint of the corresponding optimization problem is satisfied (line 5). This is done for n samples. For each sample, we construct the corresponding Hamiltonian, as described in [18] and run D-Wave's QA heuristic to assign a value $v_1 ... v_n$ corresponding to the approximation ratio of the best found solution for each A_i, B_i-pair (line 7). This is done iteratively for a specified number of generations G. In Fig. 2, the process of CE with QA and the fitness of each generation for the MEC problem Hamiltonian $\mathcal{A}$ is shown.

5 Evaluation

5.1 Experimental Setup

For the experimental evaluation, we used D-Wave's 2000Q quantum annealer (solver name: *DW_2000Q_6 (lower noise)*). We used the standard anneal schedule with $20\mu s$ and the standard energy scales $A(t)$ and $B(t)$, as stated in [4], which are required to compute the energy of a problem at a specific point in the annealing process. For embedding the problem instances on the D-Wave QPU, we used the the minorminer library [9], which tries to find an efficient embedding of the logical problem graph to the physical architecture, called chimera

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77980-1_5



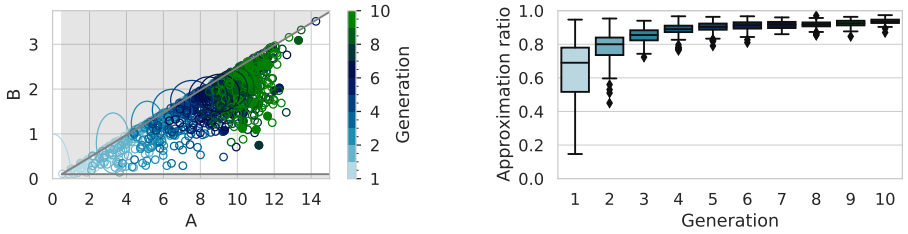
(a) *Eigenspectrum (A=0.536, B=0.107)* (b) *Eigenspectrum (A=10.38, B=2.25)*

(c) *Solution Quality (A=0.536, B=0.107)* (d) *Solution Quality (A=10.38, B=2.25)*

Fig. 1: Eigenspectra and solution qualities for MEC problem instance $\mathcal{A}$. Problem $\mathcal{A}$ consists of four subsets with numbers ranging from 1 to 4, $[(1,2),(1,3),(1,2,4),(3)]$. The BKS (marked in dark blue) is 0011, i.e. sets with index 2 and 3 are in the solution. The $g_{\min} = \delta$ marks the minimum spectral gap, which is ≈ 1 for random penalty values A and B and ≈ 3 for optimized ones, see Figs. 1a and 1b, respectively. Figs. 1c and 1d show the corresponding solution qualities, i.e. the state probability of the BKS compared to the other solutions.

graph, of the D-Wave quantum annealer. Since not every problem graph fits directly to the architecture, due to the sparse connectivity of the qubits on the QPU, physical ancillary qubits need to be used to represent one logical qubit. Additionally, it is known that not every qubit of the D-Wave QPU has the same physical quality. Thus, the embedding has an influence on the solution quality of the problem, as stated in [21]. That means, the same graph structure embedded on different physical qubits of the QPU leads to different solution qualities. In Fig. 3 one graph structure of the MEC problem instance $\mathcal{A}$ is embedded ten times, each time on different qubits of the QPU. For every embedding we used the same 100 individuals (penalty value pairs), their fitness was averaged over three runs per individual, to draw a fair comparison. One can see, that in the best embedding (the third one) the mean solution quality is around 71%, while

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77980-1_5



(a) *CE optimization (best penalty value pair found: A=10.38,B=2.25)*

(b) *Fitness of each individuum per generation*

Fig. 2: Example CE optimization with $G = 10$ is shown in Fig. 2a. The ellipses represent the μ_g and σ_g^2 of generation g. The filled circles correspond to the best ρ fraction of individuals. The best values found by CE, for this specific MEC problem instance $\mathcal{A}$, were 10.38 and 2.25 for A, respectively B. In Fig. 2b the fitness of each individuum per generation (population size is 100) computed with D-Wave's 2000Q annealer is represented by the boxplots.

Algorithm 2 Cross-Entropy Energy Penalty Optimization

function OPTIMIZE(p, Φ_0, f, ρ, n, G)
  for $g = 1 \rightarrow G$ **do**
    $B_1 ... B_n \sim p(\cdot|\Phi_{g-1}, \Gamma_B)$
    $\mathbf{B} \leftarrow B_1 ... B_n$
    $A_1 ... A_n \sim p(\cdot|\Phi_{g-1}, \Gamma_A(\mathbf{B}))$
    $\mathbf{A} \leftarrow A_1 ... A_n$
    $v_1 ... v_n \sim QA(A_1, B_1) ... QA(A_n, B_n)$
    $\mathbf{v} \leftarrow v_1 ... v_n$
    sort **A, B** according to **v**
    $\Phi_g \leftarrow \text{argmax}_\Phi \prod_{i=1}^{\lceil n\rho \rceil} p(A_i, B_i|\Phi)$
  return A_1, B_1

in the worst embedding (the second one) the mean solution quality is around 50%. Consequently, the embedding can not be influenced directly and thus may also impact the overall solution quality of the CE optimization.

In Tab. 1 the parameter settings of the CE method are listed. Since we are using a truncated normal distribution to sample from, we need to specify additional clipping parameters for both penalty values A and B. The sampling range of A is computed according to the corresponding problem constraints.

We used four problem instances for each optimization problem to test our CE approach with D-Wave's QA algorithm. The instances per optimization problem, named $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ and $\mathcal{D}$ range from 4 to 7 logical qubits, respectively. Since D-Wave's quantum annealing hardware is still in its infancy regarding the number of qubits and their connectivity, those logical problems already led to physical ancillary qubits to enable a valid embedding. Those ancillary qubits are coupled by a chain strength parameter to ensure that by measuring they collapse to the

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77980-1_5

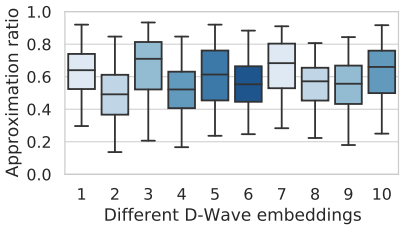


Fig. 3: Different D-Wave embeddings for MEC problem instance $\mathcal{A}$.

Table 1: Cross-Entropy Parameter Settings

CE Attributes							
G	n	ρ	γ^*	min σ^2	σ_0^2	μ_0	B sample range
10	100	0.1	0.5	0.1	1.0	0.0	$[0.1, 10.0]$

*The learning rate specifies the amount of changes from Φ_{g-1} to Φ_g.

same basis state. However, this additional parameter also influences the overall solution quality and makes it harder for CE to optimize large problem instances. That is why we rather used small instances, to demonstrate the effect of CE. However, our approach is theoretically also applicable to larger problems.

5.2 Results & Discussion

In Fig. 4 the linear correlation of the approximation ratio to the BKS and the minimum spectral gap is plotted. The number in the upper left corner is the Pearson product-moment correlation coefficient, which assumes values in the range of -1 to $+1$, where the extrema occur in case of a strong negative or positive correlation, while a value of 0 indicates uncorrelated variables. The first row represents the KP instances $\mathcal{A} - \mathcal{D}$, while the second and third row represents the MEC and SP instances, respectively. The blue circles represent the individuals (penalty value pairs) over ten generations of the CE method (in total 1000 individuals). The solution quality is given by the approximation ratio of the BKS and can be calculated by dividing the BKS counts by the number of measurements (default 100). Note that we averaged the approximation ratio for each individual of the population over three runs on the D-Wave hardware to compensate the stochasticity of the quantum system.

The results show that for each problem instance (perhaps with the exception of MEC $\mathcal{C}$), there exists a correlation between the approximation ratio computed with D-Wave's quantum annealer and the minimum spectral gap. Notice that, within the broadest range of the minimum spectral gap, the approximation ratio is in general much higher and in some cases near to 1.0 (see MEC problem instance $\mathcal{A}$ and $\mathcal{B}$), while in cases, where the minimum spectral gap is small the approximation ratio is also comparatively lower and more variance occurs (see MEC problem instance $\mathcal{C}$).

In Fig. 5 the results of QA with CE and the conventional QA approach are compared against each other. We tested both methods on different problem instances as stated in Section 5.1. The results are shown in Fig. 5a-5c for the KP, MEC, and SP, respectively.

CE was initialized with the parameter setting of Tab. 1 and D-Wave's QA algorithm was used as explained in Section 5.1. W.r.t the classical QA approach,

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77980-1_5

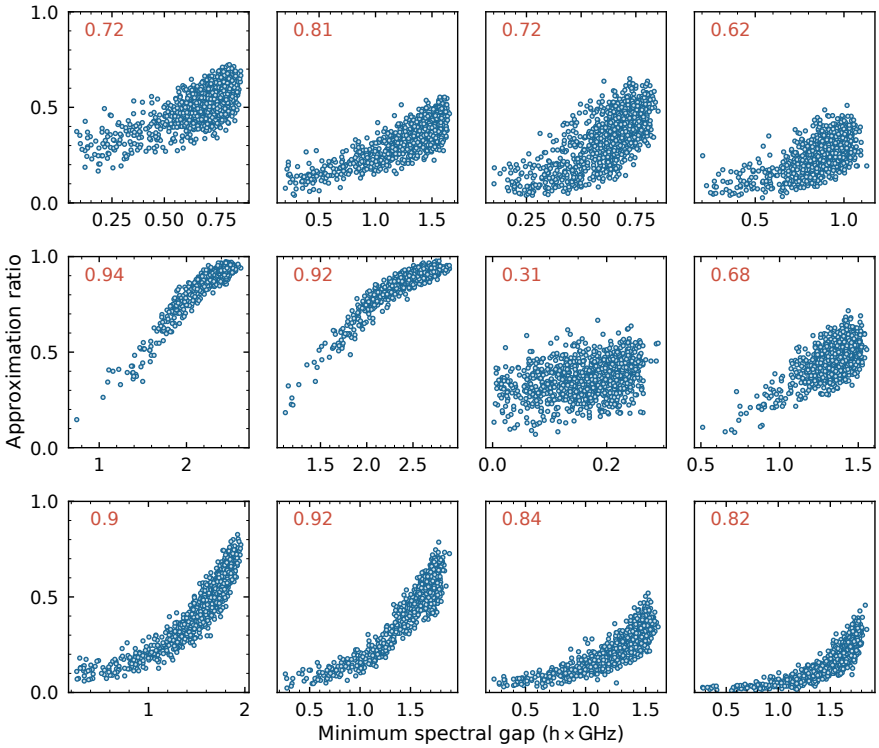


Fig. 4: Correlation plots of the approximation ratio to the BKS and the minimum spectral gap. The blue circles represent the individuals (penalty value pairs) over ten generations of the CE method (in total 1000 individuals). The first row represents the KP instances $\mathcal{A} - \mathcal{D}$, while the second and third row represents the MEC and SP instances, respectively. The red number is the correlation coefficient.

we randomly sampled five penalty value pairs of the same sampling range as stated in Tab. 1. Each penalty value pair was executed 10 times on D-Wave's 2000Q quantum annealer and the corresponding solution qualities (approximation ratio) are represented in the respective box plot "random". For the QA with CE box plots, named "optimal", we used the ρ fraction of the penalty value population of the last CE generation and calculated their fitness, i.e. solution quality.

The results show, that for the KP in Fig. 5a, the solution quality, w.r.t the mean could be increased by around 500% in the best case (see problem $\mathcal{B}$) and by around 85% in the worst case (see problem $\mathcal{D}$), by using the optimized penalty values.

Also for the MEC problem in Fig. 5b, the solution qualities were increased by around 170% in problem $\mathcal{C}$) in the best case and by 30% in the worst case in

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77980-1_5

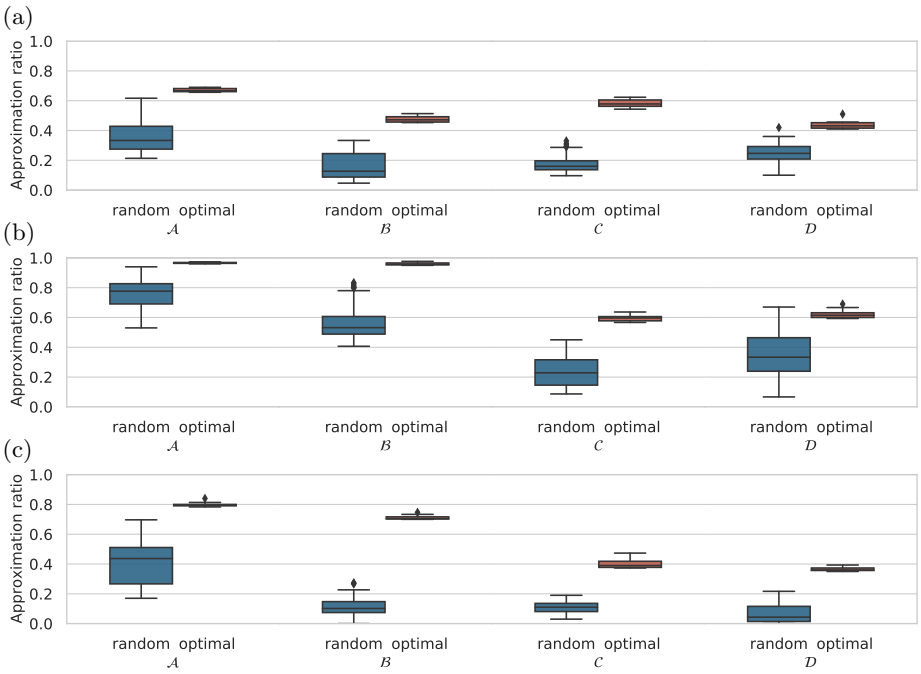


Fig. 5: Solution qualities for the four problem instances $\mathcal{A}$-$\mathcal{D}$ of the KP, MEC and SP problem instances are represented in 5a-5c, respectively. The "random" boxplots represent the approximation ratio of five randomly sampled penalty value pairs (each run 10 times), while the "optimal" boxplots represent the approximation ratio of the ρ fraction of the penalty value pair population of the 10th generation of the CE method.

problem $\mathcal{A}$. Furthermore note, that in problem $\mathcal{A}$ and $\mathcal{B}$ an approximation ratio of nearly 100% could be reached with the optimized penalty values.

In Fig. 5c a significant growth in the quality of the solution can be seen, too, for the SP problem. In the best cases (problem $\mathcal{B}$ and $\mathcal{D}$) an increase of around 600% could be achieved while in the worst case (problem $\mathcal{A}$) the optimized penalty values still led to an 80% increase in solution quality.

The overall solution quality in general decreases with the size of the problem instances. However, this is obvious since the number of possible solution and therefore the whole solution space increases.

Another feature is the comparatively small variance of the approximation ratio of the optimized penalty values, over all problem instances, which can be seen in Fig. 5. That is due to the values being picked from the best ρ fraction of individuals of the last CE generation, while the randomly sampled energy penalty value pairs may contain disadvantageously ones. However, this does not detract from the fact that our CE approach is quite stable w.rt. the stochastic quantum system.

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77980-1_5

6 Conclusion

In this paper we have presented a Cross Entropy approach to shape constrained Hamiltonians by optimizing their penalty values. We showed by the numerical computation of the eigenspectrum that this optimization leads to a scaling and shifting of the minimal spectral gap and thus makes it easier for D-Wave's quantum annealing algorithm to stay in the ground state during the anneal process. Consequently, this results in an improved overall solution quality (approximation ratio). The experiments were conducted based on the Knapsack-, Minimum Exact Cover- and Set Packing Problem. For all three constrained optimization problems we could show a significantly better solution quality compared to the conventional approach. Moreover, using the optimized penalty values of the last generation of CE results in a lower variance of the solution quality, meaning that less averages have to be taken to find good solutions. However, since the penalty values found by the CE method differed in the used problem instances, we want to use machine learning techniques in order to investigate correlations between the optimized penalty value pairs and be able to reuse them for other problem instances. Due to this fact, we currently see the strength of our approach in the optimization of the overall solution qualities, but less in achieving computational speedups. Furthermore, we want to study our approach also for larger problem instances and take the hyperparameter of the physical qubit chains and the embedding into account to even improve this approach.

Acknowledgements. This work was funded by the BMWi project PlanQK (01MK20005I).

References

1. D-Wave Launches Leap, the First Real-Time Quantum Application Environment. https://www.dwavesys.com/press-releases/d-wave-launches-leap-first-real-time-quantum-application-environment, accessed: 2021-01-10
2. D-wave systems sells its first quantum computing system to lockheed martin corporation. https://www.dwavesys.com/news/d-wave-systems-sells-its-first-quantum-computing-system-lockheed-martin-corporation, accessed: 2021-01-10
3. Ibm makes quantum computing available on ibm cloud to accelerate innovation. https://www-03.ibm.com/press/us/en/pressrelease/49661.wss, accessed: 2021-01-10
4. QPU-Specific Anneal Schedules. https://support.dwavesys.com/hc/en-us/articles/360005267253-What-are-the-Anneal-Schedule-Details-for-this-QPU-, accessed: 2021-01-10
5. Adachi, S.H., Henderson, M.P.: Application of quantum annealing to training of deep neural networks. arXiv preprint arXiv:1510.06356 (2015)
6. Albash, T., Lidar, D.A.: Adiabatic quantum computation. Reviews of Modern Physics **90**(1), 015002 (2018)
7. Amin, M.H., Choi, V.: First-order quantum phase transition in adiabatic quantum computation. Physical Review A **80**(6), 062326 (2009)

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77980-1_5

8. Boros, E., Hammer, P.L., Tavares, G.: Local search heuristics for quadratic unconstrained binary optimization (qubo). Journal of Heuristics **13**(2), 99–132 (2007)
9. Cai, J., Macready, W.G., Roy, A.: A practical heuristic for finding graph minors. arXiv preprint arXiv:1406.2741 (2014)
10. Choi, V.: The effects of the problem hamiltonian parameters on the minimum spectral gap in adiabatic quantum optimization. Quantum Information Processing **19**(3), 90 (2020)
11. Coffey, M.W.: Adiabatic quantum computing solution of the knapsack problem. arXiv preprint arXiv:1701.05584 (2017)
12. Farhi, E., Goldstone, J., Gutmann, S., Sipser, M.: Quantum computation by adiabatic evolution. arXiv preprint quant-ph/0001106 (2000)
13. Feld, S., Roch, C., Gabor, T., Seidel, C., Neukart, F., Galter, I., Mauerer, W., Linnhoff-Popien, C.: A hybrid solution method for the capacitated vehicle routing problem using a quantum annealer. Frontiers in ICT **6**, 13 (2019)
14. Grant, E.K., Humble, T.S.: Adiabatic quantum computing and quantum annealing (07 2020)
15. Kadowaki, T., Nishimori, H.: Quantum annealing in the transverse ising model. Phys. Rev. E **58**, 5355–5363 (1998)
16. Karp, R.M.: Reducibility among combinatorial problems. In: Complexity of computer computations, pp. 85–103. Springer (1972)
17. Lidar, D.A., Rezakhani, A.T., Hamma, A.: Adiabatic approximation with exponential accuracy for many-body systems and quantum computation. Journal of Mathematical Physics **50**(10), 102106 (2009)
18. Lucas, A.: Ising formulations of many np problems. Frontiers in Physics **2**, 5 (2014)
19. Martello, S.: Knapsack problems: algorithms and computer implementations. Wiley-Interscience series in discrete mathematics and optimiza tion (1990)
20. McGeoch, C.C.: Adiabatic quantum computation and quantum annealing: Theory and practice. Synthesis Lectures on Quantum Computing **5**(2), 1–93 (2014)
21. Padilha, D.: Solving np-hard problems on an adiabatic quantum computer. UNSW School of Mechanical and Manufacturing Engineering (2014)
22. Roch, C., Impertro, A., Phan, T., Gabor, T., Feld, S., Linnhoff-Popien, C.: Cross entropy hyperparameter optimization for constrained problem hamiltonians applied to qaoa. arXiv preprint arXiv:2003.05292 (2020)
23. Roch, C., Phan, T., Feld, S., Müller, R., Gabor, T., Hahn, C., Linnhoff-Popien, C.: A quantum annealing algorithm for finding pure nash equilibria in graphical games. In: International Conference on Computational Science. pp. 488–501. Springer (2020)
24. Rubinstein, R.: The cross-entropy method for combinatorial and continuous optimization. Methodology and computing in applied probability **1**(2), 127–190 (1999)
25. Rubinstein, R.Y., Kroese, D.P.: The cross-entropy method: A unified approach to monte carlo simulation, randomized optimization and machine learning. Information Science & Statistics, Springer Verlag, NY (2004)
26. Schaller, G., Mostame, S., Schützhold, R.: General error estimate for adiabatic quantum computing. Physical Review A **73**(6), 062307 (2006)
27. Su, J., Tu, T., He, L.: A quantum annealing approach for boolean satisfiability problem. In: Proc. of the 53rd Annual Design Automation Conference. p. 148. ACM (2016)
28. Weinstein, A., Littman, M.L.: Open-loop planning in large-scale stochastic domains. In: Twenty-Seventh AAAI Conference on Artificial Intelligence (2013)

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77980-1_5