

RESEARCH ARTICLE

Mean field approximation for solving QUBO problems

Máté Tibor Veszeli ^{*}, Gábor Vattay

Department of Physics of Complex Systems, Eötvös Loránd University, Budapest, Hungary

^{*} mate.veszeli@ttk.elte.hu

Abstract

The Quadratic Unconstrained Binary Optimization (QUBO) problem is NP-hard. Some exact methods like the Branch-and-Bound algorithm are suitable for small problems. Some approximations like stochastic simulated annealing for discrete variables or mean-field annealing for continuous variables exist for larger ones, and quantum computers based on the quantum adiabatic annealing principle have also been developed. Here we show that the mean-field approximation of the quantum adiabatic annealing leads to equations similar to those of thermal mean-field annealing. However, a new type of sigmoid function replaces the thermal one. The new mean-field quantum adiabatic annealing can replicate the best-known cut values on some of the popular benchmark Maximum Cut problems.



OPEN ACCESS

Citation: Veszeli MT, Vattay G (2022) Mean field approximation for solving QUBO problems. PLoS ONE 17(8): e0273709. <https://doi.org/10.1371/journal.pone.0273709>

Editor: Mohammadreza Hadizadeh, Central State University & Ohio University, UNITED STATES

Received: March 28, 2022

Accepted: August 11, 2022

Published: August 30, 2022

Copyright: © 2022 Veszeli, Vattay. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the manuscript.

Funding: The authors thank the National Research, Development, and Innovation Office of Hungary for support via the Quantum Information National Laboratory grant NKFIH-831-2/2021.

Competing interests: The authors have declared that no competing interests exist.

1 Introduction

From solid state physics [1] to social phenomena [2] Ising models can describe a wide range of complex systems. Spin models are versatile because they are simple yet able to show complex phenomena, like emergence and phase transitions [3–5]. Spin models are also important in large, real-life optimization problems which can be distilled down to finding the global minimum of a high-dimensional, nonlinear function. Most of these tasks are NP-hard [6]. Small problems up to approximately a hundred nodes for a dense problem can be solved by exact methods like the Branch-and-Bound (B&B) [7], or Branch-and-Cut [8]. Larger problems can be solved approximately with intuitive methods like the tabu search [9, 10], semi-definite programming [11–13], stochastic simulated annealing (SSA) [14–16] and quantum annealing [17–19]. Quantum annealing is also the key idea behind adiabatic quantum computers [20, 21] such as the D-Wave system [22–25]. To reduce the computational costs of stochastic methods, the mean-field annealing (MFA) [26] has been proposed. In MFA, the dynamics of spins are replaced with the evolution of their average values. As the temperature decreases, the MFA algorithm updates these averages based on their values at the previous temperature. Computation using the means reaches equilibrium faster than the corresponding stochastic dynamics, and MFA relaxes to a solution much faster than SSA, leading to an overall decrease in computational effort. Such methods have become popular lately, and special devices like coherent Ising machines [27, 28] have been developed.

In this paper, we develop the mean-field version of the quantum annealing [17] method and apply it to the maximum cut problem, which belongs to the larger set of Quadratic Unconstrained Binary Optimization (QUBO) problems. The mean-field quantum annealing introduced here is similar to [18, 19] in spirit. The novelty is the proposed quantum MFA is that the equations derived here are fully analogous to the thermal MFA equations, the tunneling rate of the quantum MFA corresponds to the inverse temperature, and the sigmoid theta function is replaced with a new and different sigmoid function. The new sigmoid function improves the performance of the quantum MFA, which is in line with recent result [29] emphasizing the importance of the shape of the sigmoid function.

The structure of this paper is the following. In section 2 we summarize the thermal mean-field annealing method for the Ising problem. Then in section 3 we present the quantum annealing problem and its proposed novel mean-field solution. In section 4 we test the new method on the maximum cut problem benchmark G-sets [30], and finally, in section 5 we compare the new method with SSA and B&B.

2 Thermal mean-field annealing

The variational principle of statistical physics is a powerful tool for examining interacting systems at finite temperatures. The most straightforward version of this principle is the mean-field approximation. This approximation works best when the number of interactions per spin is large. In real-life optimization problems, the number of nodes is usually not as large as in physics, typically just tens of thousands, and usually, they have no symmetry. The statistical physical derivation of the mean-field annealing starts from the expression of the free energy

$$F(T) = -T \ln \sum_n e^{-E_n/T}, \tag{1}$$

at finite temperature, T , where the sum goes over all the system states, and E_n is the energy. The probability of finding the system at state n is proportional to the Boltzmann weight $e^{-E_n/T}$, therefore, the free energy coincides with the ground state energy at zero temperature. The problem is that even if we know all the E_n energies, apart from the most straightforward cases, we cannot evaluate the summation due to the exponentially large number of terms. The variational principle states that the exact free energy is always smaller or equal to variational free energy

$$F^V(T) = \sum_n P_n E_n + T \sum_n P_n \ln(P_n), \tag{2}$$

where P_n is an arbitrary probability distribution. The closer the probabilities P_n are to the real ones, the lower $F^V(T)$ will be. In practice, the distribution can't be too complicated since we have to calculate the variational free energy analytically, avoiding exponential summations. The typical strategy is to consider a class of probability distributions $P_n(\mathbf{m})$ parametrized by some set of parameters \mathbf{m} , then calculate $F^V(\mathbf{m}, T)$ and finally find the minimum \mathbf{m}^* . This solution is temperature dependent and we will refer to it as $\mathbf{m}^*(T)$. The variational free energy is then $F^V(T) = F^V(\mathbf{m}^*(T), T)$.

The Ising model consist of interacting spins: $\mathbf{S} = (S_1, S_2 \dots S_N)$, with components $S_i \in \{\pm 1\}$. The model is defined by its energy:

$$E(\mathbf{S}) = -\frac{1}{2} \sum_{ij} J_{ij} S_i S_j - \sum_i h_i S_i, \tag{3}$$

where J_{ij} is the interaction between spin i and j , and h_i is the external magnetic field. In the mean field approximation, we assume that a variational distribution factorizes. In the case of spin systems

$$P^{MF}(\mathbf{S}; m) = \prod_{i=1}^N P_i(S_i; m_i) = \prod_i \frac{1 + m_i S_i}{2}. \tag{4}$$

This distribution is normalized and the expectation values of the spins are the parameters $\langle S_i \rangle = m_i$. The variational free energy is

$$F^{MF}(\mathbf{m}, T) = -\frac{1}{2} \sum_{ij} J_{ij} m_i m_j - \sum_i h_i m_i + T \sum_i \left[\frac{1 + m_i}{2} \ln \left(\frac{1 + m_i}{2} \right) + \frac{1 - m_i}{2} \ln \left(\frac{1 - m_i}{2} \right) \right]. \tag{5}$$

We have to find its minimum, so its derivative has to be zero

$$\frac{\partial F^{MF}}{\partial m_i} = -\sum_j J_{ij} m_j - h_i + T \frac{1}{2} \ln \left(\frac{1 + m_i}{1 - m_i} \right) = 0 \tag{6}$$

and the second derivative, the Hessian should be positive definite

$$\frac{\partial^2 F^{MF}}{\partial m_i \partial m_j} = -J_{ij} + \frac{T \delta_{ij}}{1 - m_i^2} \succ 0. \tag{7}$$

We can rewrite the implicit Eq (6) as a self-consistent equation:

$$m_i = \tanh \left[\frac{h_i + \sum_j J_{ij} m_j}{T} \right]. \tag{8}$$

If $\mathbf{m}(T)$ is known for a given temperature T , we can determine it for a slightly lower temperature $\mathbf{m}(T - \Delta T)$ by using the self-consistent equation iteratively as it was proposed in [26]. This method requires fewer function evaluations than SSA. Above some temperature $T > T_0$ the iterative equation

$$m_i(T - \Delta T) = \tanh \left[\frac{h_i + \sum_j J_{ij} m_j(T)}{T - \Delta T} \right] \tag{9}$$

converges towards a unique high temperature solution $m_i \approx h_i/T$. There are more solutions at lower temperatures. At T_0 , a bifurcation occurs, and the iterations after that follow one of the new post-bifurcation solutions. Later on at successively lower temperatures $T_1 > T_2 > T_3 \dots$ further bifurcations occur, and the iteration follows one of the new solutions. The bifurcation temperatures T_n are those points, where one of the eigenvalues and consequently the determinant of the Hessian (7) becomes zero

$$\det \left(-J_{ij} + \frac{T_n \delta_{ij}}{1 - m_i^2(T_n)} \right) = 0 \tag{10}$$

at the solution $m_i(T_n)$ of the self-consistent Eq (8). Due to the nature of the solution, the mean-field annealing is also called a bifurcation machine. The procedure described above leads to one of the local minima of free energy. In general, there is no guarantee that the method leads to the absolute minimum, but decreasing the size of cooling steps ΔT and increasing the overall running time in parallel, improves the chances of finding deeper minima [31].

3 Quantum mean field annealing

Instead of the free energy and temperature, we can use quantum mechanics and the adiabatic theorem to determine the system’s ground state. The adiabatic theorem asserts that if a quantum system is initially at the ground state and the corresponding time-dependent Hamiltonian operator changes sufficiently slowly, there is a gap between the ground state energy and the rest of the spectrum. The system remains at the instantaneous ground state [32]. An application of this theorem is to find the ground state of a complicated Hamiltonian. The ground state of the initial Hamiltonian H_i should be easy to prepare, and the final operator H_f is the one whose ground state we are interested in. In that case, we can compose the Hamilton operator of the adiabatic problem

$$H(t) = (1 - s(t))H_i + s(t)H_f, \tag{11}$$

where $s(t)$ is a continuous, monotonic function of time, with $s(0) = 0$ and $s(T_A) = 1$. The time T_A is the annealing time, which must be large. The simplest choice for this is $s(t) = t/T_A$. To solve the Ising problem with the adiabatic method we have to use the quantum version of the Ising model, as the final Hamilton operator

$$H_f = -\frac{1}{2} \sum_{ij} J_{ij} \sigma_i^z \sigma_j^z - \sum_i h_i \sigma_i^z, \tag{12}$$

where σ^z is the Pauli z-matrix. For the initial Hamilton operator the standard choice is the transverse magnetic field term

$$H_i = -\Delta \sum_i \sigma_i^x, \tag{13}$$

where Δ is the strength of the external field. This transverse term plays similar role to the entropy term in Eq (2) and it induces tunneling between the minima of the Ising Hamiltonian. The initial ground state is

$$|\Psi_0\rangle = \frac{\otimes_{i=1}^N (|\uparrow\rangle + |\downarrow\rangle)}{\sqrt{2}}, \tag{14}$$

which is a product state, and after the annealing the final state is the ground state of Eq (12) and from that we can determine the minimal energy spin configuration of Eq (3). For a given $H(t)$ the state of the system is governed by the Schrödinger equation.

$$i \frac{d}{dt} |\Psi(t)\rangle = H(t) |\Psi(t)\rangle. \tag{15}$$

Initially the system is at ground state: $|\Psi(t = 0)\rangle = |\Psi_0\rangle$, If the annealing time T_A is sufficiently large, then the evolving wave function reaches the ground state of the Ising model.

To develop an MFA to the quantum problem we introduce the energy functional

$$E[|\Psi(s)\rangle] = \langle \Psi(s) | H(s) | \Psi(s) \rangle. \tag{16}$$

At $s = 0$, we set the initial wave function to the initial state $|\Psi(0)\rangle = |\Psi_0\rangle$ and for $s > 0$, we want to keep it in the minimal energy state. In the mean-field approximation, just like in the thermal case, we assume that the spins are independent and the trial state vector is a product

$$|\Phi(s)\rangle = \otimes_{i=1}^N [c_{i\downarrow}(s) |\downarrow\rangle + c_{i\uparrow}(s) |\uparrow\rangle] \tag{17}$$

where the amplitudes are constrained by $|c_{i\downarrow}(s)|^2 + |c_{i\uparrow}(s)|^2 = 1$. The expectation values of the

Pauli operators can be expressed with the amplitudes as

$$\begin{aligned}
 m_i^z &= \langle \Phi | \sigma_i^z | \Phi \rangle = |c_{i\uparrow}|^2 - |c_{i\downarrow}|^2, \\
 m_i^x &= \langle \Phi | \sigma_i^x | \Phi \rangle = c_{i\uparrow}^* c_{i\downarrow} + c_{i\downarrow}^* c_{i\uparrow}, \\
 m_i^y &= \langle \Phi | \sigma_i^y | \Phi \rangle = -i c_{i\uparrow}^* c_{i\downarrow} + i c_{i\downarrow}^* c_{i\uparrow},
 \end{aligned}
 \tag{18}$$

where m_i^x, m_i^y and m_i^z are real. They satisfy $(m_i^x)^2 + (m_i^y)^2 + (m_i^z)^2 = 1$. It is easy to show that $c_{i\uparrow}$ and $c_{i\downarrow}$ can be chosen to be real, and in that case $m_i^y(s) = 0$. In terms of the spin expectation values the energy functional is

$$E[\mathbf{m}^z(s)] = s \left[-\frac{1}{2} \sum_{ij} J_{ij} m_i^z m_j^z - \sum_i h_i m_i^z \right] - (1-s)\Delta \sum_i \sqrt{1 - (m_i^z)^2},
 \tag{19}$$

where we used $m_i^x = \sqrt{1 - (m_i^z)^2}$, since $m_i^x > 0$ during the minimization process. The parameter s plays the role of the temperature. The energy functional is minimal where the derivatives in terms of the expectation values vanish

$$\frac{\partial E}{\partial m_i^z} = -s \sum_j J_{ij} m_j^z - h_i + (1-s)\Delta \frac{m_i^z}{\sqrt{1 - (m_i^z)^2}} = 0,
 \tag{20}$$

and the Hessian is positive definite

$$\frac{\partial^2 E}{\partial m_i^z \partial m_j^z} = -s J_{ij} + (1-s)\Delta \frac{\delta_{ij}}{(1 - (m_i^z)^2)^{3/2}} \succ 0.
 \tag{21}$$

The first part of Eq (20) is the same as in (6). The difference is in the second term, as it is seen in Fig 1. Now, we can rewrite the implicit Eq (20) as a self-consistent equation

$$m_i = \sigma \left[\frac{h_i + \sum_j J_{ij} m_j}{\Gamma} \right],
 \tag{22}$$

where the tunneling rate is $\Gamma = \Delta(1-s)/s$ and we dropped the superscript $m_i = m_i^z$. The new sigmoid function is

$$\sigma(x) = \frac{x}{\sqrt{1+x^2}}.
 \tag{23}$$

This new equation is the main result of the paper. It is the equivalent of the thermal self-consistent Eq (8), where the tunneling rate Γ is the analog of of the temperature T and $\sigma(x)$ replaces $\tanh(x)$.

The iterative solution of the new equation is similar to the thermal case. Given $m_i(s=0) = 0$ we would like to determine $m_i(s=1)$. The $\mathbf{m}(s)$ vector is minimizing the energy functional $E(\mathbf{m}(s))$, where we assume that the $s \mapsto \mathbf{m}(s)$ trajectory is continuous. Without the external magnetic field the Ising model has a \mathbb{Z}_2 symmetry. That means the $m_i^z = 0$ is a solution to Eq (20), and it is a minimum as long as the smallest eigenvalue of the Hessian is above zero. For the parameter s this means

$$s < \frac{\Delta}{\lambda_{\max}(\mathbf{J}) + \Delta}
 \tag{24}$$

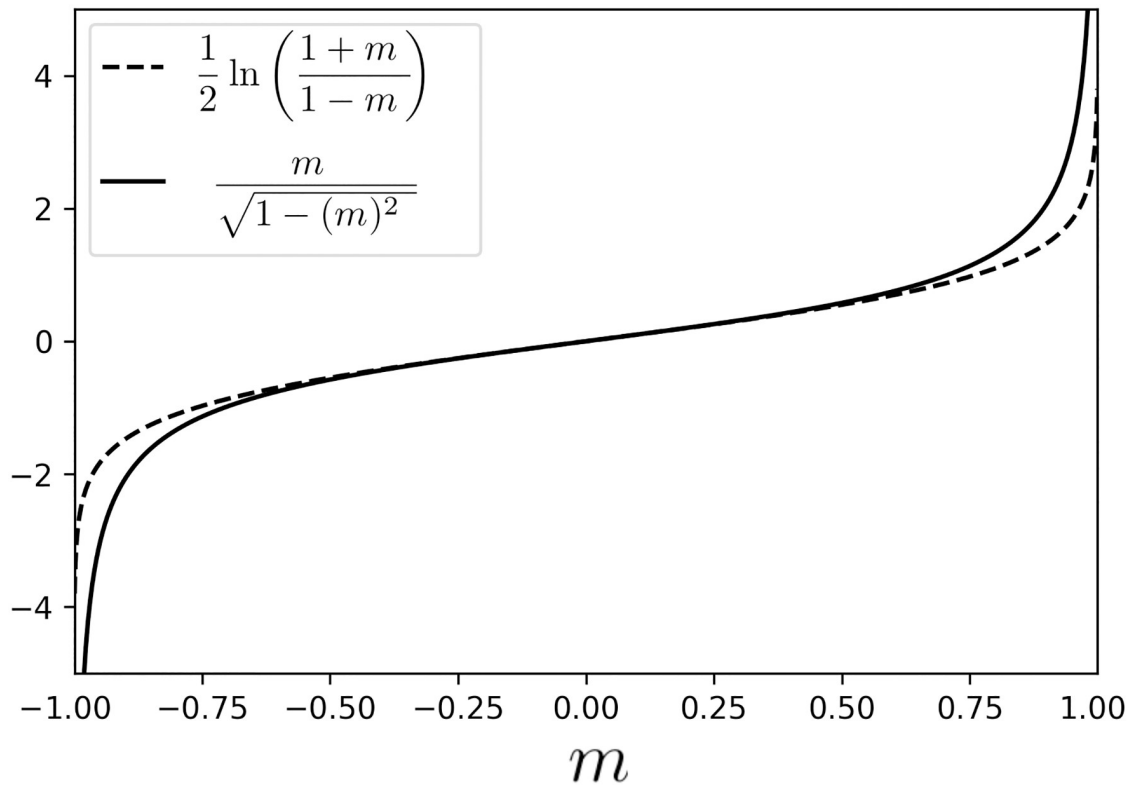


Fig 1. Comparison of the second term in the equations of the thermal (6) and the quantum (20) annealing.

<https://doi.org/10.1371/journal.pone.0273709.g001>

where $\lambda_{\max}(\mathbf{J})$ is the largest eigenvalue of the matrix J_{ij} . We can set Δ to 1 and rescale J_{ij} so that its largest eigenvalue is also 1. Now the trivial solution holds until s reaches 0.5. That means it is sufficient to start the simulation from $s = 0.5$. Once $m_i(s = 1)$ is known in the simulation, we round it to either 1 or -1. This gives us the \mathbf{S} spin configuration. Since the derivative of $E(\mathbf{m}^z)$ diverges as some m_i approaches ± 1 it is advantageous to use a different parametrization $m_i(\vartheta_i) = \cos(\vartheta_i)$, $m_i^x(\vartheta_i) = \sqrt{1 - m_i^2} = \sin(\vartheta_i)$ and then the derivative $\partial E / \partial \vartheta_i = s(\sum_j J_{ij} \sin(\vartheta_j) \cos(\vartheta_i) + h_i \sin(\vartheta_i)) + (1 - s)(-\Delta) \cos(\vartheta_i)$ is regular for all ϑ_i .

4 Benchmarking

As it was summarized by Lucas [33], many NP hard problems can be formulated with the Ising model. One of them is the QUBO problem, where we have to find the binary variables $x_i \in \{0, 1\}$, minimizing the quadratic functional $\text{argmin}\{q(\mathbf{x})\}$, where $q(\mathbf{x}) = \sum_{ij} Q_{ij} x_i x_j$, and \mathbf{Q} is a symmetric matrix. Replacing the binary variables with spin variables $x_i = (1 + S_i)/2$ leads to the Ising problem with parameters $J_{ij} = -\frac{1}{2} Q_{ij}$ and $h_i = -\frac{1}{2} \sum_j Q_{ij}$. The QUBO problem is then equivalent to finding the ground state of this Ising model. The NP-hard Maximum Cut problem [34] is often used as a benchmark for optimization algorithms. The task is to partition an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ into two subsets $(\mathcal{S}, \mathcal{V} \setminus \mathcal{S})$ such that the number of edges between these subsets is as large as possible. If the graph is defined via its adjacency matrix, where A_{ij} is one, if i and j are connected and zero otherwise, then the corresponding cut value is

$$CV = \sum_{\{i,j\} \in \mathcal{E}, i \in \mathcal{S}, j \in \mathcal{V} \setminus \mathcal{S}} A_{ij} = \sum_{\substack{ij \\ (i < j)}} A_{ij} \frac{1 - S_i S_j}{2}. \tag{25}$$

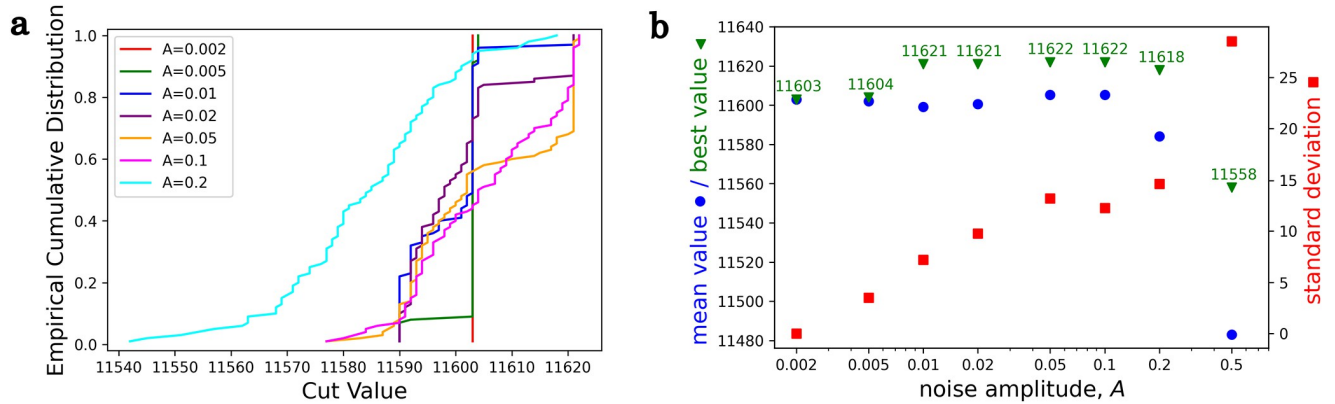


Fig 2. a, Empirical Cumulative Distribution of the Cut Values of G3 for different external magnetic field amplitudes. b, Mean value, best value and standard deviation of the Cut Values of G3 for different external magnetic field amplitudes.

<https://doi.org/10.1371/journal.pone.0273709.g002>

In (25) $S_i = +1$ means the spin i is in the subset \mathcal{S} , and $S_i = -1$ means it's in $\mathcal{V} \setminus \mathcal{S}$. Maximizing the cut value is equivalent to minimizing the Ising energy, where $J_{ij} = -A_{ij}$ and $h_i = 0$.

During the simulation, the \mathbb{Z}_2 symmetry is disadvantageous because, without the external field, the system remains in the $m_i^z = 0$ solution forever. Adding some small random external magnetic field, h_i breaks this symmetry. We can use this field as noise and run the simulation repeatedly for many random realizations. One such distribution is shown in Fig 2a for the set G3 from G-set [30], which is one of the favorite problems to benchmark the methods for solving the Maximum Cut problem. This is a random graph with 800 nodes and 19176 links. The h_i components are randomly generated uniformly from $(-A, A)$, where A is the amplitude, and one hundred trials were generated for all amplitudes. The mean values, the best values, and standard deviations are shown in Fig 2b. If the amplitude is small, we have a high average CV with a small deviation, on the other hand, if the amplitude is large, then the average CV is small, and the deviation is high. An optimal amplitude range exists, where the average is still large enough as well as the deviation, resulting in the largest best value from the sample. In the G3, this is around $A = 0.05 - 0.1$, where the best-known cut value 11622 has been achieved.

We tested our algorithm with other Max-Cut problems from the G-set. We focused on the smaller ones, i.e., the largest was the G22 graph with 2000 nodes. The results are summarized in Table 1. The last column shows the best values we could find in the literature [9, 35–39]. In most cases, our best results are the same as the best known; in the rest, it is close.

5 Comparison of the QMFA with SSA and B&B

Two widely used algorithms for discrete optimization problems are the stochastic simulated annealing (SSA), based on the Monte Carlo method, and the Branch-and-Bound method (B&B) [40]. The former is a metaheuristic algorithm, just like the QMFA, but has different hyperparameters. The latter is an exact algorithm where the tests were terminated after a fixed time due to the prior knowledge of its exponential time complexity. These two algorithms are used to compare the performance of the QMFA.

The pseudocode of QMFA is the following:

algorithm: Mean Field Annealing:
preprocess: $\underline{J} := J/\lambda_{\max}(J)$, $\Delta := 1$
input: $E(\underline{J}; \underline{h}, s)$, $\partial_{\underline{J}} E(\underline{J}; \underline{h}, s)$, $\partial_{\underline{J}} \partial_{\underline{J}} E(\underline{J}; \underline{h}, s)$, A , n
output: \underline{J}

Table 1. G-set benchmarks. QMFA: Our results with the quantum MFA method. Best: Best-known results.

	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11	G12	G22
QMFA	11624	11620	11622	11646	11631	2178	2006	2005	2050	1999	560	554	13353
Best	11624	11620	11622	11646	11631	2178	2006	2005	2054	2000	564	556	13359

<https://doi.org/10.1371/journal.pone.0273709.t001>

```

 $\underline{q} := \pi/2$ 
 $\underline{h}$  := random numbers between  $-A$  and  $A$ 
for  $k:=0$  to  $n$ 
   $s := 1/2 + k/(2n)$ 
   $\underline{q} :=$  local minimum of  $E(\underline{q}; \underline{h}, s)$  using  $\partial_{\underline{q}} E(\underline{q}; \underline{h}, s)$  and  $\partial_{\underline{q}} \partial_{\underline{q}} E(\underline{q}; \underline{h}, s)$  with the
  initial value  $\underline{q}$ 
 $\underline{m} := \cos(\underline{q})$ 
 $\underline{S} := (-1$  if  $m_i < 0$  else  $1$  for  $m_i$  in  $\underline{m}$ )
return  $\underline{S}$ 

```

We have used the Newton-conjugate-gradient method from the SciPy library for finding the local minimum. In this QMFA algorithm the noise amplitude A and the number of steps n are the hyperparameters, and once the vector \underline{S} is known, the Cut Value is $CV = \frac{1}{4} \sum_{ij} J_{ij} (S_i S_j - 1)$, with the original matrix \underline{J} . Fig 3 shows one instance from G1 how the Cut Value converges as n increases. With $n = 20$, we can be sure that we have reached the final Cut Value. One run with $n = 20$ takes $t_{QMFA} = (0.93 \pm 0.18)$ s. Unfortunately, not every \underline{h} noise gives the best Cut Value therefore, we have to run the algorithm multiple times. As shown in Fig 2 only a well-tuned noise amplitude can achieve a good result. For G1, the $A = 0.1$ is a good choice, but if we do not know this hyperparameter, we have to explore this parameter space as

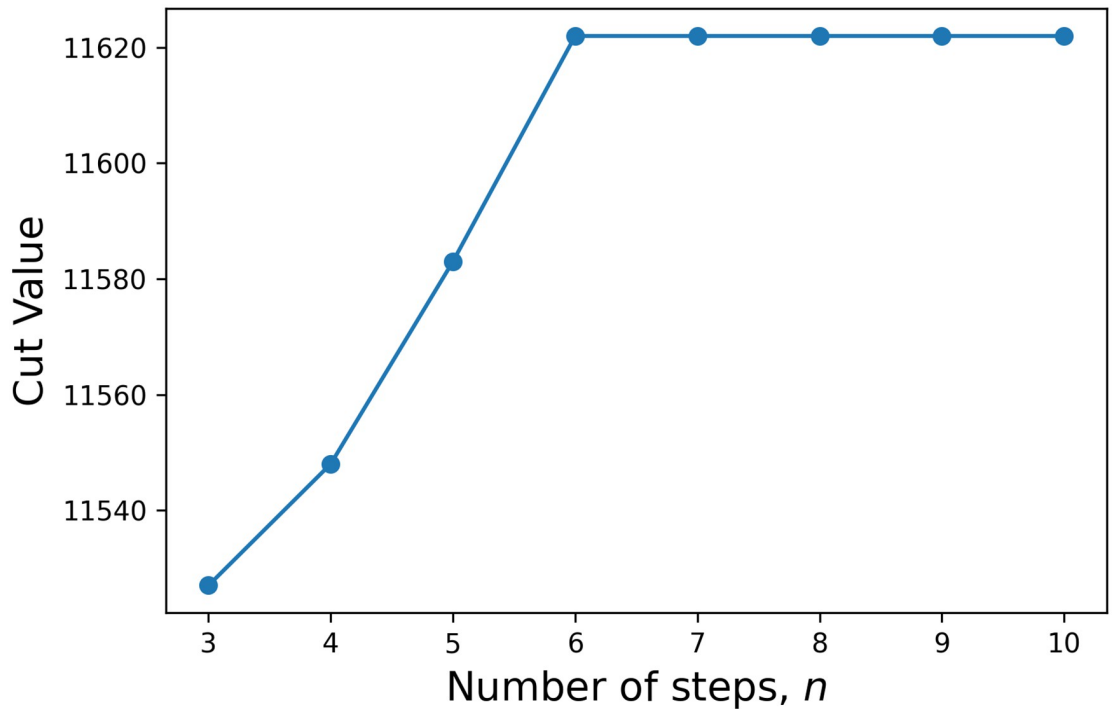


Fig 3. Convergence of the Cut Value of QFMA as a function of steps for the G1 problem.

<https://doi.org/10.1371/journal.pone.0273709.g003>

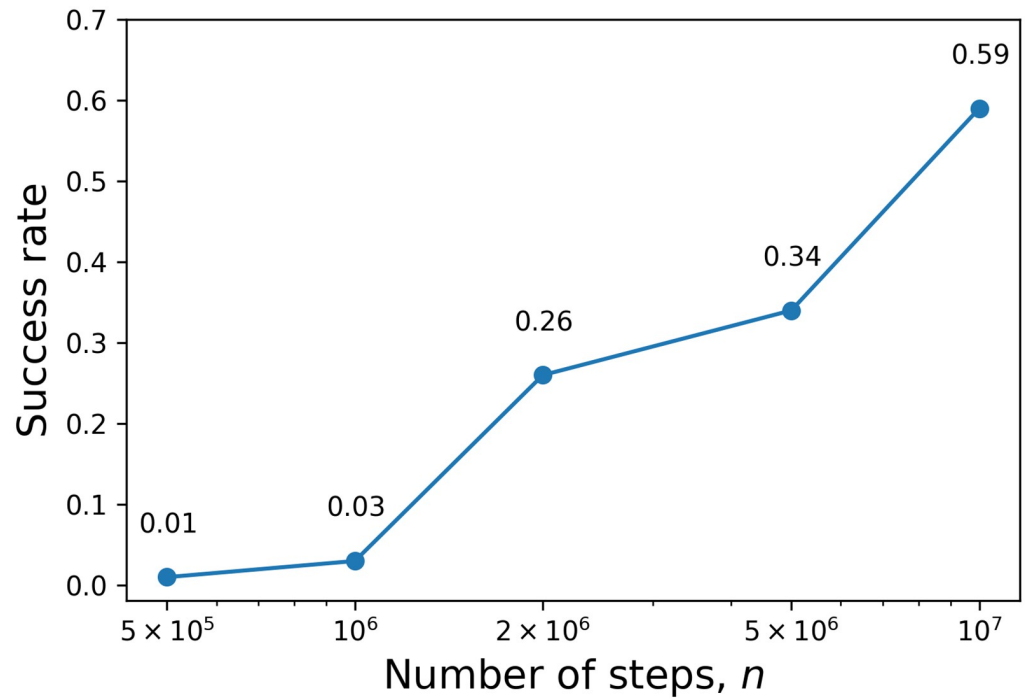


Fig 4. Success rate out of 100 trials in the SSA case.

<https://doi.org/10.1371/journal.pone.0273709.g004>

we did in Section 4. After we have determined n and A , we also need the success rate, i.e., the probability of finding the best-known Cut Value. For the G1 problem, out of 1000 trials, 61 were successful, so a rough estimation for the time needed to find the best value if the hyperparameters are known is $0.93 \text{ s} \cdot (1000/61) = 15.24 \text{ s}$.

In the case of SSA we need an annealing schedule [41], which tells us how the temperature decreases. We have chosen the geometrical cooling schedule, where $T_{k+1} = T_k \alpha$, with $\alpha \in (0, 1)$. The pseudocode of the SSA is

```

algorithm: Stochastic Simulated Annealing
input:  $E_S, T_{\text{init}}, \alpha, n$ 
output:  $\underline{S}$ 
 $\underline{S} :=$  random vector with  $\pm 1$  elements
for  $k := 0$  to  $n$ 
   $i :=$  random integer from  $[1, N]$ 
   $\Delta E_i := 2 \sum_j J_{ij} S_i S_j$ 
   $T := T_{\text{init}} \alpha^k$ 
  if  $\text{random}(0, 1) < \exp(-\Delta E_i / T)$ :
     $S_i := -S_i$ 
    
```

Here $\Delta E_i = E_{\underline{S}^{(i)}} - E_S$, and $\underline{S}^{(i)}$ is the same as \underline{S} except the i th spin is flipped. If the initial temperature is $T_{\text{init}} = 2 \max_i \Delta E_i$, then any initial trial i will be accepted with more than 50%, so it can be considered a sufficiently high temperature. The final temperature is $T_{\text{final}} = T_{\text{init}} \alpha^n$. Since in the G-sets the J_{ij} elements are -1, 0 or 1, the smallest positive energy difference is 1, hence it is reasonable to choose $T_{\text{final}} = 0.05$. Below that temperature it is unlikely to accept any state with higher energy than the current one. The parameter of the geometrical schedule is $\alpha = \sqrt[n]{T_{\text{final}}/T_{\text{init}}}$. The only remaining hyperparameter is n . Fig 4 shows the success rate for different n values for G1. In 15 seconds, 10^6 Monte Carlo steps can be carried out. It means if we

use $n = 2 \times 10^6$, then we need approximately 115 seconds to find the best known Cut Value of G1.

For the B&B method, we used the BiqCrunch implementation [40]. In one hour, the Cut Value it could find was 11524.

The success rate can be dissimilar for the other sets. For example, for G3, after the hyperparameter optimization in the case of the QMFA, it is 0.016, and in the case of SSA with $n = 5 \times 10^7$, it is 0.17. Even the running time is increased to $t_{\text{QMFA}} = (2.8 \pm 1.8)$ seconds, resulting in the expected time for finding the best known Cut Value with QMFA to 175 seconds. With SSA, this expected time is 455 seconds, and the B&B method found a solution with the Cut Value of 11538 in one hour.

In the examples mentioned above, the QMFA outperformed both the SSA and the B&B. Still, for example, for the G2 problem, the QMFA was only able to find the best known Cut Value once out of 2000 trials hence it would be inconsistent to calculate the average success time. On the other hand, the SSA found 5 times the best Cut Value out of 100 trials with $n = 5 \times 10^6$ and the expected success time is 1500 seconds. The B&B method found a state with CV = 11521 in one hour.

The G11 is special since the QMFA cannot find the best Cut Value, the SSA can find it on an average of 2500 seconds, but the B&B method needs only 512 seconds. This can be due to its simple structure since the G11 is a toroidal graph. On the other hand, an additional one hour is not enough for B&B to confirm that it is the best possible Cut Value. Since it is a toroidal graph (in this case, an 8×100 grid with periodic boundaries), we can solve it exactly with the transfer matrix method, confirming that 564 is indeed the best possible Cut Value.

Considering all of the facts above, we can state that no algorithm is universally dominant, but we can say the QMFA can find a good, in some cases the best, optimum in a relatively short time.

6 Conclusion and outlook

We have shown that the thermal and the quantum mean-field annealing are governed by similar equations, where the temperature and the tunneling rate play similar roles. This is due to the similarity of the entropic term in the free energy and the transverse term in the energy of the quantum approach. The sigmoid function in the quantum approach differs from the thermal one, and for large arguments, it changes like a power law as opposed to the exponential in the thermal case. As it was pointed out in [29] the specific nonlinearity of the sigmoid function affects the speed of the convergence of the mean-field approach. This can be one of the reasons for the relative success of the present approach in reproducing some of the best cut values for certain problems from the G-set. The approximation presented here is the most straightforward application of the variational method. The next step can be a more sophisticated version of the mean-field annealing that can be introduced by considering the correlations among the most coupled spins, which might lead to even better cut values in the future.

Author Contributions

Conceptualization: Gábor Vattay.

Formal analysis: Máté Tibor Veszeli.

Investigation: Máté Tibor Veszeli.

Supervision: Gábor Vattay.

Writing – original draft: Máté Tibor Veszeli.

Writing – review & editing: Gábor Vattay.

References

1. Nishimori H. "Statistical Physics of Spin Glasses and Information Processing: An Introduction." Oxford University Press (2001)
2. Mézard M., Parisi G., and Virasoro M., *Spin glass theory and beyond: An Introduction to the Replica Method and Its Applications*, vol. 9. World Scientific Publishing Company, 1987.
3. Ising E., "Beitrag zur Theorie des Ferromagnetismus," *Zeitschrift für Physik A Hadrons and Nuclei*, vol. 31, 02–04 1925.
4. Onsager L., "Crystal Statistics. I. A Two-Dimensional Model with an Order-Disorder Transition," *Phys. Rev.*, vol. 65, pp. 117–149, Feb 1944. <https://doi.org/10.1103/PhysRev.65.117>
5. Baxter R., *Exactly Solved Models in Statistical Mechanics*. Academic Press, 1982.
6. Barahona F., "On the computational complexity of Ising spin glass models," *Journal of Physics A: Mathematical and General*, vol. 15, no. 10, p. 3241, 1982. <https://doi.org/10.1088/0305-4470/15/10/028>
7. Rendl F., Rinaldi G., and Wiegele A., "A branch and bound algorithm for Max-Cut based on combining semidefinite and polyhedral relaxations," in *International Conference on Integer Programming and Combinatorial Optimization*, pp. 295–309, Springer, 2007.
8. Padberg M. and Rinaldi G., "A branch-and-cut algorithm for the resolution of large-scale symmetric traveling salesman problems," *SIAM review*, vol. 33, no. 1, pp. 60–100, 1991. <https://doi.org/10.1137/1033004>
9. Kochenberger G. A., Hao J.-K., Lü Z., Wang H., and Glover F., "Solving large scale max cut problems via tabu search," *Journal of Heuristics*, vol. 19, no. 4, pp. 565–571, 2013. <https://doi.org/10.1007/s10732-011-9189-8>
10. Glover F., Lü Z., and Hao J.-K., "Diversification-driven tabu search for unconstrained binary quadratic problems," *4OR*, vol. 8, no. 3, pp. 239–253, 2010. <https://doi.org/10.1007/s10288-009-0115-y>
11. Poljak S. and Rendl F., "Solving the max-cut problem using eigenvalues," *Discrete Applied Mathematics*, vol. 62, no. 1–3, pp. 249–278, 1995. [https://doi.org/10.1016/0166-218X\(94\)00155-7](https://doi.org/10.1016/0166-218X(94)00155-7)
12. Rendl F., Rinaldi G., and Wiegele A., "Solving max-cut to optimality by intersecting semidefinite and polyhedral relaxations," *Mathematical Programming*, vol. 121, no. 2, p. 307, 2010. <https://doi.org/10.1007/s10107-008-0235-8>
13. Goemans M. X. and Williamson D. P., "Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming," *Journal of the ACM (JACM)*, vol. 42, no. 6, pp. 1115–1145, 1995. <https://doi.org/10.1145/227683.227684>
14. Kirkpatrick S., Gelatt C. D., and Vecchi M. P., "Optimization by simulated annealing," *science*, vol. 220, no. 4598, pp. 671–680, 1983. <https://doi.org/10.1126/science.220.4598.671> PMID: 17813860
15. Černý V., "Thermodynamical approach to the traveling salesman problem: An efficient simulation algorithm," *Journal of Optimization Theory and Applications*, vol. 45, pp. 41–51, Jan 1985. <https://doi.org/10.1007/BF00940812>
16. Isakov S. V., Zintchenko I. N., Rønnow T. F., and Troyer M., "Optimised simulated annealing for Ising spin glasses," *Computer Physics Communications*, vol. 192, pp. 265–271, 2015. <https://doi.org/10.1016/j.cpc.2015.02.015>
17. Kadowaki T., Nishimori H., "Quantum annealing in the transverse Ising model," *Physical Review E*, 1998 Nov 1; 58(5):5355. <https://doi.org/10.1103/PhysRevE.58.5355>
18. Tanaka K. and Horiguchi T., "Probabilistic, iterated and quantum-iterated computational methods in gray-level image restoration," *Interdisciplinary Information Sciences*, vol. 8, no. 1, pp. 33–50, 2002. <https://doi.org/10.4036/iis.2002.33>
19. Suzuki S., Nishimori H. "Quantum annealing by ferromagnetic interaction with the mean-field scheme." *Physica A: Statistical Mechanics and its Applications*. 2007 Oct 1; 384(1):137–40. <https://doi.org/10.1016/j.physa.2007.04.077>
20. Steffen M., van Dam W., Hogg T., Breyta G., and Chuang I., "Experimental implementation of an adiabatic quantum optimization algorithm," *Physical Review Letters*, vol. 90, no. 6, p. 067903, 2003. <https://doi.org/10.1103/PhysRevLett.90.067903> PMID: 12633331
21. Johnson M. W., Amin M. H., Gildert S., Lanting T., Hamze F., Dickson N., et al., "Quantum annealing with manufactured spins," *Nature*, vol. 473, no. 7346, p. 194, 2011. <https://doi.org/10.1038/nature10012> PMID: 21562559

22. Harris R., Sato Y., Berkley A., Reis M., Altomare F., Amin M., et al., "Phase transitions in a programmable quantum spin glass simulator," *Science*, vol. 361, no. 6398, pp. 162–165, 2018. <https://doi.org/10.1126/science.aat2025> PMID: 30002250
23. Farhi E., Goldstone J., Gutmann S., and Sipser M., "Quantum computation by adiabatic evolution," *arXiv preprint quant-ph/0001106*, 2000.
24. Roland J. and Cerf N. J., "Quantum search by local adiabatic evolution," *Physical Review A*, vol. 65, no. 4, p. 042308, 2002. <https://doi.org/10.1103/PhysRevA.65.042308>
25. Hamerly R., Inagaki T., McMahon P. L., Venturelli D., Marandi A., Onodera T., et al., "Experimental investigation of performance differences between coherent Ising machines and a quantum annealer," *Science advances*, vol. 5, no. 5, p. eaau0823, 2019. <https://doi.org/10.1126/sciadv.aau0823> PMID: 31139743
26. Bilbro G., Mann R., Miller T., Snyder W., van den Bout D., and White M., "Optimization by mean field annealing," *Advances in neural information processing systems*, vol. 1, pp. 91–98, 1988.
27. Inagaki T., Haribara Y., Igarashi K., Sonobe T., Tamate S., Honjo T., et al., "A coherent Ising machine for 2000-node optimization problems," *Science*, vol. 354, no. 6312, pp. 603–606, 2016. <https://doi.org/10.1126/science.aah4243> PMID: 27811271
28. Haribara Y., Utsunomiya S., and Yamamoto Y., "A coherent Ising machine for MAX-CUT problems: performance evaluation against semidefinite programming and simulated annealing," in *Principles and Methods of Quantum Information Technologies*, pp. 251–262, Springer, 2016.
29. Böhm F., Van Vaerenbergh T., Verschaffelt G., Van der Sande G., "Order-of-magnitude differences in computational performance of analog Ising machines induced by the choice of nonlinearity" *Communications Physics* vol. 4, no. 1, pp 1–11, 2021.
30. Helmberg C. and Rendl F., "A spectral bundle method for semidefinite programming," *SIAM Journal on Optimization*, vol. 10, no. 3, pp. 673–696, 2000. <https://doi.org/10.1137/S1052623497328987>
31. Tatsumura K., Yamasaki M. and Goto H., "Scaling out Ising machines using a multi-chip architecture for simulated bifurcation" *Nature Electron* 4, 208–217 (2021). <https://doi.org/10.1038/s41928-021-00546-4>
32. Kato T., "On the Adiabatic Theorem of Quantum Mechanics," *Journal of the Physical Society of Japan*, vol. 5, no. 6, pp. 435–439, 1950. <https://doi.org/10.1143/JPSJ.5.435>
33. Lucas A., "Ising formulations of many NP problems," *Frontiers in Physics*, vol. 2, p. 5, 2014.
34. Karp R. M., "Reducibility among combinatorial problems," in *Complexity of computer computations*, pp. 85–103, Springer, 1972.
35. Shao S., Zhang D., and Zhang W., "A simple iterative algorithm for maxcut," *arXiv preprint arXiv:1803.06496*, 2018.
36. Ma F. and Hao J.-K., "A multiple search operator heuristic for the max-k-cut problem," *Annals of Operations Research*, vol. 248, no. 1–2, pp. 365–403, 2017. <https://doi.org/10.1007/s10479-016-2234-0>
37. Festa P., Pardalos P. M., Resende M. G., and Ribeiro C. C., "Randomized heuristics for the MAX-CUT problem," *Optimization methods and software*, vol. 17, no. 6, pp. 1033–1058, 2002. <https://doi.org/10.1080/1055678021000090033>
38. Wang T., Wu L., and Roychowdhury J., "New computational results and hardware prototypes for oscillator-based Ising machines," in *Proceedings of the 56th Annual Design Automation Conference 2019*, pp. 1–2, 2019.
39. Matsuda Y., "Benchmarking the MAX-CUT problem on the Simulated Bifurcation Machine," Dec 2019.
40. Krislock N., Malick J., and Roupin F., "Biqcrunch: A semidefinite branch-and-bound method for solving binary quadratic problems," *ACM Transactions on Mathematical Software (TOMS)*, vol. 43, no. 4, pp. 1–23, 2017. <https://doi.org/10.1145/3005345>
41. Mahdi W., Medjahed S. A., and Ouali M., "Performance analysis of simulated annealing cooling schedules in the context of dense image matching," *Computación y Sistemas*, vol. 21, no. 3, pp. 493–501, 2017.