On the hardness of quadratic unconstrained binary optimization problems

V. Mehta and F. Jin

Institute for Advanced Simulation, Jülich Supercomputing Centre,
Forschungszentrum Jülich, D-52425 Jülich, Germany

K. Michielsen

Institute for Advanced Simulation, Jülich Supercomputing Centre,
Forschungszentrum Jülich, D-52425 Jülich, Germany, and
AIDAS, 52425 Jülich, Germany, and
RWTH Aachen University, 52056 Aachen, Germany

H. De Raedt

Institute for Advanced Simulation, Jülich Supercomputing Centre,
Forschungszentrum Jülich, D-52425 Jülich, Germany and
Zernike Institute for Advanced Materials,
University of Groningen, Nijenborgh 4,
NL-9747 AG Groningen, Netherlands

(Dated: June 24, 2022)

We use exact enumeration to characterize the solutions of quadratic unconstrained binary optimization problems of less than 21 variables in terms of their distributions of Hamming distances to close-by solutions. We also perform experiments with the D-Wave Advantage 5.1 quantum annealer, solving many instances of up to 170-variable, quadratic unconstrained binary optimization problems. Our results demonstrate that the exponents characterizing the success probability of a D-Wave annealer to solve a QUBO correlate very well with the predictions based on the Hamming distance distributions computed for small problem instances.

I. INTRODUCTION

Optimization is at the heart of problem solving in science, engineering, finance, operational research etc. The basic idea is to associate a cost with each of the possible values of the variables that describe the problem and try to minimize this cost. Of particular importance is the class of so-called discrete optimization problems in which some or all the variables take values from a finite set of possibilities. Discrete optimization problems are often NP-hard [1] which, in practice and in simple terms, means that solving such a problem on a digital computer will require resources that increase exponentially with the number of variables.

Many discrete optimization problems can be reformulated as quadratic unconstrained binary optimization (QUBO) problems [2,3]. Solving a QUBO amounts to finding the values of the N binary variables $x_i = 0, 1$ that minimize the cost function

$$\text{Cost}(x_1, \ldots, x_N) = \sum_{1=i \le j=N} Q_{i,j} x_i x_j \quad, \quad x_i = 0, 1 \;, \tag{1}$$

where $Q_{i,j} = Q_{j,i}$ is a symmetric $N \times N$ matrix of floating point numbers.

The interest in expressing discrete optimization problems in the form of QUBOs has recently gained momentum by the development of quantum annealers manufactured by D-Wave systems [4,5]. In theory [6], quantum annealers make use of the adiabatic theorem [7] to find the ground state of the Ising model defined by the Hamiltonian

$$H = \sum_{1=i<j=N}^{N} J_{i,j} S_i S_j + \sum_{i=1}^{N} h_i S_i \;, \tag{2}$$

Substituting $S_i = 1 - 2x_i = \pm 1$ in Eq. (2) yields

$$H = \text{Cost}(x_1, \ldots, x_N) - C_0 \;, \tag{3}$$

where the relations between the Q's in Eq. (1) and J's, h's, and C_0 in Eqs. (2) and (3) are given in appendix A.

Obviously, the transformation $S_i = 1 - 2x_i = \pm 1$ does not change the nature of the optimization problem. In other words, minimizing the cost function of a QUBO Eq. (1) or finding the ground state of the Ising model Eq. (2) are equally difficult. Currently available quantum annealer hardware often finds ground states of Eq. (2) for fully-connected problems with about 200 variables or less in about micro seconds [4,5], which is quite fast, suggesting that as larger quantum annealers become available, they have the potential to solve large QUBOs in a relatively short real time.

Although the transformation $S_i = 1 - 2x_i = \pm 1$ does not change the nature of the optimization problem, the formulation of a particular optimization problem in terms of a QUBO can. For instance, 2-satisfiability problems (2SAT) [1] (see section II A) can be solved with computational resources that increase linearly with the number of binary variables [8–10]. However, the efficient solution of 2SAT-problem that permits its efficient solution are lost when it is expressed as a QUBO/Ising model. In fact, the equivalent QUBO becomes notoriously hard to solve by e.g., simulated annealing [11]. On the other hand, it is also not difficult to construct Ising models of which the ground state is very easy to find.

In view of the potential of quantum annealers for solving large QUBOs in the near future, it is of interest to gain some insight into the degree of success by which a quantum annealer is expected to solve a QUBO/finding the ground state of corresponding Ising model, without actually performing the experiment.

In this paper, we show that there are at least three different classes of QUBOs that distinguish themselves by:

1. The success probability a D-Wave annealer finds the ground state of the Ising model/solves the QUBO.

2. The distribution of Hamming distances between the ground state and the lowest excited states computed for relatively small, representative problem instances.

We demonstrate that the differences between the Hamming distance distributions are correlated with the size-dependent scaling of the success probabilities with which D-Wave quantum annealers find the ground state.

The structure of the paper is as follows. Section II introduces the three different classes of QUBO problems that we analyze. In section III, we briefly review the three different methods by which we solve the QUBO instances. Sections IV and V present and discuss our results for the Hamming distance and level spacing distributions, and quantum annealing experiments, respectively. In section VI, we summarize our findings.

II. QUADRATIC UNCONSTRAINED BINARY OPTIMIZATION PROBLEMS

A. 2-Satisfiability problems

The problem of assigning values to binary variables such that given constraints on pairs of variables are satisfied is called 2-satisfiability [1]. 2SAT is a special case of the general Boolean satisfiability problem, involving constraints on more than two variables. In contrast to e.g., 3SAT which is known to be NP-complete, 2SAT can be solved in polynomial time. The most efficient algorithms solve 2SAT in a time which is proportional to the number of variables [8–10].

A 2SAT problem is specified by N binary variables $x_i = 0, 1$ and a conjunction of M clauses defining a binary-valued cost function

$$
\begin{aligned}
C &= C(x_1, \ldots, x_N) \\
&= (L_{1,1} \vee L_{1,2}) \wedge (L_{2,1} \vee L_{2,2}) \wedge \ldots \wedge (L_{M,1} \vee L_{M,2}) ,
\end{aligned}
\tag{4}
$$

where the literal $L_{\alpha,j}$ stands for either $x_{i(\alpha,j)}$ or its negation $\bar{x}_{i(\alpha,j)}$, for $\alpha = 1, \ldots, M$ and $j = 1, 2$. The function $i(\alpha, j)$ maps the pair of indices (α, j) onto the index i of the binary variable x_i. A 2SAT problem is satisfiable if one can find at least one assignment of the x_i's which makes the cost function C true.

Solving a 2SAT problem is equivalent to finding the ground state of the Ising-spin Hamiltonian [2,11,12]

$$
H_{2\mathrm{SAT}} = \sum_{\alpha=1}^{M} h_{2\mathrm{SAT}}(e_{\alpha,1} S_{i(\alpha,1)}, e_{\alpha,2} S_{i(\alpha,2)}) ,
\tag{5}
$$

where $e_{\alpha,j} = +1(-1)$ if $L_{\alpha,j}$ stands for $x_i(\bar{x}_i)$ and

$$
h_{2\mathrm{SAT}}(S_l, S_m) = (S_l - 1)(S_m - 1) \quad , \quad S_m, S_l = \pm 1 .
\tag{6}
$$

Grouping and rearranging terms, Eq. (5) reads

$$H_{2\text{SAT}} = \sum_{1 \leq i < j \leq N} J_{i,j} S_i S_j + \sum_{i=1}^{N} h_i S_i + C_1 \,, \tag{7}$$

where C_1 is an irrelevant constant. Therefore, solving a 2SAT problem Eq. (4) is equivalent to solving the QUBO problem defined by Eq. (7). It may be of interest to mention that if one is given a QUBO problem without knowing that it originated from a 2SAT problem, it is not clear that the QUBO can be solved in a time linear in N.

Constructing 2SAT problems is easy but finding 2SAT problems that have a unique known ground state and a highly degenerate first-excited state quickly becomes more difficult with increasing N [11]. We have generated 15 sets of 2SAT problems that have a unique ground state and a highly degenerate first-excited state, each set corresponding to an N, with N ranging from 6 to 20 with the number of clauses $M = N + 1$. The graphs representing these problems are not fully connected, meaning that not all $J_{i,j}$'s are different from zero. For our sets of 2SAT problems, the h_i's and $J_{i,j}$'s can take the integer values between -2 and 2.

B. Fully-connected spin glass

The spin glass model is defined by the Hamiltonian Eq. (2). Computing the ground state configuration is, in general, very hard. To verify that a spin configuration has the lowest energy, one would (in general) have to go through all the $2^N - 1$ other configurations to check if it indeed has the lowest energy. The qualifier "in general" is important here for there are cases, such as when all $J_{i,j} = 0$, for which the ground state is trivial to find. To effectively rule out such trivially solvable problems, we use uniform (pseudo) random numbers in the range $[-1, +1]$ to assign values to all the $J_{i,j}$'s and all the h_i's. The probability that one of the $J_{i,j}$'s is zero is extremely small, justifying the term "fully-connected spin glass". In the following, we refer to the set of model instances generated in this manner as RAN models.

C. Fully-connected regular spin-glass model

In the course of developing an QUBO-based application to benchmark large clusters of GPUs (see appendix B), we discovered by accident that the ground state of the spin glass defined by Eq. (2) with

$$\begin{aligned} J_{i,j} &= 1 - (i + j - 2)/(N - 1) \quad, \quad i \neq j \,, \\ h_i &= 1 - 2(i - 1)/(N - 1) \,, \end{aligned} \tag{8}$$

seems to have a peculiar structure. Note that of the order of N $J_{i,j}$'s are zero.

Although we have not been able to give a proof valid for all N, up to $N = 200$ we have not found any counter example for conjecture that the ground states of the Ising model with parameters given by Eq. (8) is given by $(S_1 = -1, \ldots, S_k = -1, S_{k+1} = 1, \ldots, S_N = 1)$ or, equivalently $(x_1 = 1, \ldots, x_k = 1, x_{k+1} = 0, \ldots, x_N = 0)$ where k is the integer that minimizes

$$f(k) = \sum_{1 \le i \le j \le k} Q_{i,j} = \frac{k(N-k)(N-2k+2)}{N-1} \ . \tag{9}$$

Thus, although our conjecture is correct, its solution is very easy to find for any N. We refer to the special, fully-connected regular spin glass problems defined by Eq. (8) as REG problems.

From Eq. (2), it immediately follows that randomly reversing a spin i and replacing h_j by $-h_j$ and $J_{i,j}$ by $-J_{i,j}$ for all j does not change the ground state energy of a QUBO problem. Applying such "gauge transformation" to a set of randomly selected spins generated a set of REG problems that are mathematically equivalent. This feature, in combination with the fact that the ground state is known (for at least $N \le 200$) makes REG problems well-suited for testing and benchmarking purposes, of both conventional and quantum hardware.

III. METHODS FOR SOLVING QUBOS

We solve QUBOs using three different methods.

1. A computer code referred to as **QUBO22** which uses GPUs and/or CPUs to solve QUBOs, Polynomial Unconstrained Binary Optimization (PUBOs), and Exact Cover problems (reformulated as QUBOs). QUBO22 simply enumerates all 2^N possible values of the binary variables $x_1, \ldots, x_N$ while keeping track of those configurations of x's that yield the lowest, next to lowest and largest cost. QUBO22 obviously always finds the true ground state. The number of arithmetic operations required to solve a QUBO is proportional to $N(N-1)2^N$. With the supercomputers that are available to us, the exponential increase with N limits the application QUBO22 to problems of size $N \le 56$.

   In order to compute the Hamming distance between excited states and the ground state and level spacing distributions, it is necessary to keep track of a large number of different states. To this end, we use another code, also based on full enumeration, which in practice, can readily handle problems up to $N = 20$.

2. Heuristic methods can solve QUBOs in (much) less time than QUBO22 can. However, heuristic methods do not guarantee to return the solution of the QUBO problem (although they very often do). In our work, we use **qbsolv**, a heuristic solver provided by D-Wave, to compute the ground states of all problem instances. For those problems which QUBO22 can solve, the ground states obtained by qbsolv and QUBO22 match. For all RAN and REG problems up to $N = 200$, the ground states obtained by qbsolv and the D-Wave Advantage 4.1 Hybrid solver are also the same.

3. We have used the D-Wave Advantage 5.1 quantum annealer to solve all problem instances. We calculate the success probabilities by using the ground states obtained by **qbsolv**. The data for the success probabilities is then used to analyse the scaling behavior as a function of the problem size N.
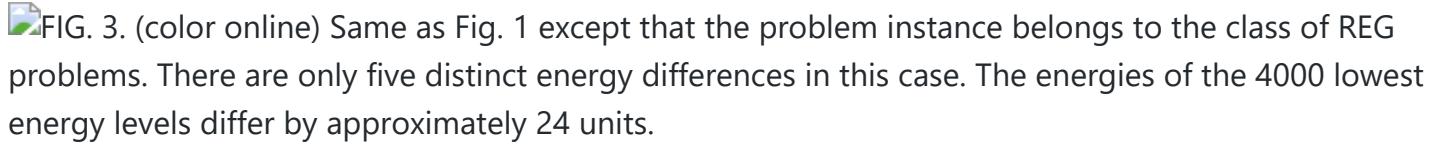
IV. HAMMING DISTANCE AND LEVEL SPACING

The Hamming distance between two bitstrings (or strings of S's) of equal length is defined as the number of positions at which the corresponding bits (S's) are different. For each of the problems in our set, we use exact enumeration to find at most 6037 states with the lowest energies. With this data we compute the Hamming distances between the ground state and these excited states.

In this section, we only present results for one representative $N = 20$ problem taken from the 2SAT, RAN and REG class, respectively. The plots for other $N = 20$ instances look similar.

From the analysis we draw the following conclusions. Recall that the first excited states used to generate Fig. 1(a) all have the same energy. By construction, according to Eq. (6), the level spacing distribution Fig. 1(b) is nonzero for $\Delta = 4$ only. Now imagine that the search process (in e.g., simulated annealing) for the ground state ends up in one of these excited states. From Fig. 1(a) it then follows immediately that the probability to reach the ground state by single-spin flipping will be very low. Indeed, most of these excited states have a Hamming distance 10-15 and it would require a miracle to have a particular sequence of single-spin flips reducing the Hamming distance to zero. In summary, from Fig. 1 it is easy to understand why this particular class of 2SAT problems is very hard to solve by simulated annealing [11,12].

For RAN problems, we conclude the following. In contrast to 2SAT problems, most of the weight of the Hamming distance distribution is centered around five. Also the level spacing distribution is very different from that of the 2SAT problems. This suggests that five or less spin flips may suffice to change the excited state into the ground state. Thus, in comparison with the class of 2SAT problems that we have selected, the RAN problems are expected to be much more amenable to simulated and quantum annealing.

FIG. 3. (color online) Same as Fig. 1 except that the problem instance belongs to the class of REG problems. There are only five distinct energy differences in this case. The energies of the 4000 lowest energy levels differ by approximately 24 units.

The Hamming distance and level spacing distribution of a REG problem are very different from the corresponding ones of the 2SAT or RAN problems. There are only five distinct energy level spacings for these problems and the Hamming distance distribution shows that many of the excited states differ from the ground state by only a few spin flips. Therefore, we may expect that of the three classes of problems considered, problems of the REG class are the least difficult to solve by simulated or quantum annealing.

V. QUANTUM ANNEALING EXPERIMENTS

We demonstrate that the conclusions of section IV, drawn from the analysis of small problem instances, correlate very well with the degree of difficulty observed when solving significantly larger problems on a D-Wave Advantage 5.1 quantum annealer.

For REG-class problems solved on a D-Wave quantum annealer, for each N, instances were generated by spin-reversal transformations, as explained above. As the problem size N increases, the success probability, that is the relative frequency with which the D-Wave yields the ground state, decreases exponentially, from $\mathcal{O}(1)$ to $\mathcal{O}(10^{-6})$. Fitting exponentials to the data reveals that the exponent changes from -0.039 to -0.090 at about $N = 80$. The larger the absolute value of the exponent, the more difficult it is to solve the QUBO by quantum annealing.

With increasing problem size N, it becomes more difficult and eventually impossible to map the fully-connected QUBO problem onto the Chimera or Pegasus lattice that defines the connectivity of the D-Wave qubits. Even a small fully-connected problems cannot be embedded on the D-Wave Advantage 5.1 Pegasus lattice without replacing logical bits by chains of physical qubits. For instance, a $N = 170$ variable REG problem maps onto about 3964 physical qubits of a D-Wave Advantage 5.1. We quantify this aspect by computing the average chain length, a measure for the average number of physical qubits that the D-Wave software uses to map a variable onto a group of physical qubits. The average chain length, computed from data obtained by solving all REG problems, only increases linearly with N and does not show any sign of the crossover observed in the scaling dependence of the success probabilities.

For RAN-class problems solved on a D-Wave quantum annealer, there are significant differences compared to REG problems. First note that there is no data point for $N = 90$ simply because after 200000 attempts, the D-Wave did not return the ground state. For $N = 100$, we were more lucky but for $N \geq 110$ we were not, in sharp contrast with the case of REG problems for which the D-Wave Advantage 5.1 returned the correct solutions up to $N = 170$. Clearly, the D-Wave Advantage 5.1 quantum annealer finds RAN problems more difficult to solve than REG problems, although both their QUBOs involve fully connected graphs. This observation is further confirmed by fitting exponentials to the data. As in the REG case, there is a crossover, not at $N = 80$ but at $N \approx 35$, with the exponent changing from -0.102 to -0.162.

The class of 2SAT problems analyzed in this paper has already been studied extensively through computer simulated quantum annealing and by means of D-Wave quantum annealers [13,14]. The general conclusion is that, in spite of the small number ($N \leq 20$) of variables, the 2SAT class contains instances which are very difficult to solve by simulated annealing [11] and quantum annealing [13,14]. For completeness, we present data for the success probabilities for the set of $N \leq 20$ problems, also used to study the Hamming distance and level spacing distributions. Focusing on problem of size $N = 14, \ldots, 20$, we find that the success probabilities, obtained by the D-Wave Advantage 5.1 quantum annealer, decrease exponentially with an exponent of -0.269, considerable larger in absolute value than the exponents -0.162 and -0.090 of the corresponding RAN and REG problems, respectively. Of course, we cannot simply extrapolate the $14 \leq N \leq 20$ exponent to larger N's but, we believe it is unlikely that this exponent will become smaller as N increases further.

For our D-Wave experiments, we used the default annealing time of $20\mu s$. One may expect that by using longer annealing times, the success probabilities will increase and more and larger RAN class problems can be solved.

VI. CONCLUSION

We have analyzed a large number of QUBOs that we have synthesized in three different ways. The first set of QUBOs was obtained by mapping a special selection of 2SAT problems onto QUBOs. These 2SAT problems are special in the sense that they possess a unique ground state and a large number of degenerate, first excited states. Findings such 2SAT problems is computationally demanding, limiting our search to instances with less than 21 variables.

The second set of problems contains Ising models in which uniform random numbers determine all the two-spin interactions and all the local fields. From a statistical mechanics point of view, such fully-connected spin-glass models exhibit frustration and computing their ground states and temperature dependent properties is known to be difficult.

Finally, the third set of problems is also of the fully-connected spin-glass type but the interaction and field values are given by a peculiar linear function of the spin indices. Our numerical experiments suggest that this problem may be solvable for any number of spins but we have not yet been able to proof this conjecture mathematically.

We have calculated Hamming distance distributions and level spacing distributions for small problem instances and also submitted the small and large QUBO instances to a D-Wave quantum annealer. Our results demonstrate that the exponents characterizing the success probability of a D-Wave annealer to solve a QUBO correlate very well with the predictions based on the Hamming distance and level spacing distributions computed for small QUBO instances.

It would be of interest to see if Hamming distance distributions for small problem instances also predict the effectiveness of simulated annealing.

ACKNOWLEDGEMENTS

Appendix A: Relation between QUBOs and Ising models

The relations between the $Q_{i,j}$'s, $J_{i,j}$'s, h_i's and C_0 are given by

$$J_{i,j} = \frac{1}{4}Q_{i,j} \quad \text{if} \quad i \neq j \quad , \quad J_{i,i} = 0 \; , \tag{A1}$$

$$h_i = -\frac{1}{2}\left(Q_{i,i} + \frac{1}{2}\sum_{j \neq i}Q_{i,j}\right) \; , \tag{A2}$$

$$C_0 = \frac{1}{2} \sum_{i=1}^{N} Q_{i,i} + \frac{1}{4} \sum_{1 \le i < j \le N} Q_{i,j}$$

$$= \sum_{1 \le i < j \le N} J_{i,j} - \sum_{i=1}^{N} \left[h_i + \sum_{j \ne i} J_{i,j} \right], \tag{A3}$$

$$Q_{i,j} = 4 J_{i,j} - 2 \delta_{i,j} \left(h_i + \sum_{k \ne i} J_{i,k} \right). \tag{A4}$$