A quantum annealing protocol to solve the maximum clique problem

Author: Alejandro García Rivas.
Facultat de Física, Universitat de Barcelona, Diagonal 645, 08028 Barcelona, Spain.

Advisor: Bruno Juliá Díaz

Abstract: We give a short introduction to quantum annealing in the transverse Ising model and formulate the maximum clique problem in this setting. We then briefly show how problems are actually implemented in D-Wave's annealers through the idea of minor-embeddings and chains. We also provide simulations for a small problem and actual results on D-Wave's annealers for larger instances as well as the dependence of the success rate on different parameters.

I. INTRODUCTION

The standard model of quantum computation, the quantum gate model, has been seen to offer speedups to classical algorithms as well as efficiently solve a larger class of problems than classical models [1]. There are, however, many other computational models that are also based on quantum effects. One of them is adiabatic quantum computation (AQC), which is extensively described in [2] and shown to be equivalent to the quantum gate model in [3] in the sense that one model can simulate the other in polynomial time and viceversa. Despite this fact, the nature of these two models looks so different that it is not strange to expect some particular type of problems to be easily solved within one model but not the other.

In this work, we begin in section II by briefly explaining AQC and its heuristic implementation, quantum annealing (QA), as well as giving an Ising formulation to the maximum clique problem. Then, in section III, we focus on explaining how problems are actually implemented in D-Wave's quantum annealers (cf. [4]) using a particular simple problem. Section IV is devoted to using D-Wave's quantum annealers to solve large instances of the maximum clique problem for which classical simulation of the annealing is no longer possible. Finally, section V provides a summary of the results and the conclusions of this work.

II. AQC AND QA ALGORITHMS

Adiabatic quantum computation consists of the following idea: *consider a pair of Hamiltonians H_0 and H_1 and functions $A, B : [0,1] \to \mathbb{R}$ (called schedule functions) satisfying $A(1) = B(0) = 0$ and $A(0), B(1) > 0$. Assume also that the ground state $|\psi_0\rangle$ of H_0 is non-degenerate. Then, by the adiabatic theorem [2, 5], for sufficiently large times t_f, the system with Hamiltonian*

$$H(t) = A(t/t_f)H_0 + B(t/t_f)H_1 \tag{1}$$

evolves from $|\psi_0\rangle$ to a state which is reasonably close to a ground state of H_1.

The adiabatic theorem thus opens an avenue to solve problems whose solution can be identified with the ground state of some Hamiltonian H_1. The general AQC algorithm consists of finding such H_1, preparing $|\psi_0\rangle$, making the system evolve with $H(t)$ and performing a measurement at the end of the evolution.

On the other hand, following [6], a QA algorithm pursues the same idea, but the final Hamiltonian H_1 is restricted to represent a classical objective function. In particular, QA algorithms solve combinatorial optimization problems. In practice, however, an actual physical implementation would take place in an open system susceptible to all kinds of noises that would cause non-adiabatic jumps yielding a final state far from the desired solution. Despite this, one might still hope to obtain reasonable solutions as long as the interactions with the environment are kept small. This is why QA is sometimes called an heuristic instead of an actual algorithm when referring to the physical realization.

We now restrict ourselves to QA in the transverse Ising model. This consists of further restrictions on the Hilbert space and the Hamiltonians H_0 and H_1. Namely, consider a system of n 1/2-spins and define Hamiltonians

$$H_0 = -\sum_{i=1}^{n} \sigma_x^{(i)}, \tag{2}$$

$$H_1 = c + \sum_{i=1}^{n} h_i \sigma_z^{(i)} + \sum_{i>j} J_{ij} \sigma_z^{(i)} \sigma_z^{(j)}, \tag{3}$$

where $\sigma_z^{(i)}, \sigma_x^{(i)}$ denote Pauli matrices acting on the ith spin and c, h_i, J_{ij} are real coefficients. A Hamiltonian of the form Eq. (3) is called an Ising Hamiltonian. Consider also the basis $\{|z_1 \dots z_n\rangle\}$, where $|z_i\rangle = |\pm 1\rangle$ denotes the up or down state in the z direction and $\sigma_z^{(i)}|z_1 \dots z_n\rangle = z_i|z_1 \dots z_n\rangle$. Notice then that H_1 is diagonal in this basis (it thus represents a classical objective function) and that $|\psi_0\rangle = \frac{1}{2^{n/2}}(1, \dots, 1)$ is the non-degenerate ground state of H_0. QA consists then of evolving the system Hamiltonian according to Eq. (1) for some schedule functions A, B and performing a measurement at the end.

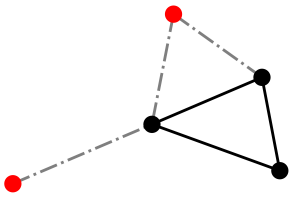

FIG. 1: Simple 5-vertex graph. A maximum clique is indicated by black nodes and solid edges.

A. Ising formulation of the maximum clique problem

A k-clique of a graph is a subgraph consisting of k vertices that are all neighbors of each other. The clique problem (CP) asks:

Given a graph, does it have a k-clique?

Notice that this is a decision problem, i.e. it is only necessary to answer 'yes' or 'no' and it is not required to find such a k-clique. This is one of the 21 Karp's NP-complete problems [7], i.e. it is as hard as the hardest NP problems in the sense that any other NP problem can be efficiently mapped to it. In other words, finding efficient algorithms to solve CP, also gives efficient algorithms to any other NP problem. Apart from this property, clique problems are also interesting by themselves as they are naturally encountered in many different topics such as computational chemistry [8] and signal processing [9]. In this work, though, we focus on solving a harder version of CP, namely the maximum clique problem (MCP):

Given a graph, find its largest clique.

This is an NP-hard problem because it solves its corresponding NP-complete decision version (i.e. CP) for every natural k.

Many NP problems have an Ising formulation [10], i.e. a Hamiltonian in Ising form (3) whose ground states exactly encode the solutions to the problem. The MCP is not an exception and, in fact, its Ising formulation is particularly simple and is a good representative of how can one encode solutions as ground states.

It goes as follows: given a graph, denote the set of vertices by $V = \{1, \dots, n\}$ and the set of edges by E. We propose working with a system of n spins. It is convenient to define $\rho_z^{(i)} = (1 + \sigma_z^{(i)})/2$. A subgraph is encoded in a state $|z_1 \dots z_n\rangle$ by the rule that it contains vertex i if, and only if, $z_i = +1$. Then, the Hamiltonian

$$H_1 = -\alpha \sum_{i=1}^{n} \rho_z^{(i)} + \beta \sum_{(i,j)\notin E} \rho_z^{(i)} \rho_z^{(j)} \tag{4}$$

does the job as long as $\alpha < \beta$ [11]. For the sake of concreteness, we just choose $\alpha = 1$ and $\beta = 2$. To give an actual expression in Ising form (3), we still need to develop Eq. (4). This gives:

$$h_i = \frac{\beta(n - 1 - e_i) - 2\alpha}{4}, J_{ij} = \frac{\beta g_{ij}}{4}, \tag{5}$$

where e_i is the number of neighbours of vertex i and $g_{ij} = 0$ if $(i, j) \in E$ and $g_{ij} = 1$ otherwise.

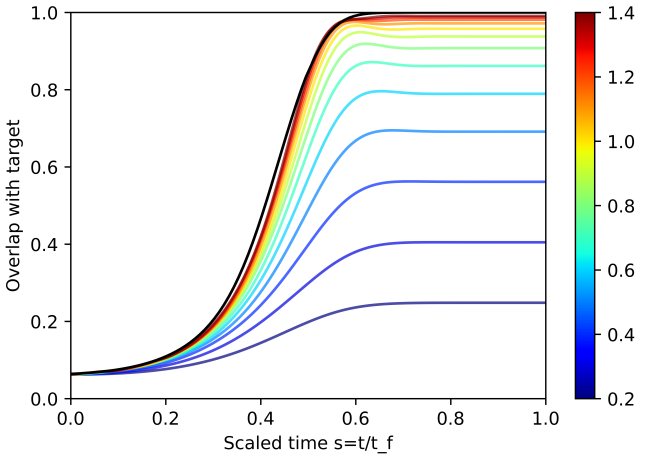

FIG. 2: Overlap of the instantaneous state with the target states throughout the evolution parametrized by $s = t/t_f$. The black line represents the ideal adiabatic evolution. Color corresponds to values of t_f ranging from 0.2 ns to 1.4 ns. Also, motivated by the behaviour of D-Wave's annealers, we have used actual schedule functions A, B for the DW_2000Q_6 annealer (cf. [12]) and coefficients h_i, J_{ij} are scaled according to section III B.

B. Numerical simulation

The simulation of the evolution is performed by numerically solving the time-dependent Schrödinger equation

$$i\hbar \frac{\partial |\psi(t)\rangle}{\partial t} = H(t)|\psi(t)\rangle. \tag{6}$$

We have used the Crank-Nicolson algorithm, as it is unconditionally stable. Namely, we set

$$|\psi(t + \Delta t)\rangle = \left(1 + i\frac{\Delta t}{2\hbar} H\right)^{-1} \left(1 - i\frac{\Delta t}{2\hbar} H\right) |\psi(t)\rangle \tag{7}$$

and iterate.

To exemplify the procedure, we present now simulation results (cf. [13]) for the 5-vertex graph of Fig. 1. In Fig. 2 we show the evolution of the overlap with target $\sum_{|\varphi\rangle} |\langle\psi(s)|\varphi\rangle|^2$, where $|\varphi\rangle$ ranges over all ground states of H_1 and $|\psi(s)\rangle$ is the instantaneous state parametrized by the scaled time $s = t/t_f$. We can see that larger annealing times t_f fit the ideal adiabatic case (black line) and give a better overlap at the end as well. Since the schedule functions A and B take maximum values of 6 GHz, we see that success probabilities > 0.95 are achieved when the dimensionless quantity $t_f \cdot \max\{A(t), B(t)\}$ is of the order of 10.

III. D-WAVE'S QUANTUM ANNEALERS

D-Wave offers to the public one minute of computation time for both of their quantum computers: DW_2000Q_6 and its successor, Advantage_system4.1. The main differences between these two are the number of available qubits (~ 2000 for the former and ~ 5000 for the latter) and the underlying graph in which the Ising Hamiltonian is embedded: DW_2000Q_6 is based on the Chimera graph and Advantage_system4.1 on the Pegasus graph, which has far more connections (cf. [14]).

A. The idea of minor-embeddings and chains

Physical realizations of QA thus have to deal with a problem we have not yet spoken about, namely, that of not having couplings between each pair of spins. D-Wave solves this by chaining physical spins (i.e. actual spins in the annealer) together and make them represent the same logical spin. To be more precise, we need a few definitions.

A minor-embedding of a graph $H = (V_H, E_H)$ in another graph $G = (V_G, E_G)$ is a map $\phi : V_H \to \mathcal{P}(V_G)$ such that

(i) the subgraphs $\phi(v)$ are connected for every $v \in V_H$,

(ii) the sets $\phi(v)$ and $\phi(u)$ are disjoint for $v \neq u$, and

(iii) if $(v, u) \in E_H$, then there has to be at least one edge in E_G joining $\phi(v)$ and $\phi(u)$.

In this setting, the vertices in V_H and V_G depict the logical and the physical spins, respectively. The subset $\phi(v)$ is said to be a chain representing v. Heuristics to find such minor-embeddings and chains can be found in [15].

The coefficients h_i, J_{ij} for the physical system are then determined by fairly distributing the logical values h_v, J_{vu} among all physical spins and available connections. For instance, if v is represented by the chain $\phi(v) = \{i_1, \ldots, i_m\}$, then $h_{i_k} = h_v/m$ for every $k = 1, \ldots, m$. Furthermore, in order to penalize chained physical spins not having the same value, a negative enough coupling $J_{ij} = -\texttt{chain\_strength} < 0$ is added whenever i and j belong to the same chain and $(i, j) \in E_G$.

B. What value of `chain_strength` should be used?

According to the explanation above, larger values of `chain_strength` should always give better results. However, D-Wave's annealers do not accept arbitrary values of h_i and J_{ij}. More precisely, h_i has to lie inside $[-2, +2]$ for the DW_2000Q_6 and $[-4, +4]$ for the Advantage_system4.1; and J_{ij} has to be in $[-1, +1]$ for both annealers.

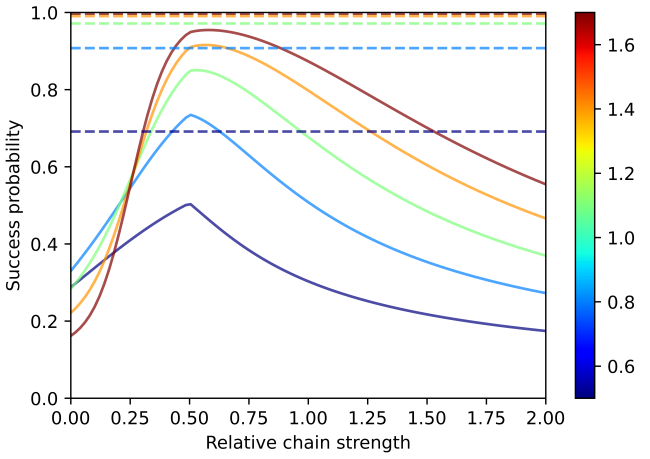

FIG. 3: Solid lines represent success probability vs RCS for different annealing times ranging from 0.5 ns to 1.7 ns distinguished by color. Dashed lines represent success probability of the logical problem as in section II B.

Thus, the input physical coefficients h_i, J_{ij} are actually divided by

$$\max\left\{\frac{|h_i|}{\texttt{h\_range}}, \frac{|J_{ij}|}{\texttt{J\_range}}\right\} \tag{8}$$

before the annealing starts. As a consequence, setting too large values of `chain_strength` yields an Ising Hamiltonian with negligible problem coefficients h_i, J_{ij}, i.e. a smaller minimum gap, more non-adiabatic jumps and, hence, less probability to obtain an actual solution at the end of the annealing.

On the other hand, too low values of `chain_strength` may not even correspond to a Hamiltonian whose ground state represents a solution to the problem. Therefore, a finite and non-zero optimal `chain_strength` is expected to exist.

Following recent work [16], we define the relative chain strength

$$\text{RCS} = \frac{\texttt{chain\_strength}}{\texttt{max\_strength}}, \tag{9}$$

where $\texttt{max\_strength} = \max\{|h_i|, |J_{ij}|\}$. The intention of this definition is to express the optimal `chain_strength` in terms of a less problem-specific parameter.

Simulations of the evolution on the graph of Fig. 1 with this new setup give the results shown in Fig. 3. Notice that no matter the annealing time, there is always a spike at RCS $= 0.5$. This is just a consequence of the fact that, with our expression of RCS (9), the scaling factor (8) changes definition precisely at RCS $= 0.5$. Notice also that any solid line lies below its corresponding dashed line, i.e. success rates with this new setup are slightly worse than when we just solved for the logical case without chains in section II B.

In Fig. 4, we can see the corresponding results on the actual DW_2000Q_6 annealer (cf. [13]) with annealing

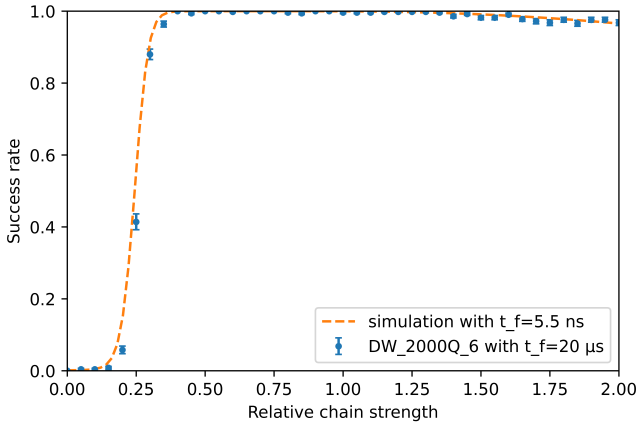

FIG. 4: Success rate vs RCS on DW_2000Q_6 with 20 μs annealing time and 500 repetitions together with simulation results with 5.5 ns annealing time.

time $t_f = 20$ μs together with the simulation outcome with annealing time $t_f = 5.5$ ns. Despite the huge difference in t_f, the dashed line fits the blue points remarkably well. This has an accidental meaning and can be justified by saying that our simulations are of an ideal closed system, while D-Wave's annealers are open and susceptible to noise, as they are actual physical realizations. In fact, recent simulations of the annealing together with a heat bath using several models give better agreement with D-Wave data [17].

One thing, though, that does not seem accidental and agrees with simulations is having a maximum success rate at RCS ~ 0.5.

IV. LARGER PROBLEMS

In this section we test D-Wave's annealers at solving the maximum clique problem for graphs with sizes ranging from 15 to 60 vertices. All graphs are randomly generated according to the Erdös-Rényi model [18]. In this model, for every pair of vertices there is an edge joining them with fixed probability p. We consider three values of p, 0.2, 0.5 and 0.8, which correspond to sparse graphs, balanced ones and dense graphs, respectively.

We first scan over RCS with fixed $t_f = 20$ μs and then over annealing times with fixed optimal RCS.

A. Scanning RCS

Results are shown in Fig. 5 and Fig. 6. Several conclusions can be drawn from these:

(1) As expected because of its higher connectivity, Advantage_system4.1 gives better results than DW_2000Q_6, although there are some exceptions. The former even finds solutions for dense graphs with 60 vertices, while the latter cannot deal with 45-vertex graphs.

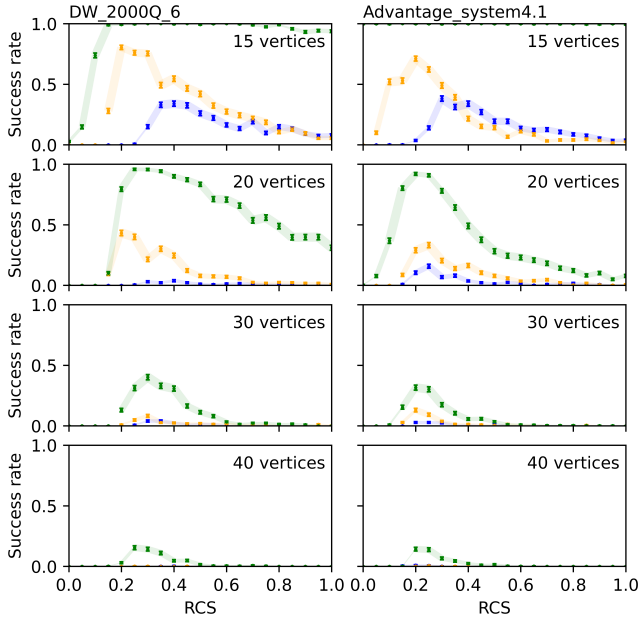

FIG. 5: Success rate vs RCS for different graph sizes and connectivity on both D-Wave annealers. Green data corresponds to dense graphs ($p = 0.8$), orange to balanced ones ($p = 0.5$) and blue data to sparse graphs ($p = 0.2$). All instances are solved with 20 μs annealing time and 500 repetitions.

(2) Dense graphs are easier to deal with than sparse graphs. This is because the denser a graph is, the more coefficients J_{ij} vanish (cf. Eq. (4)) and the less physical spins are needed. This is on the other hand rather surprising because classical algorithms work the other way around: large cliques on sparse graphs are easier to find [19]. This suggests using a hybrid algorithm that somehow deals with sparse components classically and with dense components quantically.

(3) The quality of the annealing heavily depends on the value of RCS, but optimal RCS seems to lie in the range $[0.2, 0.3]$, especially for large graphs. When solving the MCP we can thus restrict ourselves to this range (or similar ones) instead of scanning RCS over huge intervals.

B. Scanning annealing times

Results for dense graphs ($p = 0.8$) with 30 and 40 vertices are show in Fig. 7. Annealing times lie in the interval $[1.0\ \mu\text{s}, 2000.0\ \mu\text{s}]$, which is the available D-Wave range. Better performance with larger annealing times is generally observed for Advantage_system4.1, but it is not as clear with DW_2000Q_6. Another aspect to notice is that using larger annealing times does not seem to pay off. Indeed, take for example the 30-vertex instance solved with Advantage_system4.1. The success rate with

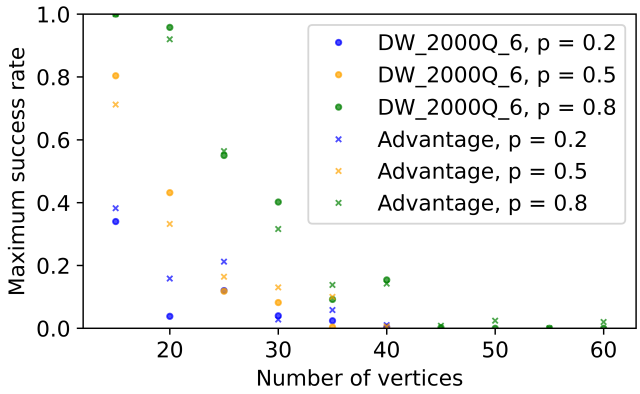

FIG. 6: Maximum success rates for graphs of different sizes and connectivity on both DW_2000Q_6 and Advantage_system4.1 with 20 μs annealing time and 500 repetitions.

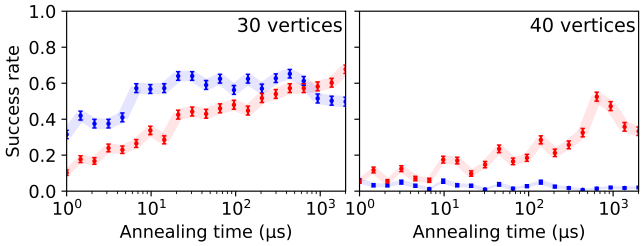

FIG. 7: Success rate vs annealing time for dense graphs of 30 and 40 vertices on both D-Wave annealers. Blue data corresponds to DW_2000Q_6 and red data to Advantage_system4.1. Statistics are gathered over 400 repetitions for each annealing time.

$t_f = 1$ μs is 0.1, while for $t_f = 2000$ μs is 0.6, The annealing is thus 2000 times longer, but the success rate is only enhanced by a factor of 6, which ends up being worse in terms of computing time.

V. SUMMARY AND CONCLUSIONS

In this work we have studied quantum annealing at solving the maximum clique problem. We first gave simulation results of DW_2000Q_6 on a simple 5-vertex graph and already got high success probabilities (> 0.95) for annealing time $t_f \gtrsim 1.0$ ns.

In reality, however, this simple problem needs a minor-embedding in the Chimera graph to be physically solved by the DW_2000Q_6 annealer. Simulations of this embedded problem gave worse results, but success probabilities were still > 0.95 for annealing time $t_f \gtrsim 1.7$ ns. Actual results on the annealer with $t_f = 20$ μs, though, seemed to fit simulation outcome with $t_f = 5.5$ ns. This behaviour is justified by introducing a finite temperature through a heat bath as environment in the simulation [17].

Finally, we have used DW_2000Q_6 and Advantage_system4.1 to solve larger random instances of MCP. Results generally show better success rate with the latter than with its predecessor. Also, although optimal RCS depends on the problem, it is seen to generally lie in $[0.2, 0.3]$ for large instances. Another remarkable observation is the fact that contrary to classical algorithms, QA works best with dense graphs.

Acknowledgments

I would like to thank Dr. Bruno Julià Díaz and Ivan Morera Navarro for their orientation and advise as well as for the useful meetings we had. I also appreciate the minutes of computation D-Wave Systems has offered me to complete this work.

[1] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition* (Cambridge University Press, 2010).
[2] T. Albash and D. A. Lidar, Rev. Mod. Phys. **90**, 015002 (2018).
[3] D. Aharonov, W. van Dam, J. Kempe, Z. Landau, S. Lloyd and O.Regev, in *45th Annual IEEE Symposium on Foundations of Computer Science* (2004) pp. 42-51.
[4] D-Wave Systems: `https://www.dwavesys.com/`
[5] S. Jansen, M. Ruskai, R. Seiler, J. Math. Phys. **48**, 102111 (2007).
[6] C. C. McGeoch, *Adiabatic Quantum Computation and Quantum Annealing: Theory and Practice* (Morgan & Claypool, 2014).
[7] R. M. Karp, in *Complexity of Computer Computations*, edited by R. E. Miller, J. W. Thatcher, and J. D. Bohlinger (Springer US, Boston, 1972), pp. 85-103.
[8] S. Butenko and W. E. Wilhelm, Eur. J. Oper. Res. **173**, 1-17 (2006).
[9] A. Douik, H. Dahrouj, T. Y. Al-Naffouri and M. Alouini, Proc. IEEE **108**, 583-608 (2020).
[10] A. Lucas, Front. Phys. **2**, 5 (2014).
[11] G. Chapuis, H. Djidjev, G. Hahn and G. Rizk, J. Signal Process. Syst. **91**, 363-377 (2019).
[12] D-Wave Physical Properties and Schedules: `https://docs.dwavesys.com/docs/latest/doc_physical_properties.html`
[13] The entire code can be found in my GitHub.
[14] D-Wave Topologies: `https://docs.dwavesys.com/docs/latest/c_gs_4.html`
[15] J. Cai, W. G. Macready and A. Roy, arXiv:1406.2741v1.
[16] D. Willsch, M. Willsch, C. D. Gonzalez Calaza, F. Jin, H. De Raedt, M. Svensson and K. Michielsen, arXiv:2105.02208v1.
[17] M. Willsch, D. Willsch, F. Jin, H. De Raedt and K. Michielsen, Phys. Rev. A **101**, 012327 (2020).
[18] B. Bollobás, *Random Graphs* (Cambridge University Press, 2001).
[19] E. Tomita, A. Tanaka and H. Takahashi, Theor. Comput. Sci. **363**, 28-42 (2006).