Quantum annealing of a hard combinatorial problem

Daniel Lecina Casas
Supervisor: Matteo Palassini
Departament de Física Fonamental, Facultat de Física, Universitat de Barcelona, Diagonal 647, 08028 Barcelona

We present the numerical results obtained using quantum annealing (QA) in a hard combinatorial problem: the coloring problem (q-COL) of an Erdős-Rényi graph. We first propose a quantum coloring Hamiltonian, natural extension of q-COL, based on the quantum Ising model in a transverse field. We then test several QA schemes and find the one that solves the highest number of graphs in the smallest number of iterations. Our results suggest that the computation time of QA scales exponentially in the size and it does not improve the results obtained by thermal annealing (TA) for q-COL.

Keywords: constraint satisfaction problem, coloring, quantum coloring problem, quantum annealing

I. INTRODUCTION

There is a great interest in finding solutions of highly constrained combinatorial problems since they appear in many fields of study. In a constraint satisfaction problem (CSP) one must decide if a set of N discrete variables can fulfill M constraints. An example is the q-coloring problem (q-COL): given a graph with N nodes and M edges, the aim is to find a configuration in which connected nodes have different colors chosen among q possibilities.

A decision problem involving discrete variables is said to belong to the P class if it can be solved, in the worst case, in a time scaling polynomially in its size and to the NP class if it is possible to verify in a polynomial time whether an assignment of the variables is a solution or not. Thus, NP contains P. The hardest instances of NP form the NP-complete class, to which any other NP problem can be reduced in a polynomial time [1, 2]. It is widely believed that NP-complete problems cannot be solved, in the worst case, in polynomial time on classical computers. Proving this conjecture is the most important open problem in theoretical computer science.

Statistical mechanics and quantum mechanics provide new approaches to study combinatorial problems. In this report our main interest is to find if we can take advantage of them to solve q-COL [3]. q-COL belongs to the NP-complete class, so an efficient algorithm to solve it would also solve efficiently any other NP problem. Furthermore, it is quite interesting for physicists since it can be rephrased in terms of the antiferromagnetic Potts model by defining a Hamiltonian equal to the number of unsatisfied constraints:

$$\mathcal{H} = \sum_i \sum_{j \in \partial i} \delta_{\sigma_i \sigma_j} \tag{1}$$

where $\sigma \in \{1, \dots, q\}$ and $\sum_{j \in \partial i}$ sums over all the j nodes connected to i.

A naive brute-force search of a ground state configuration implies q^N assignments of colors, which becomes impractical for large graphs. Using quantum mechanics we are going to build a scheme of annealing to reach faster the ground state.

This report is organized as follows. Section II is devoted to explain the two stochastic methods that we use to find the ground state of q-COL. In Section III we propose a quantum coloring model. The code and the followed procedure are analyzed in Section IV. Finally, we expose the results in Section V and present the main conclusions in Section VI.

II. QUANTUM AND THERMAL ANNEALING

Consider a system evolving with a local stochastic dynamics such as the Metropolis algorithm converging to the Boltzmann distribution at temperature T. If the configuration space is rugged and has many local energy minima (metastable states) the system can get trapped if T is small in comparison to the energy barriers. Finding a ground state configuration becomes then a challenge and an annealing scheme improves notably the search.

The idea of thermal annealing (TA), also known as simulated annealing, is to start sampling the Boltzmann distribution from a high temperature (i.e. large fluctuations) and lower it slowly so that the system leaves local minima in order to reach deeper minima. If we follow the correct schedule it is guaranteed that the system eventually reaches the global minimum, but, if T is decreased too quickly, the system might get stuck in a shallower minimum. The problem is that the prescribed time is exponential in the size of the system (see [4, 5]).

Quantum annealing (QA) [6] is another proposal based on quantum tunneling as well as quantum adiabatic evolution. Quantum tunneling induces transitions between states, thereby crossing easily high barriers when they are narrow. This may help the system leave the metastable states and explore faster the rest of configurations.

The quantum adiabatic theorem (QAT) states that a physical system remains in an instantaneous eigenstate of its Hamiltonian if a perturbation acts on it slowly enough and if there is a gap between the eigenvalue and the rest of the Hamiltonian spectrum. The QAT can be used to find the ground state of a

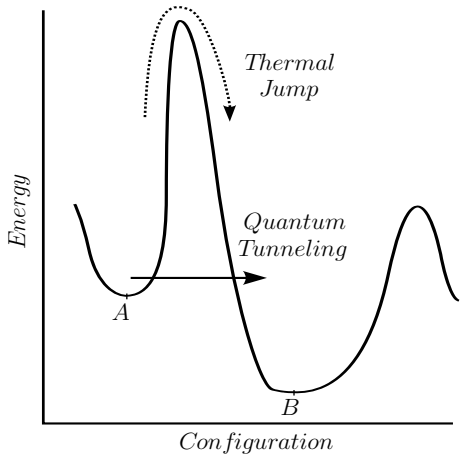

Figure 1. To reach the ground state we need to leave configuration A and reach the deeper minimum B. Classically, the only way to reach it is with thermal fluctuations, but quantum mechanics gives us the chance to tunnel through the energy barrier.

classical Hamiltonian, $\mathcal{H}_1$, by: 1) converting $\mathcal{H}_1$ into an analogous quantum mechanical system $\widehat{\mathcal{H}}_1$; 2) designing a suitable $\widehat{\mathcal{H}}_2$ that does not commute with $\widehat{\mathcal{H}}_1$ and such that we know how to prepare the system in an eigenstate of $\widehat{\mathcal{H}}_2$; 3) adding it to $\widehat{\mathcal{H}}_1$, $\widehat{\mathcal{H}} = \widehat{\mathcal{H}}_1 + \Gamma\widehat{\mathcal{H}}_2$; 4) starting at $\Gamma \gg 1$ and lowering Γ adiabatically to 0.

For example, consider the quantum Ising model (defined, for now, on an arbitrary graph): $\widehat{\mathcal{H}}_1 = -J\sum_i \sum_{j\in\partial i} \hat{\sigma}_i^z \hat{\sigma}_j^z$, where $\hat{\sigma}_j^z$ is the z Pauli spin operator on the j-th spin, and J is the coupling between spins. It is the natural extension of the classical Ising model and it is diagonal in the $\{|\underline{\sigma}\rangle\}$ basis. Taking advantage of the QAT we can find the ground state of $\widehat{\mathcal{H}}_1$ by adding a transverse field:

$$\widehat{\mathcal{H}} = \widehat{\mathcal{H}}_1 + \Gamma\widehat{\mathcal{H}}_2 = -J\sum_i \sum_{j\in\partial i} \hat{\sigma}_i^z \hat{\sigma}_j^z - \Gamma\sum_i \hat{\sigma}_i^x \quad (2)$$

where $\Gamma \geq 0$. The terms $\hat{\sigma}^x$ are off-diagonal in the $|\sigma\rangle$ basis, and consequently introduce quantum-mechanical tunneling. When $\Gamma \gg 1$ the ground state of $\widehat{\mathcal{H}}$ corresponds to the superposition of all classical configurations. The QAT guarantees that lowering Γ "slowly"[1] we will finally reach the ground state of $\widehat{\mathcal{H}}_1$ when $\Gamma = 0$.

This strategy fails if there is a crossing of energy levels of the ground state and the first excited state. In particular, quantum phase transitions (QPT) may lead to a crossing of energy levels. While classical phase transitions usually happen at a finite temperature and order is destroyed by thermal fluctuations, QPT occur at T=0 and fluctuations are driven by Heisenberg's uncertainty principle. For example, the quantum Ising model in a transverse field has a certain $\Gamma = \Gamma_c$ at which spontaneous magnetization appears at $T = 0$ [7] and another phase transition driven by temperature for the model on a lattice of dimension larger than two [8].

As we already saw, Eq. (1) describes q-COL. In sections III A and III B we will give a natural quantum extension and add a proper non-commuting term.

TA and QA have different inherent natures and limitations, and sometimes it may not be clear which performs better. Comparisons between TA and QA for the Ising model with transverse field, precursor of our model, have been reported diagonalizing the Hamiltonian at $T = 0$ [9]. In contrast, we will compare both annealings by sampling the quantum partition function at $T > 0$.

III. QUANTUM MONTE CARLO FOR COLORING

A. Quantum Ising Model in a transverse field

Let us consider the Ising model in a transverse field defined above. The whole system is described by a Hilbert space spanned by 2^N vectors: $|\underline{\sigma}\rangle = |\sigma_1\rangle \otimes \cdots \otimes |\sigma_N\rangle$, where $\underline{\sigma}$ denotes a single configuration.

The partition function is defined by $Z = \mathrm{Tr}\left(e^{-\beta\widehat{\mathcal{H}}}\right)$ and using the Suzuki-Trotter formula [10, 11] can be written as (see Appendix A for derivation):

$$Z = \lim_{m\to\infty} K_1 \sum_{\{\underline{\sigma}^{tot}\}} \prod_{\tau=1}^{m} e^{K_2 \sum_{i,j\in\partial i} \sigma_i^\tau \sigma_j^\tau} e^{\tilde{\Gamma}\sum_i \sigma_i^\tau \sigma_i^{\tau+1}} \quad (3)$$

where we have introduced m copies of the system (or layers) and $K_1 = \left(\frac{\sinh\frac{2\Gamma\beta}{2}}{2}\right)^{\frac{Nm}{2}}$, $K_2 = \frac{J\beta}{m}$ and $\tilde{\Gamma} = -\frac{1}{2}\ln\tanh\frac{\Gamma\beta}{m}$.

The partition function (3) is akin to the partition function of a classical Ising model with Nm spins, with Hamiltonian:

$$-\beta\mathcal{H} = K_2 \sum_{i,j\in\partial i,\tau} \sigma_i^\tau \sigma_j^\tau + \tilde{\Gamma}\sum_{i,\tau} \sigma_i^\tau \sigma_i^{\tau+1} \quad (4)$$

Hence, for instance, the d-dimensional quantum Ising model (a special case of Eq. 2 in which the graph is a d-dimensional lattice) is mapped into an anisotropic (d+1)-dimensional classical Ising model. The new dimension is called imaginary time because when we assign $t = -i\hbar\beta$, the density matrix operator $e^{-\beta\widehat{\mathcal{H}}}$ can be interpreted as the Hamiltonian $e^{-i\frac{\widehat{\mathcal{H}}t}{\hbar}}$.

When $\Gamma \to 0^+ (K_3 \to +\infty)$ all the layers are perfectly correlated and we have exact replicas of the initial system along the imaginary time direction (this is must be so, since the quantum Hamiltonian reduces to the classical one). In the other limit $\Gamma \to +\infty (K_3 \to 0^+)$ the layers are totally uncorrelated.

[1] See [2] for a criterion to choose the rate of variation of Γ.

B. Quantum q-coloring model

We are now ready to build a q-coloring quantum Hamiltonian. Let us consider N nodes, each described by the eigenvectors $|\sigma_i\rangle$ with eigenvalues $1, 2, \ldots, q$. The q^N-dimensional Hilbert space is spanned by the vectors $|\underline{\sigma}\rangle = |\sigma\rangle \otimes \cdots \otimes |\sigma\rangle$.

In analogy with the quantum Ising model, we propose a Hamiltonian $\widehat{\mathcal{H}} = \widehat{\mathcal{H}}_1 + \Gamma\widehat{\mathcal{H}}_2$ whose first part is analogous to the energy of the classical model (thus, it must be diagonal) $\widehat{\mathcal{H}}_1 = \sum_{|\sigma_i\rangle,|\sigma_j\rangle} |\sigma_i\rangle|\sigma_j\rangle\langle\sigma_i|\langle\sigma_j|\delta_{\sigma_i,\sigma_j}$.

To build the operator $\widehat{\mathcal{H}}_2$ we observe that it should not commute with $\widehat{\mathcal{H}}_1$. Considering single-node flips (i.e. changes of color) and assuming that there are no privileged colors, we propose: $\widehat{\mathcal{H}}_2 = -\sum_i \sum_{\sigma_i \neq \sigma_j} |\sigma_i\rangle\langle\sigma_j|$. In particular, for $q = 2$: $\widehat{\mathcal{H}}_2 = -\sum_i \sigma_i^x$, and for $q = 3$: $\widehat{\mathcal{H}}_2 = -\sum_i (L_{+,i} + L_{-,i} + |1_i\rangle\langle 3_i| + |3_i\rangle\langle 1_i|)$, where $L_{+,i} = |1_i\rangle\langle 2_i| + |2_i\rangle\langle 3_i|$, and $L_{-,i} = |2_i\rangle\langle 1_i| + |3_i\rangle\langle 2_i|$. It can be easily shown that $[\widehat{\mathcal{H}}_1, \widehat{\mathcal{H}}_2] \neq 0$.

With the Suzuki-Trotter formalism, proceeding in the same way as in the previous section, we find the analogous classical Hamiltonian:

$$-\beta\mathcal{H} = -K_2 \sum_{\tau,i,j\in\partial i} \delta_{\sigma_i^\tau \sigma_j^\tau} + \tilde{\Gamma} \sum_{\tau,i} \delta_{\sigma_i^\tau \sigma_i^{\tau+1}} \qquad (5)$$

where $K_2 = \frac{\beta}{m}$ and $\tilde{\Gamma} = -\ln\left(1 - \frac{q}{(q-1)+e^{\frac{q\beta\Gamma}{m}}}\right) > 0$. Exactly the same analysis in terms of Γ, discussed in the previous section, applies here.

C. Erdős-Rényi ensemble

We consider the Erdős-Rényi ensemble of random graphs: given a set of N nodes, we assign an edge to every pair of nodes (i, j) with probability p, independently of all the others[2]. Thus, the probability to have M edges is a binomial:

$$P(M) = \binom{\frac{N(N-1)}{2}}{M} p^M (1-p)^{\frac{N(N-1)}{2}-M} \qquad (6)$$

with a mean number of edges $\langle M\rangle = \frac{N(N-1)}{2}p$ and variance $\mathrm{Var}(M) = \frac{N(N-1)}{2}p(1-p)$.

This ensemble of graphs has been well characterized (see [12]). They have loops of length $\mathcal{O}(\log(N))$, which introduce frustration. When the density of constraint $\alpha = \frac{M}{N}$ is small all the edges are likely to be satisfied. For large N, there is a sharp transition at α_c that separates the phase $\alpha < \alpha_c$ where almost all the instances are unsatisfiable (with probability one as $N \to \infty$) from the phase where almost all are satisfiable [3]. The region near the transition includes the hardest problems to solve [13].

IV. PROCEDURE

We have written a C++ program that, given an Erdős-Rényi graph with N nodes and constraint density α interacting with the Hamiltonian (1), uses the Metropolis algorithm following a TA or a QA scheme to find a ground state configuration for $q = 3$. We deployed about 6 months of computational time in Quad-Core Intel Xeon @2.26 GHz processors.

For each value of N and α, we simulate one hundred solvable instances (i.e. instances for which the ground state energy is zero) so that we know when we have reached the ground state. These were obtained in [14] using a parallel implementation of the complete Davis-Putnam-Logemann-Loveland (DPLL) algorithm. The number of nodes is $N = 64, 128, 256$ and the range of α includes the hardest instances: $2.10 \leq \alpha < \alpha_c = 2.34$.

For both QA and TA, the program initializes the system in a random configuration and iterates choosing a single random node at a time until it reaches the ground state[3] or a preset maximum number of Monte Carlo steps per node (MCS), τ_{MCS}.

In TA, the program samples the Boltzmann distribution for the Hamiltonian (1) at temperature varying during the simulation time t as $T(t) = T_{\mathrm{in}}\left(1 - \frac{t}{\tau_{\mathrm{MCS}}}\right)$, where T_{in} is the initial temperature and $t \in [0, \tau_{\mathrm{MCS}})$.

In QA, it samples the Boltzmann distribution for the Hamiltonian (5) at a fixed number of layers m and temperature T and at $\tilde{\Gamma}(t) = \tilde{\Gamma}_{\mathrm{in}}\left(1 - \frac{t}{\tau_{\mathrm{MCS}}}\right)$, where $\tilde{\Gamma}_{\mathrm{in}}$ is the initial $\tilde{\Gamma}$ and $t \in (0, \tau_{\mathrm{MCS}}]$. We choose $T = m^{-1}$ so that $K_2 = 1$ in (5).

In order to find the optimal $\tilde{\Gamma}_{\mathrm{in}}$, we fix $m \sim N$ and run simulations for different $\tilde{\Gamma}_{\mathrm{in}}$. Afterwards, with the optimal $\tilde{\Gamma}_{\mathrm{in}}$ fixed, we run several simulations for different m, yielding an optimal m value. We continue iterating until both values become stable.

Once we have chosen the optimal parameters we will study how the fraction of solved graphs increases with τ_{MCS} and compare it with TA.

V. RESULTS

A. Behavior of QA as a function of $\tilde{\Gamma}_{\mathrm{in}}$

Figure 2 shows the fraction of solved graphs as a function of $\tilde{\Gamma}_{\mathrm{in}}$ for different τ_{MCS}. As we can see,

[2] This may give a disconnected graph. We avoid this choosing only the largest connected component.

[3] In QA, when all the edges within a single layer are satisfied, which was indeed the original goal, we can build a configuration with zero energy by choosing an even permutation of the colors (e.g. $(1, 2, 3) \to (2, 3, 1)$) in all the subsequent neighboring layers.

there is a peak at $\tilde{\Gamma}_{\rm in} \simeq 2$ which separates two different regimes. We need $\tilde{\Gamma}_{\rm in} \gtrsim K_2 \left(= \frac{\beta}{m} = 1\right)$ in order to tunnel through the barrier necessary to flip one variable (see Eq. (5)). For $\tilde{\Gamma}_{\rm in} \lesssim 2$, the system cannot tunnel and can only evolve via thermal fluctuations. For $\tilde{\Gamma}_{\rm in} \gg 2$, the system will spend too much time at $\tilde{\Gamma}(t) \gg 1$, and not enough time in the region of low $\tilde{\Gamma}(t)$, therefore it will not sample sufficiently $\widehat{\mathcal{H}}_1$.

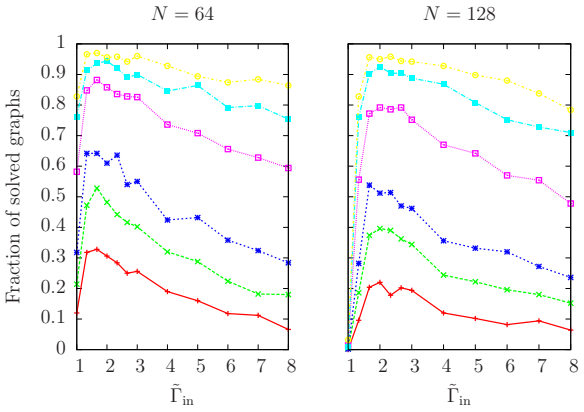

Figure 2. Fraction of solved graphs as a function of $\tilde{\Gamma}_{\rm in}$ for different $\tau_{\rm MCS}$. To compute each point we average the fraction of solved graphs over a total amount of 500 instances: 100 for each value $\alpha = 2.10$, 2.20, 2.25, 2.30 and 2.32 at $m = 4$. Each line represents a $\tau_{\rm MCS}$ value. From bottom to top, $N = 64$: $\tau_{\rm MCS} = 3 \cdot 10^3$, $6 \cdot 10^3$, 10^4, $3 \cdot 10^4$, $6 \cdot 10^4$, 10^5 MCS; $N = 128$: $\tau_{\rm MCS} = 3 \cdot 10^4$, $6 \cdot 10^4$, 10^5, $3 \cdot 10^5$, $6 \cdot 10^5$, 10^6 MCS. The maximum efficiency is found at $\tilde{\Gamma}_{\rm in} \simeq 2$.

As we increase $\tau_{\rm MCS}$, the peak becomes a plateau. This is because there are hard instances that require a larger $\tau_{\rm MCS}$ to be solved. To prove their existence, in Figure 3 we plot the number of MCS that it takes to solve all the instances that we used in Figure 2 for two values of $\tilde{\Gamma}_{\rm in}$ ($N = 64$, $\tilde{\Gamma}_{\rm in} = 1.3, 3.0$). We observe that, although there is no general correlation, there is a large number of common graphs not solved with either value of $\tilde{\Gamma}_{\rm in}$ (or solved by just one with $\tau \lesssim \tau_{\rm MCS}$). Moreover, the typical time to solve a graph increases with $\tilde{\Gamma}_{\rm in}$, since it takes more time to reach smaller $\tilde{\Gamma}$.

It is also noteworthy that, for each $\tilde{\Gamma}_{\rm in}$, there is a gap (shaded region in the Figure 3 for $\tilde{\Gamma}_{\rm in} = 1.3$) with no solved graphs within $\tau_{\rm MCS}$ and the maximum number of MCS that took to solve a graph. This is due to the fact that, for t close to $\tau_{\rm MCS}$, $\tilde{\Gamma}(t)$ is too small for the system to tunnel through energy barriers.

B. Behavior of QA as a function of m

The performance of QA as a function of m is harder to analyze due to the superposition of various effects. In Figure 4 we plot the fraction of solved graphs versus m at a constant number of total single-node updates, $\tau_{\rm it}$, by choosing $\tau_{\rm MCS} = \frac{\tau_{\rm it}}{Nm}$, at $\tilde{\Gamma}_{\rm in} = 2$. We observe two peaks, one at $m = 4$ and another at

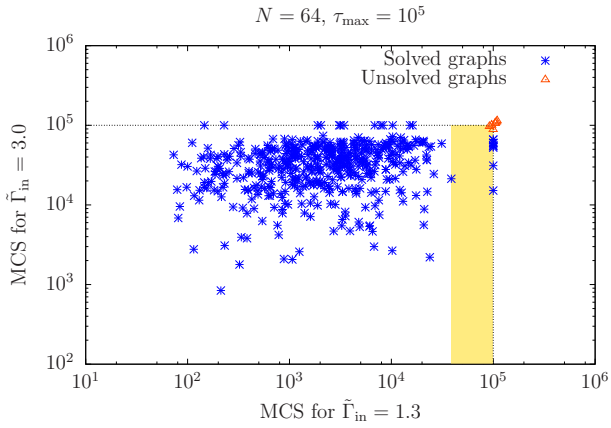

Figure 3. Scatter plot of the number of MCS that took to solve a given graph by QA with $\tilde{\Gamma}_{\rm in} = 1.3$ and 3.0 at $m = 4$. Each point represents one of the 500 graphs included in Figure 2. When a graph is not solved we assign it the number of MCS $\tau_{\rm MCS}$ (dashed line). We added some noise to the number of MCS of the graphs not solved by any of the two values (see the upper right corner).

$m \sim N$ that shifts to higher m as $\tau_{\rm it}$ increases.

The first effect is the rate of variation of $\tilde{\Gamma}(t)$. At small m, $\tilde{\Gamma}(t)$ is varied more slowly, since $\tau_{\rm MCS}$ is larger. That explains, in part, the first peak. On the other hand, a higher m might cause $\tilde{\Gamma}(t)$ to vary too fast, if $\tau_{\rm it}$ is too small. That is why, as we increase $\tau_{\rm it}$, the second peak shifts to higher m.

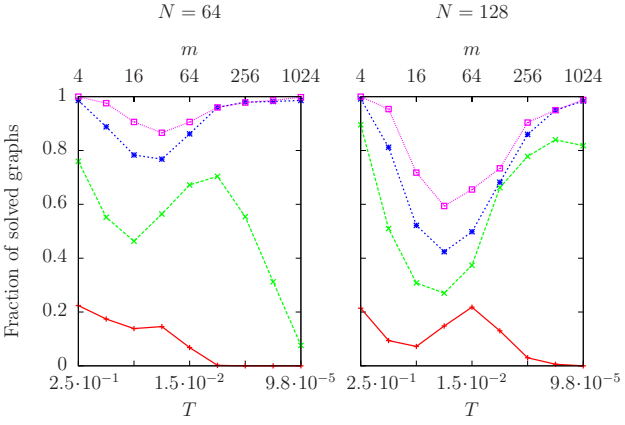

Figure 4. Each point represents the fraction of graphs solved of the same ensemble of 500 instances that we used for the Figure 2, 100 for $\alpha = 2.10$, 2.20, 2.25, 2.30 and 2.32. Each line has a constant $\tau_{\rm it}$ for different m. From bottom to top: $N = 64$, $\tau_{\rm it} \simeq 5.2 \cdot 10^5$, $4.2 \cdot 10^6$, $6.7 \cdot 10^7$, $5.4 \cdot 10^8$; $N = 128$, $\tau_{\rm it} \simeq 1.7 \cdot 10^7$, $2.7 \cdot 10^8$, $2.1 \cdot 10^9$, $1.7 \cdot 10^{10}$.

The second effect is that a larger m permits a better exploration of the configuration space, since we explore simultaneously m different configurations of the original system. In the second peak, this multiple exploration makes up for the fast lowering of $\tilde{\Gamma}(t)$.

The last effect is that at small m , T $(= m^{-1})$ is

large enough to create thermal fluctuations, and these help the system to explore better the configuration space, i.e. in this region we have the superposition of quantum and thermal exploration. However, we expect this effect to vanish for large N, since the probability to visit the ground state by fluctuations vanishes exponentially with N for any finite T. Then, the peak at $m = 4$ is the crossover from a thermal to a quantum regime. For larger N, the only choice in order to increase the probability to visit the ground state is $T \sim N^{-1}$, so that $\delta N \sim \mathcal{O}(1)$.

It is important to note that the fraction of solved graphs increases faster with $\tau_{\rm it}$ in the first peak than in the second one (e.g. for $N = 64$, $\tau_{it} = 6.7 \cdot 10^7$ — $4.3 \cdot 10^{10}$, the performances are different in the first peak, while they are similar in the second). We will analyze this point further in the next section.

C. Behavior of QA as a function of $\tau_{\rm it}$

To study the growth of the fraction of solved graphs with $\tau_{\rm it}$, we apply QA to solve an ensemble of 100 graphs for all the combinations of $N = 64, 128$ and $\alpha = 2.10, 2.20, 2.25, 2.30, 2.32$ for a wide range of $\tau_{\rm it}$ at $\tilde{\Gamma}_{\rm in} = 2$, and at the two m that showed the best performance, $\frac{m}{N} = 2^4$ and $m = 4$. In order to study the behavior with N, we also study $N = 256$ at $m = 4$.

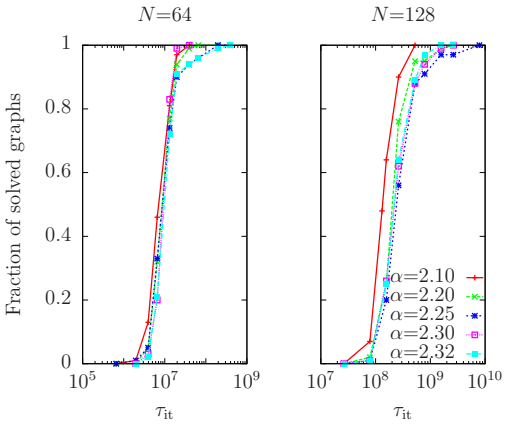

Figure 5. The symbols in the right panel are the same as in the left one. Each point represents the fraction of graphs solved of the same 100 instances for different values of $\tau_{\rm it}$ and fixing $\tilde{\Gamma}_{\rm in} = 2$, $\frac{m}{N} = 2^4$.

Figure 5 shows long tails[4] indicating that there is a small fraction of instances that take a much longer time than the median $\tau_{\rm in}$ to solve.

The plot at $m = 4$ (Figure 6) shows tails of equal length and the fraction of solved graphs increases more progressively than at $\frac{m}{N} = 2^4$. It suggests that thermal fluctuations help the system to reach a ground state configuration for these moderate values of N. However, the required $\tau_{\rm it}$ to solve all the instances are similar.

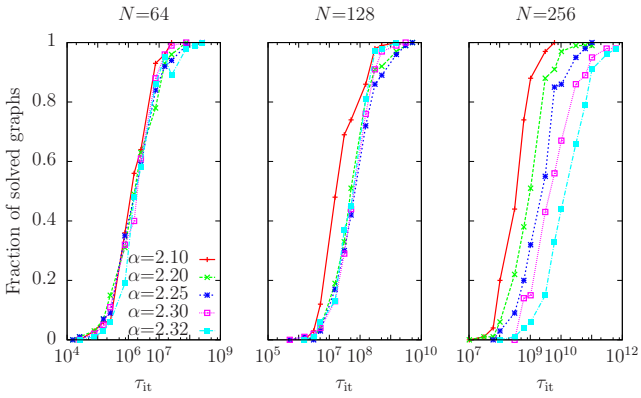

Figure 6. The symbols in the left panel are the same as in the central and the right ones. Each point represents the fraction of graphs solved of the same 100 instances for different $\tau_{\rm it}$ and fixing: $\tilde{\Gamma}_{\rm in} = 2$, $m = 4$.

As we increase N the difference for the various α also increases. To understand this behavior we compute the number of iterations that it takes to solve an instance with a certain probability: we fix a threshold of solved graphs and obtain by interpolation the smallest number of iterations to achieve it. The results are shown in Figure 7. For small probabilities, the behavior for small α seems to be a power law, while for larger α it is not inconsistent with an exponential. For example, for $P = 0.5$, $\alpha \leq 2.30$ seems to have a power law behavior, while $\alpha = 2.32$ seems to behave exponentially. On the other hand, when we increase the probability, an exponential behavior arises for smaller α. A possible explanation of this is that: 1) For all α, there are easy and hard instances that can be solved in a polynomial and an exponential number of iterations respectively. 2) The ratio of harder instances increases with α. The exponential behavior and the different ratios of harder and easy instances lead to the increasing difference for the various α. These hard instances make the algorithm scale exponentially in the size.

D. Comparison with TA

To compare the results with TA, we try to solve the previous ensemble of graphs for different $\tau_{\rm it}$ and following the same annealing schedule. The results are shown in Figure 8.

The results are consistent with those obtained in the previous subsection for $m = 4$: the probability shows no long tails and the difference for the various α also increases with N. Therefore, it agrees with the fact that at $m = 4$ solutions are found mainly due to thermal fluctuations. Nevertheless, QA at $m = 4$ is

[4] By tail we mean the region where the fraction of solved graphs is either $\gtrsim 0$ or $\lesssim 1$

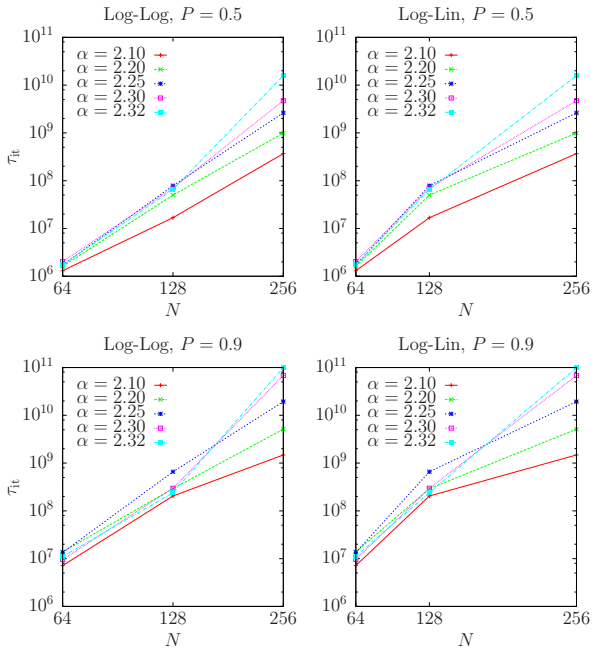

Figure 7. Plot of $\tau_{\rm it}$ required to solve a graph with probabilities 0.5 and 0.9 respectively at $m = 4$.

less efficient due to the cost associated to the extra $m - 1$ layers.

On the other hand, QA at $\frac{m}{N} = 2^4$ has a longer tail at large t, and the fraction of solved graphs increases more rapidly. The long tails make QA not to improve over TA.

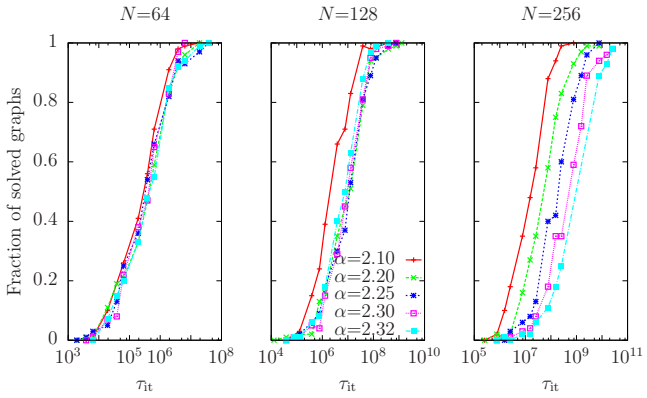

Figure 8. The symbols in the central panel are the same as in the left and the right ones. Each point represents the fraction of graphs solved by TA of the same 100 instances that we used for QA, for different $\tau_{\rm it}$ and fixing $T_{\rm in} = 2$.

VI. CONCLUSIONS

We have implemented a new approach to tackle a CSP: we have rewritten the classical coloring problem in terms of a quantum Hamiltonian in order to use quantum statistical mechanics to find a ground state configuration. Using the Suzuki-Trotter formula we have found a Hamiltonian which can be used to simulate the quantum coloring problem. We have tested the algorithm for a wide range of values of the different parameters and found the optimal ones for the coloring problem. This choice is crucial to the efficiency of the algorithm. The initial value $\tilde{\Gamma}_{\rm in}$ shows a peak which is the middle ground of a value high enough to tunnel through energy barriers and low enough not to spend too much time at $\tilde{\Gamma}(t) \gg 1$. The number of layers m shows two peaks, a first one at $m = 4$, due to thermal fluctuations, and a second one at $m \sim N$, which is the middle ground of a value high enough to explore fast the configuration space and low enough so that $\tilde{\Gamma}(t)$ does not vary to fast. We have pointed the existence of hard instances, the quantity of which increases with α, which suggest that the algorithm scales exponentially in the size. However, it does not improve over TA.

Further studies should focus on studying the behavior of QA at larger N. The most interesting point would be the ratification of the expected exponential behavior of $\tau_{\rm it}$ with N. Moreover, it would be interesting to confirm numerically that the probability to find a ground state configuration at $m = 4$ tends to zero for large N.

ACKNOWLEDGMENTS

I would like to thank Matteo Palassini for the guidelines throughout the project and for his help in the analysis of the results and Salvatore Mandrà for the fruitful discussions, for his advice in the process of writing the code and for providing the ensemble of graphs.

Appendix A: Suzuki-Trotter Formula

The Suzuki-Trotter formula states that $\widehat{\rho}_m = \left\{ e^{\widehat{A}_1/m} \cdots e^{\widehat{A}_p/m} \right\}^m \to \exp\left(\sum_i^p \widehat{A}_i \right)$ as $m \to \infty$ for a set of operators $\{\widehat{A}_j\}$. By means of this formulation we can compute:

$$\begin{aligned} Z &= \textstyle\sum_{\{\underline{\sigma}\}} \langle \underline{\sigma} | e^{-\frac{\beta}{m}(\widehat{\mathcal{H}}_1 + \widehat{\mathcal{H}}_2)m} | \underline{\sigma} \rangle = \\ &= \lim_{m\to\infty} \textstyle\sum_{\{\underline{\sigma}\}} \langle \underline{\sigma} | \left(e^{-\frac{\beta}{m}\widehat{\mathcal{H}}_1} e^{-\frac{\beta}{m}\widehat{\mathcal{H}}_2} \right)^m | \underline{\sigma} \rangle \end{aligned} \tag{A1}$$

Introducing the identity $\mathbb{I} = \sum_{\{\underline{\sigma}^i\}} |\underline{\sigma}^i\rangle\langle\underline{\sigma}^i|$ between consecutive terms[5] leads to:

$$Z = \lim_{m\to\infty} \sum_{\{\underline{\sigma}^{tot}\}} \prod_{\tau=1}^{m} \langle\underline{\sigma}^\tau|e^{-\frac{\beta}{m}\widehat{\mathcal{H}}_1} e^{-\frac{\beta}{m}\widehat{\mathcal{H}}_2}|\underline{\sigma}^{\tau+1}\rangle = \\ = \lim_{m\to\infty} \sum_{\{\underline{\sigma}^{tot}\}} \prod_{\tau=1}^{m} e^{\frac{J\beta}{m}\sum_{i,j\in\partial i}\sigma_i^z\sigma_j^z} \langle\underline{\sigma}^\tau|e^{-\frac{\beta}{m}\widehat{\mathcal{H}}_2}|\underline{\sigma}^{\tau+1}\rangle \tag{A2}$$

where $\sum_{\{\underline{\sigma}^{tot}\}} \equiv \sum_{\{\underline{\sigma}^1\}} \cdots \sum_{\{\underline{\sigma}^m\}}$ and $|\underline{\sigma}^{m+1}\rangle \equiv |\underline{\sigma}^1\rangle$. The matrix elements can be easily written:

$$\langle\sigma_i^\tau|e^{-\frac{\beta}{m}\hat{\sigma}_i^x}|\sigma_i^{\tau+1}\rangle = \delta_{\sigma_i^\tau,\sigma_i^{\tau+1}} \cosh\frac{\Gamma\beta}{m} + \delta_{\sigma_i^\tau,-\sigma_i^{\tau+1}} \sinh\frac{\Gamma\beta}{m} \tag{A3}$$

Taking into account that σ_i^τ can only take the values ± 1, we can rewrite Eq. (A2) in an Ising-like fashion, as in Eq. 3.

[1] M. Mézard, Montanari, *Information, Physics and Computation*, Oxford Graduate Texts (2009)
[2] A. Das, B. Chakrabarti, Rev. of Mod. Phys., Vol. 80 (2008)
[3] M. Mézard, M. Palassini, O. Rivoire, Phys. Rev. Lett. 95, 200202 (2005)
[4] B. Berg, *Markov chain Monte Carlo simulations and their statistical analysis*, World Scientific (2004)
[5] D. Landau, K. Binder, *A guide to Monte Carlo simulations in Statistical Physics*, Cambridge University Press (2005)
[6] E. Farhi, J. Goldstone, S. Gutmann, J. Lapan, A. Lundgren, D. Preda, Science, Vol. 292 no. 5516 pp. 472-475 (2001)
[7] S. Sachdev, *Quantum Phase Transitions*, Cambridge University Press (1998)
[8] F. Krzakala, A. Rosso, G. Semerjian, F. Zamponi, Phys. Rev. B 78, (2008)
[9] T. Kadowaki, H. Nishimori, Phys. Rev. E, Vol. 50, No. 5 (1998)
[10] M. Suzuki, S. Miyashita, A. Kuroda, Prog. Thero. Phys., Vol. 58, No.5 (1977)
[11] M. Suzuki, Prog. Thero. Phys., Vol. 56, No.5 (1976)
[12] A. Braunstein, R. Mulet, A. Pagnani, M. Weigt, and R. Zecchina, Phys. Rev. E 68, 036702 (2003)
[13] P. Cheesmman, B. Kanefsky, and W. M. Taylor, in Proceedings of the 12th International Joint Conference on Artificial Intelligence (IJCAI-91), Morgan Kaufmann, San Mateo, CA, pp. 331-337 (1991)
[14] S. Mandrá, M. Palassini, in preparation
[15] H. De Raedt, A. Lagendijk, Phys. Reports (Review Section of Physics Letters) 127, No.4, (1985)
[16] S.L. Sondhi, S.M. Girvin, J.P. Carini, D. Shahar, Rev. Mod. Phys., Vol. 69, No.1 (1997)
[17] H. de Raedt, B. de Raedt, Phys. Rev. A, Vol. 23, No.6, (1983)
[18] M. Kardar, *Statistical Physics of Fields*, Cambdridge University Press (2007)
[19] M. Vojta, Rep. Prog. Phys 66, pp. 2096-2110, (2003)
[20] J. Ardelius, L. Zdeborová, Phys. Rev. E 78, 040101R (2008)
[21] S. Cook, Proceedings of the Third Annual ACM Symposium on Theory of Computing. pp. 151-158 (1971)
[22] D. Griffiths, *Introduction to quantum mechanics*, Pearson Prentice Hall (2005)

[5] This step is done the other way around in the 1-D Ising model with periodic boundary conditions when it is solved with the transfer matrix method [18]. Hence, we can also understand the Suzuki-Transfer formalism as a transfer matrix method.