Quantum Annealing on NP Problems

Author: Roger Uceda Ruiz.
Facultat de Física, Universitat de Barcelona, Diagonal 645, 08028 Barcelona, Spain.

Advisors: Bruno Julià Díaz and Abel Rojo Francàs

Abstract: In this work we make use of quantum annealing simulations to solve a well-known NP-Problem, the N-Queens problem, by describing it with a Hamiltonian and finding its ground-state. We use two different methods and a biased Hamiltonian to compare the results obtained with a biased solution of the 3×3 grid to later solve for the non-biased Hamiltonian and see what we obtain. Also, we analyze the energy spectrum of both 3×3 and 4×4 grids.

I. INTRODUCTION

Quantum Annealing (QA) has proven to be an effective method for solving NP-Problems that can be described as a spin system with global interactions [1] as the system becomes larger and larger and classical computation becomes impossible. In this work we are using Quantum Annealing (QA) to find the solutions of a well-known NP-Problem, the N-Queens problem.

Before Quantum Annealing (QA), Simulated Annealing (SA) was used to try finding the solution to those NP-Problems that could be reduced to finding a ground-state of a spin system. The transitions of SA were due thermal fluctuations rather than quantum fluctuations as it is done in QA. QA was proposed by Kadowaki and Nishimori in 1984 [2]. To prove QA was better than SA, they tested with the quantum version of the Ising Model and proved this technique got better results than SA in the same conditions.

Throughout this thesis, we are comparing the results of QA with the actual solutions of a well-known NP-Problem. To do so, we use two approaches: the first approach is used by Kadowaki and Nishimori [2] and O. Promio [3], the second approach is used by D-wave [4], who is one of the most important companies that are working with QA.

The main goal of this work is to solve the N-Queens problem using QA. The N-Queens problem is a NP-Problem that consists on having a N×N grid where you must put N queens in a way they do not kill each other. The solutions of this problem are well known: for a 2×2 grid you can only put 1 Queen, for a 3×3 grid you can only put 2 Queens and for N $\geq$ 4, you can put N Queens.

This work is organized as follows. First in Section II we present the model, then in Section III we solve the 3×3 grid using the first method, after that, in Section IV we develop the second method and also present the energy spectrum of the 3×3 and 4×4 grids, and finally in Section V we summarize the work and explain the final conclusions.

II. MODEL

Th first thing we need to do is to find a mapping between the solution to the N-Queens problem and an Ising Hamiltonian [1]. For the N-Queens problem, we are able to write a Hamiltonian that does the job,

$$\mathcal{H}_{\text{target}} = \delta \sum_{ij,i'j'}^{N\times N} X_{ij}X_{i'j'} - \epsilon \sum_{ij}^{N\times N} X_{ij}, \qquad (1)$$

$$\delta, \epsilon > 0,$$

where X_{ij} is our Ising variable and can only be 1 if there is a queen at position (i, j) and 0 if there is not. The first term of $\mathcal{H}_{\text{target}}$ increases the energy by δ if at position (i, j) there is a queen and can be captured. The second term is our constrain, adding a queen decreases the energy by ϵ.

We also need to define the units of our problem. For the energy, δ is used; and for the time t_0 defined as $t_0 \equiv \hbar/\delta$ is used. If we write Eq. (1) in a way that we can actually see the units:

$$\mathcal{H}_{\text{target}} = \delta \left(\sum_{ij,i'j'}^{N\times N} X_{ij}X_{i'j'} - \frac{\epsilon}{\delta} \sum_{ij}^{N\times N} X_{ij} \right). \qquad (2)$$

We can appreciate that ϵ also is an energy, as we expected, but having an ϵ/δ ratio explicitly written will help us finding the best weight both terms should have.

A. Optimizing the ϵ/δ ratio

The N-Queen problem requires N queens placed in a N×N grid, $\epsilon/\delta < 1$ because if it is bigger or equal than 1, the ground-state of $\mathcal{H}$ could have more than N queens and that is not the problem we are trying to solve.

Knowing that $\epsilon/\delta < 1$ we must determine what is the value that increases the most the gap between the energy of the ground-state and the energy of the first excited state. If we use Eq. (1), the ground-state must have $E_{gs} = -N\epsilon$, and we have two different candidates to be the first excited energy: For the first candidate, we remove a queen. Now, we have $N - 1$ queens, thus increasing the energy in ϵ. The energy of this state is $E_\epsilon = -\epsilon(N-1) = E_{gs}+\epsilon$. And, for the second candidate, we add a queen, but now one of them can be captured.

So, we decrease the energy in ϵ but we increase it in δ. In this case we have $E_{\epsilon\delta} = -\epsilon(N+1) + \delta = E_{gs} - \epsilon + \delta$.

Dividing both candidates by ϵ we get $E_\epsilon/\epsilon = -N + 1$ and $E_{\epsilon\delta}/\epsilon = -N - 1 + \delta/\epsilon$, imposing $E_\epsilon = E_{\epsilon\delta}$ we get $\delta/\epsilon = 2$, thus $\epsilon/\delta = 0.5$. Having this ratio makes the difference between the energies of the ground-state and the first excited state as big as possible reducing the probability of transitioning from the ground-state to an excited state while we are doing the Annealing process.

B. Method

The dynamics of the problem follow the time-dependent Schrödinger equation,

$$i\hbar\frac{\partial\,|\phi(t)\rangle}{\partial t} = \mathcal{H}\,|\phi(t)\rangle\,, \tag{3}$$

Eq. (3) is the equation we want to solve, and to do so we are making use of the Cranck-Nicholson algorithm [3] that approximates $|\phi(t+\Delta t)\rangle$ in the following way:

$$|\phi(t+\Delta t)\rangle = \left(1 + i\frac{\Delta t}{2}\mathcal{H}\right)^{-1}\left(1 - i\frac{\Delta t}{2}\mathcal{H}\right)|\phi(t)\rangle\,. \tag{4}$$

We are using two different methods to find the solutions. Each method has a different $\mathcal{H}(t)$. For the first method we have:

$$\mathcal{H}(t) = \mathcal{H}_{\text{target}} - \Gamma(t)\delta\sum_{i=1}^{N\times N}\sigma_i^x, \tag{5}$$

where $\mathcal{H}_{\text{target}}$ is exactly Eq. (3) and $\Gamma(t)$ is a monotone descendant function where at $t = 0$ has a very big value making $\delta\sum_{i=1}^{N}\sigma_i^x \gg \mathcal{H}_{\text{target}}$. We need a δ before $\sum_{i=1}^{N}\sigma_i^x$ to be coherent with the units. And for the second method we have:

$$\mathcal{H}(t) = a(t)\mathcal{H}_{\text{target}} + b(t)\delta\sum_{i=1}^{N\times N}\sigma_i^x, \tag{6}$$

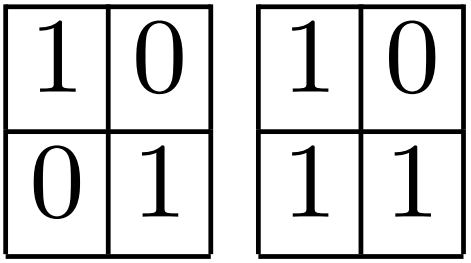
where $a(t) = 1, b(t) = 0$ when $t \longrightarrow \tau$ and $a(0) = 0, b(0) = 1$ and are called schedules. We are discussing how these methods work on the next sections.

Both method share a common term, $\delta\sum_{i=1}^{N\times N}\sigma_i^x$, this is the initial Hamiltonian, also known as the transverse Hamiltonian (we must have a δ to be coherent with units). This initial Hamiltonian forces all spins to point at the x direction, and our goal with QA is to get from the ground-state of this transverse Hamiltonian to the ground-state of $\mathcal{H}_{\text{target}}$.

In order to get better results, we must follow the Adiabatic Theorem [5] when using both methods. This theorem states that, if we start in the ground-state of the Hamiltonian at $t = 0$ and the time evolution is sufficiently slow, we will go through all the ground-states of $\mathcal{H}(t)$. For the method using $\Gamma(t)$ this implies that $\Gamma(t)$ must go slowly to zero, and for the method using schedules, the τ value must be large.

C. The states and grid notation

A state is one of the possible configurations that the grid can have. For example, take a 2×2 grid, two of the possible states could be:



These two states are written in grid notation. For each position we have two possibilities, to have a queen and to not have, so, for a generic N×N grid, we have $2^{N\times N}$ possible states. In our 2×2 example we have $2^{2\times 2} = 2^4$ possible configurations. It looks like we can translate the grid notation to a binary number notation. Using our examples, the first configuration (left) can be written as $1\cdot 2^3 + 0\cdot 2^2 + 0\cdot 2^1 + 1\cdot 2^0 = 9$, and the second configuration (right) can be written as $1\cdot 2^3 + 0\cdot 2^2 + 1\cdot 2^1 + 1\cdot 2^0 = 11$. Respectively, they will finally be written as $|9\rangle$ and $|11\rangle$.

III. SOLVING THE 3×3 GRID USING A $\Gamma(t)$ FUNCTION

This method, that has as a $\mathcal{H}(t)$ Eq. (5), does not actually reach our target Hamiltonian, but one as close as the target as we want. In general, we start from $t = 0$ with $\mathcal{H}(0) \approx \Gamma(0)\delta\sum_{i=1}^{N\times N}\sigma_i^x$, and we end up at $t = \tau$ with $\mathcal{H}(\tau) \approx \mathcal{H}_{\text{target}}$. Ideally, if $\tau \longrightarrow \infty$ we would strictly get $\mathcal{H}(\tau) = \mathcal{H}_{\text{target}}$. We are using dimensionless Γ functions as all the units are included in both the target and the transverse Hamiltonian.

We are interested in the fidelity and infidelity of the states we find throughout all the Annealing process and its expected energy. The fidelity is $|\langle\phi(t)|\phi_{\text{target}}\rangle|^2$ and the infidelity is $1 - |\langle\phi(t)|\phi_{\text{target}}\rangle|^2$. The expected energy is $\langle\phi(t)|\mathcal{H}(t)|\phi(t)\rangle$.

As we know the solutions of a 3×3 grid, we put a bias towards the state-solution $|12\rangle$ to make clearer comparisons of this method. Rather than having an energy of $E = -\delta$ as every other state-solution, $|12\rangle$ will have an energy of $E = -3\delta$. Our target state now is $|12\rangle$, $\mathcal{H}_{\text{target}}$ becomes $\mathcal{H}_{\text{biased}}$ and $\mathcal{H}(t)$ becomes $\mathcal{H}_{\text{biased}}(t)$.

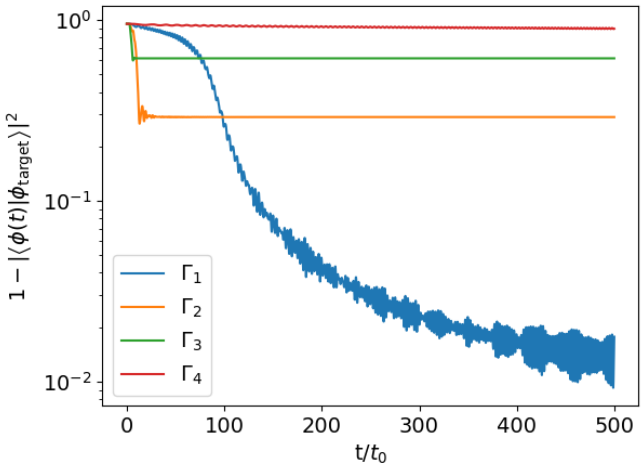


FIG. 1: Comparing the infidelity for different $\Gamma(t)$ depending on how fast this function goes to 0. Each plot is done using the same $\Delta t = 0.2t_0$.

In Fig. 1 we are studying how four different Γ functions perform. The functions are, also we can see them in Table I, $\Gamma_1(t) = \frac{50}{t+1}$, $\Gamma_2(t) = \frac{50}{t^2+1}$, $\Gamma_3(t) = \frac{50}{e^t}$ and $\Gamma_4(t) = \frac{50}{\sqrt{t+1}}$. They all start with the same Hamiltonian at $t = 0$ being $\mathcal{H}(0) = -50\delta \sum_{i=1}^{N\times N} \sigma_i^x$. The Γ functions that go to zero faster (i.e. Γ_2 and Γ_3) arrive to a lower infidelity value faster, but then they remain constant, they do not improve. Given that Γ_1 and Γ_4 are the slowest ones we can observe that, in the case of Γ_1, goes slower to an infidelity value than Γ_2 and Γ_3 but reaches a lower value, thus being better at the end. This should also be true for Γ_4 but it is much more slower than the other ones making the process very long in time just and we are not able to see the behavior that should be similar to the other three.

TABLE I: Expected energy and fidelity using Different $\Gamma(t)$. $\mathcal{H}_{target}(\tau)$ is different for each $\Gamma(t)$. Data is from Fig. 1. τ is the final time, and is the same value for all $\Gamma(t)$ at $\tau = 500t_0$.

$\Gamma(t)$	$\langle\phi(\tau)\|\mathcal{H}_{\text{biased}}(\tau)\|\phi(\tau)\rangle$ (δ)	$\|\langle\phi(\tau)\|\phi_{\text{target}}\rangle\|^2$
$\Gamma_1(t) = \frac{50}{t+1}$	-2.98	0.98
$\Gamma_2(t) = \frac{50}{t^2+1}$	-1.91	0.71
$\Gamma_3(t) = \frac{50}{e^t}$	0.14	0.39
$\Gamma_4(t) = \frac{50}{\sqrt{t+1}}$	-17.61	0.10

On Table I we compare the expected energy and fidelity for the states found at $t = 500t_0$ for all different Γ functions. As we can also see from Fig. 1 Γ_1 has a better fidelity and its expected energy differ only in 0.02δ from the target energy of -3δ, a difference of energies 10^{-2} orders of magnitude lower than Γ_2, whose difference of energy with the target is 1.09δ, the next best function.

IV. SOLVING THE 3×3 GRID USING SCHEDULES

here we use the method of Eq. (6) and has a clear difference with the method using $\Gamma(t)$. The difference is that using schedules we actually start at $t = 0$ with $\mathcal{H}(0) = \delta \sum_{i=1}^{N\times N} \sigma_i^x$ and end up at $t = \tau$ with $\mathcal{H}(\tau) = \mathcal{H}_{\text{target}}$. Both $a(t)$ and $b(t)$ are also dimensionless for the same reason as the Γ functions studied previously.

We are also interested in analyzing the fidelity, infidelity and expected energies of the found states like we did with the previous section.

A. Solving for $\mathcal{H}_{\text{biased}}$ using linear schedules

As we did in the previous section, we put a bias to the state-solution $|12\rangle$ with an energy $E = -3\delta$. We want to compare how τ affects the final result using the same schedules.

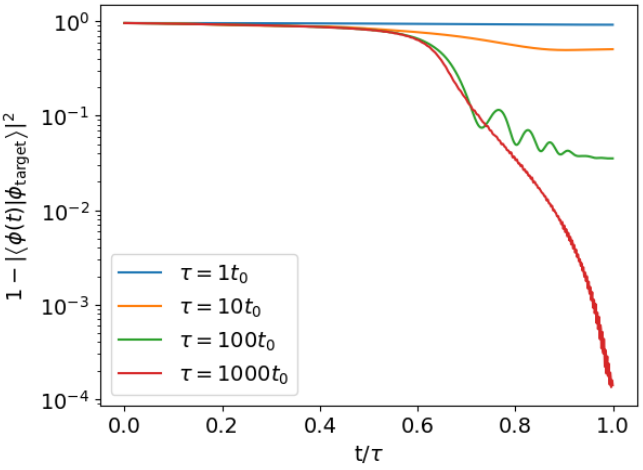


FIG. 2: Infidelity of the state found at t/τ depending on the value of τ. We use $\Delta t = \tau/10^4$ for all plots.

On Fig. 2 we can observe that all τ behave the same for $t/\tau < 0.5$ almost having a constant infidelity value at, approximately, 1, but after that the bigger the value of τ is, the faster the infidelity goes to zero.

TABLE II: Expected energy and fidelity for different τ values. As $\mathcal{H}(\tau) = \mathcal{H}_{\text{biased}}$ our ground-state energy is -3δ. Data is from Fig. 2.

τ/t_0	$\langle\phi(\tau)\|\mathcal{H}_{\text{biased}}\|\phi(\tau)\rangle$ (δ)	$\|\langle\phi(\tau)\|\phi_{\text{target}}\rangle\|^2$
1	2.2	0.08
10	-1.0	0.49
100	-2.86	0.97
1000	-2.999996	0.9999992

On Table II we can appreciate better the data from Fig. 2 at $t/\tau = 1$. Lower τ values imply a worse fidelity

and a worse expected energy from the state found. We can also note that the difference on fidelity between $\tau = 100t_0$ and $\tau = 1000t_0$ is much smaller than the difference between $\tau = 10t_0$ and $\tau = 100t_0$.

B. Solving for $\mathcal{H}_{\text{biased}}$ using different schedules

By leaving the τ value constant, now we are interested in comparing how the schedules affect the final result.

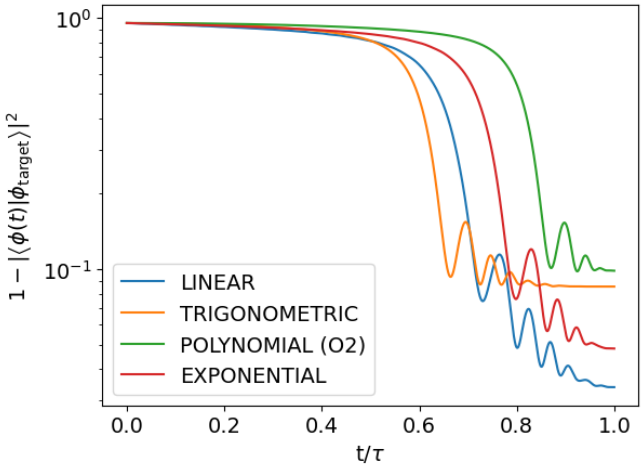


FIG. 3: Comparison of the infidelity for different schedules. We use $\tau = 100t_0$ and $\Delta t = \tau/10^4$ for all plots.

Fig. 3 shows how different schedules go to lower values of infidelity. We can see that, like in Fig. 2, all schedules have a similar value of infidelity at, approximately, 1 when $t/\tau < 0.5$. However, each schedule go to lower infidelity values at a different pace being the Linear ones the best at $t/\tau \approx 1$.

The schedules studied in Fig. 3 and Table III are Linear: $a(t) = t/\tau, b(t) = 1 - t/\tau$; Trigonometric: $a(t) = sin(\frac{\pi t}{2\tau})^2, b(t) = cos(\frac{\pi t}{2\tau})^2$; Polynomial (O2): $a(t) = (t/\tau)^2, b(t) = 1 - (t/\tau)^2$ and Exponential: $a(t) = 2^{t/\tau} - 1, b(t) = 2 - 2^{t/\tau}$.

TABLE III: Expected energy and fidelity using different schedules. As $\mathcal{H}(\tau) = \mathcal{H}_{\text{biased}}$ our ground-state energy is -3.0 δ. Data is from Fig. 3.

Type of schedule	$\langle\phi(\tau)\|\mathcal{H}_{\text{biased}}\|\phi(\tau)\rangle$ (δ)	$\|\langle\phi(\tau)\|\phi_{\text{target}}\rangle\|^2$
LINEAR	-2.86	0.97
TRIGONOMETRIC	-2.67	0.91
POLYNOMIAL (O2)	-2.62	0.90
EXPONENTIAL	-2.81	0.95

On Table III we analyze at $t/\tau = 1$ the expected energy and fidelity of the different schedules used. Note that all states found at $t/\tau = 1$ from the different schedules have a fidelity value equal or bigger than 0.9 making them good candidates for an Annealing process. Nevertheless, the Linear ones followed by the Exponential ones get better results while the Trigonometric and Polynomial (O2) differ only in 0.01 between each other and are far away from the results found with the Exponential schedules fidelity wise.

C. Solving for the target Hamiltonian

The main goal of having a biased Hamiltonian was to see better the differences that appear when we modify certain parameters of the QA. Now, we solve, using the linear schedules, the non biased Hamiltonian (i.e. $\mathcal{H}_{\text{target}}$). As there is no unique solution, the solution is 8-times degenerate, we can not compute a realistic fidelity.

The result we get is a superposition of all possible states, specially the solution-states. We can write this result state as $|\phi\rangle = \sum_i c_i |i\rangle$, where $|c_i|^2$ is the probability of measuring the i-th state, so $\sum_i |c_i|^2 = 1$.

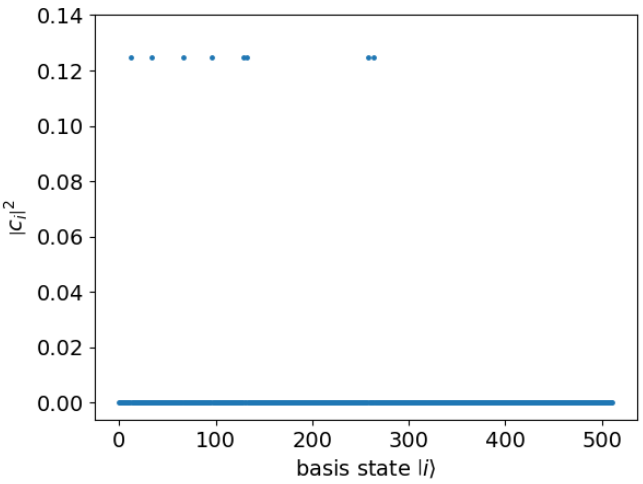


FIG. 4: Probability of measuring each state from the result found solving $\mathcal{H}_{\text{target}}$. The plot has been done with 10,000 points, $\tau = 100t_0$ and using the linear schedules. The energy found is $E = -0.998\delta$ and the "target" energy is $E = -\delta$.

If we take a look at Fig. 4, we can see that 8 states have almost the same $|c_i|^2$. The other states do have a non-zero value but are from order 10^{-5} or lower.

The 8 states that differ so much from the others all share the same $|c_i|^2$ value of 0.1247. So, if we sum all of them we get that the probability of measuring one of those 8 states is $\mathcal{P}_{8-states} = 0.998$. But the important thing they all share is the associated energy at $E = -\delta$, the "target" energy. Those 8 states are actually rotations of one of them as the system has so many symmetries, also, they have only two queens and are in a position they cannot capture each other, and are: $|12\rangle, |33\rangle, |66\rangle, |96\rangle, |129\rangle, |132\rangle, |258\rangle, |264\rangle$.

D. Energy spectrum of the 3×3 and 4×4 grids

In the previous subsections we were interested on how the overlap of our $|\phi(t)\rangle$ with the target state went to 1, but now we are interested on how the energy levels evolve as t/τ goes from 0 to 1. We are going to see this time-dependent spectre for the 3×3 grid and for the 4×4 grid.

Knowing this time-dependent spectre can help us identify points or regions susceptible of having transitions to higher energy levels and not ending up at the ground-state of the system but an excited energy level. If the method was ideally adiabatic, then, we would always go through the ground-state of $\mathcal{H}(t)$.

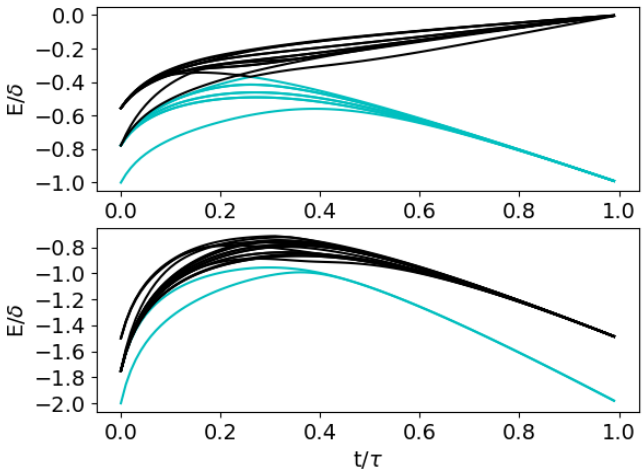


FIG. 5: Energy spectrum of both 3×3 (top) and 4×4 (bottom) grids for $t/\tau = 0$ to $t/\tau = 1$. The energy levels that end up to have the same energy as the ground-state of $\mathcal{H}_{\text{target}}$ are represented in blue, and the others, in black.

We get to see that in Fig. 5 there are at least five energy levels that go to the ground-state of the target Hamiltonian of the 3×3 grid and only two that go to the ground-state of the 4×4 grid. Indeed, only two energy levels go to the ground-state of the 4×4 grid as it only has two solutions. But, for the 3×3 grid, we get five energy levels and we should be getting eight. The reason that we can only see five is that at three levels are degenerate.

If we are following the ground-state of the transverse Hamiltonian, we should not worry about the evolution of the 3×3 grid as other seven (only four visible) levels go to the same energy. For the 4×4 grid, we must be more cautious as only two energy levels go to the ground-state of $\mathcal{H}_{\text{target}}$ and the energy levels are much closer throughout all $t \in [0, \tau]$.

V. SUMMARY AND CONCLUSIONS

Using QA we solved the N-Queens problem for the 3×3 grid, we already knew there was no possible way of putting 3 queens in a 3×3 grid without them killing each other, and found that only two queens can be put in a position where they cannot kill each other.

Both methods, $\Gamma(t)$ and the schedules, have helped us compare the results we obtained from QA with the true solution. The results obtained with the Annealing protocols is in agreement with the solution to the problem, which for the case considered can be checked easily. Additionally, we made the time-dependent energy spectre for both 3×3 and 4×4 grids to observe the differences and see critical points and regions.

In a future work, it would be nice to consider larger systems and implementing the protocol on a real annealer, such as the one of dwave systems or the future one at the Barcelona Supercomputing Center.

Acknowledgments

We want to thank B.Julià Diaz and A.Rojo Francàs for their time and help. Without them, this work would not have been possible at all. Also, a special mention to my family for their support.

[1] A. Lucas, Frontiers in Physics 2, 5 (2014).
[2] T. Kadowaki and H. Nishimori, Phys. Rev. E 58, 5355 (1998).
[3] Oscar Promio Muñoz, *Quantum Annealing in the transverse Ising Model* (2018) .
[4] `https://www.dwavesys.com/learn/quantum-computing/`.
[5] T. Albash and D. A. Lidar, Rev. Mod. Phys. 90, 015002
[6] `https://lyndenlea.uk/nqueens/?page=solutions`