Optimization through quantum annealing: theory and some applications

D.A. BATTAGLIA† and L. STELLA*‡

International School for Advanced Studies (SISSA), via Beirut 2–4, I-34014, Trieste, Italy, and INFM Democritos National Simulation Center, via Beirut 2–4, I-34014, Trieste, Italy

(Received 1 December 2005; in final form 8 June 2006)

Quantum annealing is a promising tool for solving optimization problems, similar in some ways to the traditional (classical) simulated annealing of Kirkpatrick et al. Simulated annealing takes advantage of thermal fluctuations in order to explore the optimization landscape of the problem at hand, whereas quantum annealing employs quantum fluctuations. Intriguingly, quantum annealing has been proved to be more effective than its classical counterpart in many applications. We illustrate the theory and the practical implementation of both classical and quantum annealing – highlighting the crucial differences between these two methods – by means of results recently obtained in experiments, in simple toy-models, and more challenging combinatorial optimization problems (namely, Random Ising model and Travelling Salesman Problem). The techniques used to implement quantum and classical annealing are either deterministic evolutions, for the simplest models, or Monte Carlo approaches, for harder optimization tasks. We discuss the pro and cons of these approaches and their possible connections to the landscape of the problem addressed.

1. Introduction

Many tasks of considerable practical importance can be rephrased as optimization problems, i.e. as problems in which the goal of finding the global optimum (or optima) of some specific function is being pursued. A useful metaphor, which will run as a sort of Ariadne’s thread all along our exploratory journey, is the concept of landscape. Imagine the awesome view from the top of some high mountain in the Alps. You will see a vast portion of land, carved by hundreds of valleys; some of them are very broad, occupied by glaciers, others more narrow, surrounded by rocky and fractured ridges. You will also see majestic gorges and canyons carved patiently over the centuries by the silent action of the melting snow. Even if it may sound bizarre, the flow of these small creeks finding their way through the corrugated and complex geometry of this alpine landscape might be seen as performing an optimization task. The water will stop flowing only when its gravitational potential energy is a minimum, taking into account the constraints imposed by the territory conformation (see figure 1). This may occur when the deepest possible valley (global optimum) is reached, or when the stream is not be strong enough to overcome the barrier created by lateral or end moraines, or even perhaps by an artificial dam (local optimum).

The evolution of many other physical systems is also governed by energy minimization. Molecules assume very specific configurations after a relaxation driven by many competing forces – a conspicuous success of recent advances in condensed matter physics has been to deal with...
the problem of efficiently simulating this structural optimization – and a freezing liquid crystallizes into a solid.

Other materials assume remarkable magnetic properties in some of their phases. A very instructive example is represented by the well-known Ising model of the ferromagnetic transition. The microscopic magnetic moments existing in a solid material (whose exact description and explanation would require a rather sophisticated quantum-mechanical treatment of the problem) are just modelled by binary spins, that can assume just two possible values, \( s = \pm 1 \). In the simplest version of the model, a spin is situated in each site of a cubic lattice and interacts with its neighbours according to the Hamiltonian:

\[
H_{\text{Ising}} = -J \sum_{\langle i,j \rangle} s_i s_j, \tag{1}
\]

where the sum is understood over all the pairs of neighbouring sites. If \( J \) is positive, energy can be minimized by aligning all the spins in the same direction. Two equivalent ground states exist, characterized by a large absolute value of the overall magnetization. Nevertheless, the number of excited configurations is enormously larger. A simple Ising model defined over a \( 32 \times 32 \) square lattice has about \( 10^{308} \) configurations against just two possible ground states minimizing its energy! The exact configuration assumed by the system will depend then on a delicate balance between entropy \( S \) and energy \( E \). Two regimes, paramagnetic, dominated by entropy (disorder wins, no long-range alignment of the spins is visible), and ferromagnetic, dominated by energy (global alignment of spins), will be observed at different temperatures \( T \). In particular, the equilibrium state will be such to minimise the so-called free-energy function, \( F = E - TS \), at the given system temperature.

The free-energy of the Ising model has very simple properties (see for instance [1]). As a function of the magnetization, it can be modelled by a fourth order polynomial, having just a single minimum (with null magnetization) above the Curie temperature and two equivalent minima with opposite magnetization below it (in this case the system is forced to choose one of the minima during its evolution toward the equilibrium, breaking then spontaneously the symmetry between the ‘up’ and the ‘down’ state).

A small modification of the Ising model accounts for enormously more complex free-energy landscapes. Let us consider the so-called spin glass Hamiltonian:

\[
H_{\text{SG}} = -J \sum_{\langle i,j \rangle} J_{ij} s_i s_j, \tag{2}
\]

Here the couplings themselves are link dependent and they are drawn randomly from some probability distribution (e.g. a uniform distribution). Nevertheless, they do not evolve in time and so, in this case, one typically speaks of ‘quenched disorder’. Now, the resulting free-energy function outside of the paramagnetic phase takes a very different form: it is characterized by an exponential number of local minima, separated between them by very large energy barriers. In order to relax into one of its ground states, the system must traverse several ranges of mountains, looking for hidden and rare passes, and at the risk of getting trapped for long times in narrow and deep valleys which are still far remote from a true global optimum (figure 1). It may also happen that the true equilibrium state is actually reached only after such long times, that it will never be observed in practice!

It is known that a better approach to finding the low-energy (optimum or near-optimum) states of such a system – which could be a simulated spin glass, or a metal that we want to produce with a particularly good crystalline structure – is
to cool it slowly starting from the molten state. At high temperatures, relatively large thermal fluctuations will allow the system, if it gets trapped in some local minimum, to jump over the surrounding barriers and so to search out better valleys. On the other hand, reducing the temperature will allow a detailed probing of a more restricted region, so that the system does not miss a good minimum simply by ‘flying above it’. Such protocol is called thermal (or classical) annealing – the use of the term originates in metallurgy.

Simulated thermal annealing is an efficient optimization method for finding the minima of arbitrary functions, not just the energy; this was understood for the first time by Kirkpatrick et al. in 1983 [2]. The most important ingredient is the fact that, at every simulated time step, there is a non-zero probability, \( p(\Delta E) \propto \exp(-\Delta E/T) \), to evolve in a direction which increases the energy (and free-energy, too) by a positive amount, \( \Delta E \); instead of always reducing it (for this reason, simulated annealing is also referred to as a ‘hill-climbing procedure’). This lack of (complete) ‘greediness’ – governed by the time-dependent ‘pseudo-temperature’ \( T \) – accounts for the ability of simulated annealing to escape from local minima, which would invariably trap other optimization heuristics such as the steepest-descent or conjugate gradient algorithms.

With this introduction, we can now see that quantum mechanics opens up a new possibility for optimization. Instead of simply jumping over an energy barrier, a quantum particle could attempt to tunnel across it (see figure 2)! Such a particle is sometimes said to be executing a ‘quantum walk’, by analogy with the random walk of classical Brownian motion. To see how this works, let us consider a quantum version of the spin glass Hamiltonian:

\[
\mathcal{H}_{\text{QSG}} = - \sum_{i,j} J_{ij} \sigma_i^z \sigma_j^z - \Gamma \sum_i \sigma_i^z.
\]

Here the classical Ising spins are replaced by Pauli operators, aligned along the \( z \) direction. A transverse external field \( \Gamma \) along the \( x \) direction is then added, acting as a source of quantum fluctuations: since the \( x \) and \( z \) components of the quantum spin operator do not commute, the effect of \( \Gamma \) is to induce flips of \( \sigma_i^x \), with a transition probability that increases with the magnitude of \( \Gamma \).

So, the action of the transverse field is to produce spin-flips; the natural question is then whether or not the exploration of an energy landscape directed by such quantum fluctuations can be more efficient than in the purely thermal case. If a system can be relaxed by gradually cooling it, one could also devise a quantum annealing strategy where, even at zero temperature, a system is driven to equilibrium by gradually decreasing the intensity of an external transverse field (or another suitable source of quantum fluctuations).

Recent experiments carried out by Gabriel Aeppli and collaborators [22] tried to address precisely this feature. They concentrated their observations on the relaxation of the doped magnetic compound \( \text{LiHo}_{0.44}\text{Y}_{0.56}\text{F}_4 \), in which magnetic impurities, randomly scattered across the sample, favour anti-parallel or parallel alignment according to their respective separations and orientations. The Hamiltonians equations (2) and (3) then constitute a first rough approximate model of the material’s magnetic behaviour. This material was allowed to relax towards its ground state under two different protocols. In the first (classical annealing, CA), the sample was slowly cooled subjected to a constant external transverse field, and eventually the external transverse magnetic field was suddenly varied. In the second one (quantum annealing, QA), an initially large field was instead gradually reduced at a constant room temperature, and only at the end of the field ramp was the system quenched to a low temperature. The results (based on an analysis of the magnetic susceptibility’s evolution during the annealing process) were consistent with the hypothesis that the quantum route to magnetic relaxation was faster than the classical one. We will return to the Aeppli’s experiment and its relation to the (simulated) QA of an Ising spin glass in section 3.

Finally, we want to mention some basic relationships between QA and the rapidly growing field of quantum computation (QC – for an introductory review on this topic, see [4]). In the standard approach to QC, information is stored in qubits – quantum two-level systems – instead of the customary \{0, 1\} classical bits, but the overall computation still looks ‘digital’ (i.e. it employs quantum gate operations in analogy with the classical theory of computation). However, this is not the only possible approach. For instance, adiabatic QC [4], is characterized by a completely analogue way of manipulating quantum information. As in QA, this protocol starts from an initial state prepared as the ground state of a ‘simple’ quantum Hamiltonian and eventually, through an almost adiabatic sequence of unitary evolutions, provides the solution of the ‘complex’ problem under inspection (e.g. the optimization of a rough cost energy). These unitary steps can be either elementary gate

Figure 2. Sketch of the way classical and quantum annealing can relax from a minimum to another one, separated by a barrier.
operations or ‘natural’ evolutions of less abstract objects, e.g. quantum magnets. In other words, this route to QC is similar to the early mechanical analogue computation, in a spirit very similar to Vannevar Bush’s pioneering work on the differential analyser [5]. Nobody yet knows what form a possible future implementation of a QA-based quantum computer might take. Maybe, we will be able one day to build artificial systems of spin-like degrees of freedom coupled in a reconfigurable, controllable way and also subject to a tunable external field (see, for instance [6]). The input to a computational task might then be the configuration of the hardware itself and the computation would simply correspond to waiting for a time long enough to allow the quantum relaxation to take place in a variable field. Unfortunately, such quantum devices (like any other implementation of a really useful quantum computer) still belong to the realm of science-fiction.

It is also worthwhile remembering here – to avoid confusion between the goals of QA and QC – that QA remains a valuable, generic scheme to devise optimization algorithms which can run effectively (with respect to other known algorithms) on the usual ‘classical’ computer. Indeed, many different kinds of QA have been tested so far – in the rest of the paper we shall present many of them. In particular, from the next section on, we will meet some of the highly non-intuitive behaviours of a quantum walker in a simple multi-minima landscape. This excursion will also lead us through the properties of two of the most important equations of mathematical physics, namely the Fokker–Planck and Schrödinger equations. Only then will we be ready for the exploration of more fascinating, but also dangerously unknown, complex energy landscapes.

2. A toy-model for classical and quantum annealing: the double-well potential

We will focus for now on continuous optimization problems only. Given a potential \( V(x) \) (where \( x \) is a vector of arbitrary dimensionality), we want to determine the absolute minimum position, \( x_{\text{opt}} \), such that \( E_{\text{opt}} = V(x_{\text{opt}}) \leq V(x) \) \( \forall x \). For the sake of simplicity, in the rest of this section we shall suppose that this absolute minimum is unique; the extension to the degenerate case can be obtained by a straightforward generalization.

If the potential had only a single minimum, simple methods taking advantage of the knowledge of the (opposite) gradient of \( V \) (such steepest-descent or conjugate-gradient [7]) would be effective. But a more challenging goal is to find the absolute minimum of a potential which displays many local minima (e.g. metastable states); now, the application of the simple gradient-based techniques would very probably lead to trapping into one of those metastable minima.

Usually, many attempts are needed (changing, for instance, the initial conditions) in order to maximize the probability of hitting the global minimum. Indeed, chance is a key ingredient of many efficient optimization methods. When the potential is rough, it might be better to explore with a certain degree of randomness, instead of making too strict a hypothesis about where the minimum should be.

In the next section we shall introduce some basic notions required to describe classical and quantum randomized local searches. We will proceed then, to their comparison, analysing in some detail the toy example of the simplest multi-valley landscape, the double-well potential.

2.1 The theory of classical annealing: the Fokker–Planck equation

As a natural consequence of the randomized search point-of-view, one can consider the dynamics of the distribution \( P(x, t) \), i.e. the probability of finding a particle at position \( x \) at time \( t \), instead of a deterministic evolution over the phase space. An equation of motion for \( P(x, t) \) which is appropriate to implement a stochastic optimization method is the so-called Fokker–Planck (FP) equation (9):

\[
\frac{\partial}{\partial t} P(x, t) = \text{div}(P \nabla V) + D(T) \nabla^2 P. \tag{4}
\]

In a nutshell, the Fokker–Planck equation is made of two terms which account for two different physical phenomena. If the second term is neglected, equation (4) becomes:

\[
\frac{\partial}{\partial t} P(x, t) = \text{div}(P \nabla V), \tag{5}
\]

which is equivalent to the well-known continuity equation with a probability current, \( J_P = -P \nabla V \), which is proportional to the opposite of the potential gradient (i.e. the ‘force’). Therefore, equation (5) reflects just the deterministic flow of the system according to the simple differential equation \( \dot{x} = -\nabla V \). This could be interpreted as a probabilistic mathematical description of a pure steepest-descent dynamics.

On the other hand, by neglecting the first term of equation (4), one obtains the well-known heat equation:

\[
\frac{\partial}{\partial t} P(x, t) = D(T) \nabla^2 P, \tag{6}
\]

where \( D(T) \) is the so-called diffusion constant, which governs the rate of free diffusion under the influence of fluctuations from the environment.

A crucial quantity is the asymptotic (i.e. large-\( t \)) equilibrium distribution associated to the FP equation. It turns out that this equilibrium is described by the well-known Boltzmann–Gibbs distribution:

\[
P_{\text{eq}}(x) = \frac{e^{-V(x)/(k_B T)}}{\int dx' e^{-V(x')/(k_B T)}}, \tag{7}
\]
provided that \( k_B T = D(T) \) (in the following discussion the Boltzmann constant will be set for simplicity to \( k_B = 1 \)). As we know from the basic statistical mechanics, such a distribution is characterized by peaks centred at every minimum of the potential, with a height depending on the minimum depth and a width depending on the fixed parameter \( T \) (once more, the temperature is simply understood as an tunable external parameter which governs fluctuations).

In the context of optimization, it may not be completely clear what actually is the thermal reservoir interacting with the system. Nevertheless, completely artificial thermal-like fluctuations turn out to be the key ingredient of the simulated classical annealing (CA) scheme we want to implement. CA is mimicked by initially equilibrating the system for a long enough time at a large temperature \( T_0 \) (leading to an essentially flat initial distribution), and then gradually decreasing \( T \) to zero in a given annealing time \( \tau \).

We can model this schedule by assuming that

\[
T(t) = T_0 f(t/\tau),
\]

where \( f(y) \) is some chosen monotonically-decreasing function of \( y \in [0, 1] \), with \( f(0) = 1 \) and \( f(1) = 0 \). The simplest case is the so-called linear schedule, with \( f(y) = 1 - y \), as sketched in figure 3.

As an important consequence, the diffusion constant \( D(T) \) in equation (4) becomes time-dependent, \( D(t) = D(T(t)) \), and so the whole FP equation becomes a time-dependent partial differential equation. This is a source of some technical complication and, in particular, it is no longer guaranteed that a Boltzmann–Gibbs distribution will be the final output of CA.

The main quantity of interest, the final average potential compared to the absolute minimum \( E_{\text{opt}} \) (the residual energy), is simply given by:

\[
\epsilon_{\text{res}}(\tau) = \int dx \, V(x) \, P(x, t = \tau) - E_{\text{opt}} \geq 0. \tag{9}
\]

Following an infinitely slow annealing schedule, the system would relax in each infinitesimal interval so that the probability density \( P(x, t) \) could be considered almost equal to the instantaneous equilibrium distribution (i.e. \( P(x, t) \approx P_{\text{eq}}(x; T(t)) \)). If this adiabatic condition were always preserved, the system would be perfectly localized in the absolute minimum at the end of every annealing process (and so \( \epsilon_{\text{res}}(\tau) = E_{\text{opt}} \)). For \( T = 0 \), indeed, the Boltzmann–Gibbs distribution becomes a Dirac delta function.

Unfortunately, this is only an asymptotic condition, (i.e. it holds only for \( \tau \to \infty \)). For finite-time annealings, the system’s evolution will generally lose its adiabaticity at some finite critical temperature, \( T_c > 0 – a – phenomenon \) similar to the freezing of a disordered system [9]. Nevertheless, the quality of the outcome distribution achieved by means of CA is often acceptable, and in many practical cases CA deserves to be employed, at least as a first optimization attempt [2].

2.2 Quantum annealing in real and imaginary time

Quantum annealing can be modelled by using a fictitious Schrödinger dynamics for the probability distribution \( P(x, t) \), instead of a FP one. The Schrödinger equation:

\[
i \hbar \frac{\partial}{\partial t} \psi(x, t) = \left[ -\frac{\hbar^2}{2m} \nabla^2 + V(x) \right] \psi(x, t) = H \psi(x, t), \tag{10}
\]

is, more precisely, an equation of motion for the probability amplitude \( \psi(x, t) \). However, this quantity is associated with a probability distribution through the equation:

\[
P(x, t) = \frac{\left| \psi(x, t) \right|^2}{\int dx' \left| \psi(x', t) \right|^2}. \tag{11}
\]

In the following, we shall consider a slightly modified version of the Schrödinger equation, which reads:

\[
-\hbar \frac{\partial}{\partial t} \psi(x, t) = \left[ -\frac{\hbar^2}{2m} \nabla^2 + V(x) \right] \psi(x, t) = H \psi(x, t). \tag{12}
\]

This second form, equation (12) – similar to the heat equation (see equation (6)) – is easily obtained from equation (10) by the change of variable \( t \to -it \), and, for this reason, it will be referred to as the imaginary time Schrödinger equation (IT), in order to be distinguished from the more usual real time version (RT), equation (10).
To mimic QA, we allow for a time-dependent particle mass and so we write \( \Gamma(t) = \frac{d}{dt} \). To understand why this is, notice that the kinetic energy operator, \(-\Gamma(t) \nabla^2\), does not commute with the potential operator \( V(x) \), and it therefore can be seen as the source of the quantum fluctuations in configuration space (i.e. as responsible, for instance, for the zero-point energy of a quantum particle confined in a well – reminiscent of the effect of transverse magnetic field in the experiment of [22]). As a consequence, the coefficient \( \Gamma(t) \) of the Laplacian in equation (10) plays the same role as the diffusion coefficient in the FP equation; it will be taken as our (quantum) annealing parameter. (We set \( \hbar = 1 \).

As with the temperature in CA, \( \Gamma(t) \) will start from a large value \( \Gamma_0 \) at \( t = 0 \) – corresponding to a small particle mass, and hence to large quantum fluctuations – and will decrease down to \( \Gamma(t = \tau) = 0 \), corresponding to a particle of infinite mass (hence unaffected by quantum fluctuations). Once again we will select a monotonically decreasing function \( \Gamma(t) \), say linear. In the case of QA, a suitable initial condition would be \( \psi(x, t = 0) = \psi_0(x; \Gamma_0) \), where \( \psi_0(x; \Gamma_0) \) is the ground state of the system when the annealing parameter is equal to \( \Gamma_0 \). For such a large \( \Gamma \), the ground state should be approximately the ground state of the kinetic term only, which corresponds to a simple completely delocalized probability distribution. Finally, the residual energy after QA can be still computed by means of equation (9), where the probability \( P(x, t = \tau) \) is now given by equation (11).

In order to test fairly the relative efficiency of CA and QA, one needs a simple and controlled potential landscape. In the next section, we shall present such comparison, stressing the non-trivial behaviour which characterizes the QA.

### 2.3 Annealing in double-well potential

Let us illustrate the behaviour of the CA and QA in the simplest non-trivial case: a one-dimensional double well. We will take a barrier separating two non-equivalent minima, one deeper than the other. A simple potential of this type can be written:

\[
V(x) = \begin{cases} 
V_0 \frac{(x_0^2 - x^2)^2}{a_x^2} + \delta x & \text{for } x \geq 0 \\
V_0 \frac{(x_0^2 - x^2)^2}{a_x^2} + \delta x & \text{for } x < 0 
\end{cases}
\]

(13)

with \( V_0 > 0, a_x > 0, a_- > 0 \), and \( \delta \) real constants (the possible discontinuity in the second derivative at the origin is of no consequence in our discussion). In absence of the linear term \( \delta = 0 \), the potential has two degenerate minima located at \( a_- \) and \( a_+ \), separated by a barrier of height \( V_0 \). When a small linear term \( \delta > 0 \) is introduced, on the other hand, the two minima are located (to linear order in \( \delta \)) at \( x_\pm = \pm a_\pm - \delta a_\pm^2/(8 V_0) \), and the splitting between them is given by \( \Delta E \approx \delta (a_+ + a_-) \) (favouring slightly the minimum at \( x_+ \approx a_- \)). Moreover, whenever \( a_- \neq a_+ \), the curvature of the two wells will be different. The second derivatives at the minima – again to lowest order in \( \delta \) – are given by:

\[
V''(x = x_\pm) = \frac{8 V_0}{a_\pm^2}.
\]

(14)

For reasons that will be clear in a moment, it is useful to consider this general asymmetric configuration. In particular, we shall examine the case in which the metastable ‘valley’ at \( x_+ \) is ‘wider’ than the absolute minimum at \( x_- \), which is obtained by setting \( a_+ > a_- \). The Fokker-Planck and the Schrödinger equation (both in RT and in IT) were integrated numerically using a standard fourth-order Runge-Kutta method (see, for instance, [7]). In both cases, we set a linear schedule for the reduction of the annealing parameter.

Figure 4 shows the results for the final annealed probability distribution \( P(x, t = \tau) \) at different values of the annealing time, \( \tau \), for both the Fokker-Planck (CA, panel (a)) and the Schrödinger imaginary-time case (QA-IT, panel (b)), for a double-well potential, equation (13), with \( V_0 = 1 \), \( a_+ = 0.75 \), \( a_- = 1.25 \), \( \delta = 0.1 \). Figure 4(c) summarizes the results obtained for the residual energy \( \epsilon_{\text{res}}(\tau) \).

First of all, we notice that in both cases, for large enough value of \( \tau \) (i.e. very slow annealing) QA provides a smaller residual energy than CA (see figure 4(c)). Moreover, the RT-QA and its IT counterpart have the same asymptotic behaviour, but they have a different appearance at intermediate values of the annealing time (figure 4(c)).

Let us first focus on CA. Starting from an initially broad Boltzmann-like distribution, \( P(x, t = 0) \) (solid lines), at a high temperature, \( T_0 = V_0 \), the system sharpens its localization and quickly develops a \( P(x, t = \tau) \) with two well-defined and quite narrow peaks located around the two minima, \( x_\pm \), of the potential. If we denote by \( p_\pm \) the integral of each of those two narrow peaks (with \( p_+ + p_- = 1 \)), it becomes clear that the problem has effectively been reduced to a discrete two-level system. After this point, any further increase in the annealing time \( \tau \) will cause only a very slow increase in the probability \( p_- \), and this behaviour represents a crucial limitation of CA performance. Every interested reader is referred to the work of Huse and Fisher, [9], which showed that the asymptotic value of the residual energy (apart from logarithmic corrections) is given by:

\[
\epsilon_{\text{res}}(\tau) = \Delta E + \frac{\tau}{2} B, \quad \text{for } \tau \gg 1,
\]

(15)

where \( \Delta E \) is the energy drop between the two minima and \( B = V_0 - V(x_-) \) is the energy drop between the metastable minimum and the barrier top. The solid line in figures 4(c) is an interpolation of the numerical data based on the Fokker-Planck equation and the expected behaviour,
equation (15). We found good agreement between theory and simulations. We also emphasize that, by reducing the linear term coefficient $\delta$ in equation (13), one can make the exponent in equation (15) vanishingly small, eventually leading to an exceedingly slow CA dynamics.

The behaviour of the QA evolution is remarkably different. The initial squared wavefunction $|\psi(x, t=0)|^2$ (solid line in panel (e)) corresponds to a quite small mass ($\Gamma_0 = 0.5V_0$), and is broad and delocalized over both minima. If now we perform a relatively fast annealing – that is, if $\tau > \tau_c$, with a characteristic time $\tau_c$, which depends on the annealing kind, RT or IT (see figure 4(c)) – the final wavefunction will become mostly concentrated on the wrong minimum (see figure 4(b), QA-IT with $\tau = 100$!). For $\tau > \tau_c$, on the other hand, the final wavefunction gets more localized around the deepest minimum (see figure 4(b), QA-IT with $\tau \geq 460$). Therefore, QA is able to localize the global minimum of this simple potential landscape only if the annealing schedule is slow enough. This behaviour is rather unexpected, since the final probability distribution in the case of CA always concentrates more weight in the deeper minimum, although it remains bimodal (see figure 4(a)). One can conclude that some non-trivial mechanism is acting in the case of QA. The simplicity of the double-well potential allowed us to derive a full theoretical explanation, which is described in the next section.

2.4 Curvature-induced effects and Landau–Zener tunnelling

A well established fact about elementary quantum mechanics is that a quantum particle confined in a narrow box has a zero-point energy larger than that of an identical particle confined in a wider box. In a similar way, for small masses (i.e. large $\Gamma_0$), the kinetic energy difference between two wavefunctions localized in different valleys of our asymmetric $V(x)$, can even overcome the potential energy drop between the minima of these two valleys. As a consequence, for large $\Gamma_0$, the instantaneous ground state of the Hamiltonian $H$, defined in equation (10), will prefer to localize into the metastable minimum in order to minimize the total energy. This effect is visible in figure 5(b), fixing our attention only on the ground state obtained for $\Gamma = 0.04$. On the other hand, for very large masses (i.e. very small $\Gamma$) the kinetic energy is vanishing small and a quantum particle will prefer to localize in the deepest minimum. This qualitative behaviour is confirmed by numerical diagonalization results reported in figure 5(b), namely for $\Gamma = 0.035$ and $\Gamma = 0.005$. In particular, in the last case, the wavefunction is completely localized in the deepest minimum, unlike the CA, which always showed a bimodal final probability distribution.

These findings strongly suggest that at intermediate values of $\Gamma$ there should be a level crossing in the instantaneous energy spectrum and it is reasonable to expect that...
this phenomenon should govern the tunnelling process between the two potential wells. Looking at figure 5(a) for $/C_0' = 0.038$ (indicated by an arrow in the figure), one can see (magnified in the inset) that there is not a true crossing, but an avoided one (due to the absence of any symmetry in the potential). Put another way, the degeneracy of the states in the two wells near the transition point is split by the tunnelling amplitude between them. Such avoided level-crossings have been known for many years (see, among the others, [10]) and get their name from Laundau and Zener, who described their consequences in the early 1930s.

It turns out that the presence of a Landau–Zener (LZ) crossing in the low-lying part of the spectrum is the main source of adiabaticity loss. In other words, during QA, the wavefunction $\psi(x, t)$ remains almost equal to the instantaneous ground state $\psi_0(x; \Gamma)$ of the Hamiltonian $H$ until the avoided level crossing is approached. Then there is a finite probability of missing the correct ground state and going straight on into the first exited state (see figure 6).

Adiabaticity loss due to LZ effect occurs with a probability $P_{ex} = e^{-t_{LZ}/t}$, where $t_{LZ}$ is the characteristic time of the LZ process [10]. For our realization of the double-well potential, we analytically found $t_{LZ} \approx 18,980$ [11], which remarkably enough is equal to the characteristic time $\tau_c$ introduced above for the RT-QA (see figure 4(c)).

We stress that a true LZ crossing occurs only for an asymmetric double-well potential. The spectrum of the symmetric case, when $a_\pm = a_\pm$ and considered in [11], presents only a gentle approach of the two low-lying eigenvalues (varying $/C_0'$) and the squared instantaneous ground state is always concentrated in the deeper minimum, even for very fast annealing.

The interested reader can find a detailed account of the LZ phenomenon in Landau’s textbook, [10], or in Zener’s original article, [12], which is particularly clear. Since these original works were devoted to molecular mechanics, it is not surprising that LZ theory applies unmodified to the RT-QA case. Nevertheless, it is also possible to extend those results slightly to explain the QA-IT behaviour also.

In a nutshell, changing from RT to IT implies that we consider a non-unitary evolution operator, $e^{-iH't}$, instead of the usual unitary one, $e^{-iHt}$. This non-unitary IT evolution does not preserve the norm of the wavefunction. In particular, it will filter out the excited states, giving further help to the system in following the instantaneous ground state. Let us consider, for instance, the static case (i.e. $/C_0$ fixed).

Given a complete set of eigenvectors, $|\varphi_n\rangle$, of the Hamiltonian $H(\Gamma)$, one can expand the initial state as follows:

$$\psi(x, t = 0) = \sum_n a_n \psi_n(x),$$

Figure 5. Spectral properties of the Quantum Annealing. Instantaneous eigenvalues (a) and ground state wavefunctions (b) of the Schrödinger problem $H\psi = \omega \psi$ for different values of $\Gamma$, in the case of the double-well potential, equation (13), obtained by substituting $a_\pm = 0.75, a_\pm = 1.25, V_0 = 1$ and $\delta = 0.1$ in equation (13). Notice the clear Landau–Zener avoided crossing in (a), indicated by the arrow and magnified in the inset.

Figure 6. Sketch of Landau–Zener tunnelling. Passing through the LZ point (i.e. minimal gap, indicated by a vertical arrow) there is a finite possibility to lose adiabaticity, passing from the ground state to the first excited one (process depicted by a diagonal arrow).
where \( \{ a_n \} \) are real coefficients. Therefore the wavefunction at time \( t' \), according the IT evolution, is given by:
\[
\psi(x, t = t') = \sum_n \tilde{a}_n \psi_n(x),
\]
where \( \tilde{a}_n = e^{-\tau} a_n \) and \( \epsilon_0 \leq \epsilon_1 \leq \epsilon_2 \cdots \) are the eigenvalues of the Hamiltonian \( H \). Finally, under the stronger assumption that \( \epsilon_0 < \epsilon_1 \) (which holds in the case under inspection) one finds that \( \tilde{a}_n/\tilde{a}_0 \to \tilde{a}_{n,b} \). We always found that the IT residual energy is less than or equal to the RT one, \( \epsilon_{res}^T \leq \epsilon_{res}^IT \), and also that their residual energies show the same asymptotic behaviour (an analytical fit, \( \epsilon_{res}(\tau) \propto \tau^{-1/3} \), is indicated by the dashed line in figure 4(c)). We emphasize that this exponent is universal (interested readers will find more details about the appearance of this rather strange exponent in [11]) and it is not influenced by the bare geometry of the double-well potential (e.g. the barrier height, minimum energies, curvatures, etc.). It refers, indeed, to the quantum relaxation into a single well.

Summarizing, we have found that QA and CA proceed in a remarkably different way. CA is sensitive to the height of the barrier, and more precisely to the ratio \( \Delta V/\beta \) between the energy offset \( \Delta V \) of the two minima, and the barrier height \( \beta \). On the other hand, QA depends crucially on the tunnelling probability between the two valleys, which is reflected in the LZ (avoided crossing) time, \( \tau_{LZ} \). This quantity itself depends on the gap between the ground and the first excited state close to the (avoided) crossing: the smaller is this gap, the more severe will be the adiabaticity loss during QA. Finally, we noticed that RT and IT proceed with somewhat different characteristic times, but have the same asymptotic behaviour.

The main drawback of the semi-analytic approach presented in this section is the impossibility of straightforwardly extending it to less trivial potentials, such as those in real-world optimization problems. We will see in the following how the QA of actual complex systems can be simulated using approximate Monte Carlo schemes.

3. Quantum annealing of an Ising glass

In the introduction we mentioned Aeppli’s experiment on the magnetic properties of the disordered magnetic compound LiHo\(_{0.44}\)Y\(_{0.56}\)F\(_4\) as an example of the experimental implementation of QA. Here, we shall present a way to simulate that system. The simplest model that captures the complexity of its magnetic behaviour is an Ising spin glass in a transverse magnetic field, which has been already introduced at the beginning of this paper. Recall that its Hamiltonian is
\[
\mathcal{H}_{QSG} = -\sum_{\langle i,j \rangle} J_{i,j} \sigma_i^+ \sigma_j^- - \Gamma \sum_i \sigma_i^z = H_d + H_{Ising},
\]
where \( J_{i,j} \) are random couplings between nearest-neighbour (say, on a square lattice) sites \( \{ i, j \} \); \( \sigma_i^+ \) and \( \sigma_i^- \) are Pauli matrices for the spin at site \( i \), \( H_{Ising} = -\Gamma \sum_i \sigma_i^z \) and \( \Gamma \) is the transverse field inducing quantum fluctuations. We also recall that determining the ground state of an Ising spin glass (even without transverse field) can be an extraordinarily difficult task since the disorder and the frustration (i.e. the possibility of having either positive or negative couplings in equation (18)) lead to a proliferation of metastable minima [13]. On the other hand, for large \( \Gamma \), the quantum ground state of the Hamiltonian is well approximated by a field-aligned ferromagnetic solution in which all the spins are independently aligned to the external field, which has the simple form:
\[
|\psi_{Para} \rangle \propto \prod_{i=1}^N (|\uparrow \rangle + |\downarrow \rangle),
\]
where \( i = 1, \ldots, N \) runs over the lattice sites and \( \{ |\uparrow \rangle, |\downarrow \rangle \} \) represent the ‘up’ and ‘down’ states of the \( i \)-th spin, respectively. Since the ground state at very high transverse field is simple, one can try to apply a QA scheme using the transverse field as annealing parameter. Unfortunately, a deterministic approach similar to that presented in the previous sections is applicable only to very small instances of the problem (recall from the introduction that the number of possible states of an Ising model can be astronomical). Since the actually hard instances of a typical optimization problem involve a large number of variables (one can consider up to 80 \( \times \) 80 square lattices), there is an obvious need for an alternative strategy.

As pointed out by Kirkpatrick et al. in their first application of CA [2], a Monte Carlo (MC) scheme can be used to overcome this difficulty. More recently, quantum MC algorithms have been successfully employed as a tool for performing QA simulations of various systems [14–17].

We first review briefly the basic MC approach, the Metropolis algorithm, and present its basic applications. Let \( x_i \) be the configuration of the system at time \( t_s \) and \( \{ x_0, x_1, \ldots, x_s, \ldots \} \) the random sequence obtained according to the MC evolution (time is discretized in arbitrary units). The transition probabilities from \( x_i \) to any given \( x_{i+1} \) are supposed to be Markovian, i.e. they do not depend on the previous history \( (t < t_s) \) of the system, and they are summarized by a transition matrix, \( T(x_{i+1}, x_i) \), which completely describes the process. Provided that such a matrix satisfies a few constraints [19], there is considerable freedom in choosing its form.

As a matter of principle, \( T(x_{i+1}, x_i) \) could be a very complicated function, but – in every practical implementation – it is customary to deal with a simplified transition matrix, \( \tilde{T}(x_{i+1}, x_i) \), which is often called the proposal move matrix. It generates a set of proposed changes to the configuration \( x_i \). The algorithm produces a probability distribution that converges to the right equilibrium state if
every proposed displacement from \( x_i \) to \( x_{i+1} \) is finally accepted with a probability
\[
A(x_{i+1}, x_i) = \min \left( 1, \frac{T(x_{i+1}, x_i) P_{eq}(x_{i+1})}{T(x_i, x_{i+1}) P_{eq}(x_i)} \right)
\] (20)

From a rigorous mathematical point of view, this choice is equivalent to the decomposition \( T(x_{i+1}, x_i) = A(x_{i+1}, x_i) T(x_{i+1}, x_i) \) \( \forall i \neq j \) of the exact transition matrix. Moreover, the so-called detailed balance condition [18] \( T(x_{i+1}, x_i) P_{eq}(x_i) = T(x_i, x_{i+1}) P_{eq}(x_{i+1}) \) must hold.

Whenever \( P_{eq}(x) \propto e^{-\beta H(x)} \), the Metropolis algorithm provides a valuable method to study the (classical) thermal properties of a given physical system. In particular, it turns out that the thermodynamical average of every local observable \( O(x) \) can be obtained by the formula:
\[
\langle O \rangle = \lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} O(x_i)
\] (21)
given that at time \( t_i \) the system is already equilibrated (i.e. \( P(x_i) = P_{eq}(x_i) \) if \( t_i > t_j \)).

In order to implement a simulated CA, the Metropolis algorithm requires only minor changes. For a Ising spin glass (without transverse field), one can propose single spin-flips and accept or reject them according to the aforementioned Metropolis scheme. When a number of spin-flips equals to the size of the system (a so-called MC sweep) has been proposed, the (classical) annealing parameter, \( T \), is usually decreased following a given (say, linear) annealing schedule. A new sweep is then performed, and so on until \( T \) goes to zero. Not surprisingly, the equilibrium state at high temperature is paramagnetic and can be prepared by taking \( \langle |1\rangle \rangle \) or \( \langle |0\rangle \rangle \) with equal probability. Since the Metropolis algorithm is inherently stochastic, one takes many repetitions of every annealing process, starting from different (random) initial conditions, in order to collect reliable statistics from which one can further extract the desired averages (e.g. the residual energy). For a more detailed discussion of the error due to the unavoidable statistical correlations, the reader is referred to [18].

Analogously, in order to implement QA in a stochastic fashion one can resort to a quantum MC algorithm. Different schemes for quantum MC simulations have been proposed; the path-integral MC (PIMC) based QA has proven to be very effective in the case of Ising spin glass in transverse field [19].

PIMC is intended to simulate the equilibrium behaviour of a quantum system at finite temperature \( T \), taking advantage of a very clever rewriting of the quantum partition function:
\[
Z(T, \Gamma) = \text{Tr} e^{-\beta (H_\omega + H_{\omega})} = \sum_{\{s\}} \langle \{s\} | e^{-\beta (H_\omega + H_{\omega})} | \{s\} \rangle
\] (22)
where \( s \) denotes a generic configuration of all the \( N \) spins. We recall that \( H_\omega \) is the source of the quantum fluctuations and does not commute with \( H_\omega \), the Hamiltonian of the optimization problem (i.e. the potential energy landscape).

The idea is to reduce equation (22) to a classical partition function which can be sampled in the usual way employing, for instance, a Metropolis MC scheme. To do this, we split up the exponential of the Hamiltonian, appearing in equation (22), into products of other exponentials. This is allowed by the Trotter theorem, which states that:
\[
e^{-\beta (H_\omega + H_{\omega})} = \lim_{P \to \infty} \left( e^{-\beta H_\omega} e^{-\beta H_{\omega}} \right)^P.
\] (23)

Using this equation, and inserting resolutions of the identity between the various exponentials, after few simplifications [19, 20] one gets:
\[
Z(T, \Gamma) \propto \sum_{\{s\}} \prod_{P} e^{-\frac{P}{\beta} S} \tag{24}
\]
\[
S = -\sum_{k=1}^{P} \left( \sum_{\langle i,j \rangle} J_{ij} s_i s_j + J_{\Gamma} \sum_{i=1}^{\infty} s_i s_{i+1} \right), \tag{25}
\]
which represents the partition function of a classical (anisotropic) Ising system at temperature \( P/\beta = P T \), with an extra dimension. This new transverse coordinate, labelled by \( k \), accounts for the copies \( s^k \) \((k = 2, \ldots, P)\), often referred to as Trotter replicas of the original configuration \( s \) introduced by means of equation (23). The system has now couplings \( J_{\Gamma} \) along the original lattice bonds (the same for all Trotter slices), and \( J_{\Gamma} = d \ln (\tanh \beta(P)) \) (the same for all sites \( i \)) along the extra dimension; along this new direction, the system has a finite length, \( P \), and periodic boundary conditions. The appearance of multiple copies of a classical system is strongly reminiscent of the path-integral interpretation of quantum mechanics, according to which a quantum system is not just following the least action trajectory, as in the classical case, but many trajectories simultaneously (with different phase factors).

In order to implement the PIMC numerically, a finite number of Trotter replicas \( P \) is required. This leads to a new approximation, whose error is proportional to the square of the Trotter break-up time, \( O(P^2) \) [20].

To implement a quantum annealing method (PIMC-QA) within this approach, we externally control, during the PIMC dynamics, the value of the transverse field \( \Gamma \) – leaving the temperature untouched – in the same way as one externally controls \( T \) in CA [2]. This approach does not lead, therefore, to the simulation of a true quantum mechanical dynamics, of the type implied by equation (10), but only to a MC annealing dynamics.

Here results on the two-dimensional (2D) 80 \( \times \) 80 lattice case are reported as they appear in [16, 19], where the choice of the random couplings \( J_{\Gamma} \) drawn from a flat distribution in the interval \((-2, 2)\), is also discussed in
4. Beyond physics towards quantum heuristics

PIMC-QA proved to be a fairly good model of the quantum relaxation of a Ising spin glass, being able to achieve configurations with smaller energies than in the case of CA and working more quickly. As in the case of its classical counterpart, nothing prevents us from applying QA to the optimization of arbitrary Hamiltonians, even those not directly inspired by physical intuition. An important advantage (but often also one of the main drawbacks) of the use of statistical spin models is that they are sufficiently general and abstract to be applied to capture the essential features of problems from many disparate fields, ranging from condensed matter physics to biochemistry, genomics, ecology, socioeconomics, etc.

We will focus in this section on some application of QA to combinatorial optimization, basing our discussion on the celebrated example of the Travelling Salesman Problem (TSP). Let suppose that a salesman has to visit \( N \) cities to sell and advertise his products. Since this kind of work can become extremely boring over the years, the salesman is interested in visiting each site once and only once, but, on the other hand, fuel is expensive and he wants also to minimize the total length of his tour. TSP can be described as the optimization task of finding the shortest route visiting exactly once all the \( N \) cities – whose coordinates and corresponding distance matrix have been pre-assigned – and returning to the original city. Obviously the given distances between each city are fixed for a given instance of the problem – the salesman cannot displace Aberdeen closer to Plymouth simply because he doesn’t want to spend too much money travelling.

When the number of cities becomes large, this simple-to-state problem can become extremely difficult to solve. The total number of possible tours (corresponding to the total number of permutations) grows extremely fast and an exhaustive enumeration of all the paths in order to find the shortest one is very quickly ruled out. To date, an algorithm able to determine the shortest path for every random instance of the problem in a time which is bounded by a polynomial in \( N \) is not known. On the other hand, it is quite easy to evaluate the length of a path and, once a good guess is provided, the verification that the proposed path is shorter than every other path previously tested can be done in polynomial time [23].

Theoretical computer science has provided a classification of combinatorial optimization problems in classes of complexity. Problems for which every instance can be solved in polynomial time are said to belong to the class \( \mathbf{P} \) (polynomial-time solvable by a deterministic machine), while problems for which polynomial resolvability has not yet been achieved, but for which verification of optimality can be done in polynomial time (once a guess has been provided), are said to belong to the class \( \mathbf{NP} \) (polynomially-solvable by a non-deterministic machine and \( \text{not}\) ‘non-polynomial’, as frequently stated). For a precise introduction to complexity theory, have a look at [23].

The attribution of a problem to a given class depends on the present state of the art in algorithm research. With today’s methods, factoring an integer takes typically a time which is exponential in the size of the integer (i.e. in the

Figure 7. The average best residual energy obtained by PIMC-QA and CA versus the total annealing time \( \tau \). This picture refers to the 80 \( \times \) 80 instance of the random Ising model studied in [16, 19]. CA and PIMC-QA time units are sweeps of the entire lattice.
number of its digits). Since it is very easy to verify whether or not a set of primes is a valid factoring of our integer, factorization belongs today to the class $\mathsf{NP}$. The same used to be considered true for the simpler task of deciding primality, i.e. of deciding if an integer is a prime or not. But in August 2002, Agrawal and his students were able to provide an astonishingly simple polynomial algorithm for this task [24]. Factorization itself is still in $\mathsf{NP}$.

Finding an efficient way of factoring large numbers have also important technological and economic implications since it would break many widespread encryption standards. The search for this is still going on and again QA and QC (see the Introduction) pursue similar goals in this respect. Indeed, the discovery in 1994 by Peter Shor [25] of his celebrated quantum algorithm for polynomial-time factoring of integers was a fundamental contribution to the establishment of QC as one of the most promising open fields in contemporary research in theoretical physics.

After the successful application of QA to the Ising spin glass, our group decided to try the optimization of a really hard instance of the TSP [17]. First of all, we need to understand how TSP can be translated in terms of Ising spins. We have to choose a representation for the classical potential energy $H_{\text{pot}}$ of a given configuration (i.e. the cost of a tour), and – most crucially – a suitable source of quantum fluctuations, $H_{\text{kin}}$. TSP can be mapped to a highly constrained spin-glass-like model in which each configuration of the system (a valid tour) is associated to a $N \times N \{0, 1\}$-matrix, $T$. For every ordered sequence of cities, $T_{i,j} = 1$ if the tour visits city $i$ immediately after city $j$, and $T_{i,j} = 0$ otherwise. For numerical experiments, a symmetric TSP instance in which the distances matrix is such that $d_{i,j} = d_{j,i}$ (the cost of moving from city $i$ to city $j$ is the same as that of moving from $j$ to $i$) has been selected. The directed tour represented by a $T$, and the reversed tour, represented by the transposed matrix $T^t$, then have exactly the same cost, and so it is possible to use the symmetric matrix, $U = T + T^t$, as representative of an undirected tour. The length of a tour can then be written as:

$$H_{\text{pot}}(U) = \sum_{\langle i,j \rangle} d_{i,j} \hat{U}_{i,j},$$

where $\langle i, j \rangle$ stands for counting each link only once. The advantage of working with the symmetrized matrix is now evident, since the resulting potential is a good Hermitian operator.

A suitable kinetic Hamiltonian term, $H_{\text{kin}}$, should now be chosen in order to induce fluctuations generating all the important elementary ‘moves’ of the problem. Deciding which configurations are direct ‘neighbours’ of a given configuration is once again a crucial step, because it determines the problem’s effective landscape. Energy barriers are related to the set of the allowed moves, and changing the set of possible moves will change the barriers seen by a local search method. Mathematical intuition or, more often, chance contribute to a choice of the ‘good’ set of proposed moves. A very common and high-performing choice for the TSP is the so-called 2-opt move, which consists in eliminating two links in the current tour, $(c_1 \rightarrow c_2)$ and $(c_1 \rightarrow 0)$, and rebuilding a new tour in which the connections are exchanged, $(c_1 \rightarrow c_0)$ and $(c_2 \rightarrow c_0)$ (figure 8 shows how this move can sometimes shorten

![TSP instance with 8 cities and corresponding matrices](image)

Figure 8. Left: Representation of an 8-city tour, with the corresponding matrix $T_{\text{in}}$. The final tour obtained when a 2-opt move is performed (right), with a whole section reversed (dotted line), is shorter than the original tour (left). The matrix $T_{\text{fin}}$ is shown, the circles indicating the entries that have been switched (0→1) by the 2-opt move. (Adapted from [17].)
non-optimal portions of a tour). We will not enter in the
details of the derivation of a quantum term generating
2-opt-move-like fluctuations. It is enough to know that
the trick consists in translating into spin language the
Hamiltonian equation (26), by associating Ising variables
\( S_{ij} \) to each entry of the symmetrized
tour matrix. At this point the kinetic term, once again
proportional to a transverse field, will be written as a
combination of raising and lowering operators (this is left
as an exercise to the curious reader, who can also find some
hints by consulting [17]).

We will have instead a brief look at the results of
the optimization of a hard instance with \( N = 1002 \) cities
extracted from a standard repository of benchmarks (the
‘TSPLIB’, see http://www.iwr.uni-heidelberg.de/groups/
comopt/software/TSPLIB95), whose optimal tour length
\( L_{\text{opt}} \) is known exactly. The comparison with CA, shown in
figure 9, shows results similar to those in the Ising spin glass
case: QA performs better and it is faster, even when the cost
of the simulation of the quantum computation on a ‘classical
computer’ is taken into account.

Unfortunately, it is impossible to conclude that QA is a
universally efficient tool for optimization, looking just at two
positive examples. Exactly as in the case of the simple double
well, counter-intuitive behaviour can be observed for hard
problem optimizations. For instance, we have found some
combinatorial optimization problems (notably hard K-SAT
instances, [26]) in which QA performs dramatically worse
than Classical Annealing. To date, no general theory has
been yet derived. Still, we have accumulated enough hints to
motivate further explorations of the field of quantum and
semi-classical approaches to combinatorial optimization.

5. Conclusions and perspectives

In this paper we introduced different concepts. First of all,
we pointed out that there are many physical systems such
that the determination of their ground state properties is
highly non-trivial, e.g. spin glasses. Since condensed matter
physics is a source of complex landscapes, techniques
borrowed from classical and quantum statistical mechanics
can be used equally well for tackling hard optimization
tasks in general. This has been notably successful in the
case of simulated annealing. Moreover, recent experimental
results have suggested the possibility of exploiting also
quantum relaxation pathways.

By studying the optimization of the ‘simplest among the
complex problems’, the double-well potential, with both
classical and quantum annealing, we have verified that the
residual energy found in the quantum case is asymptotically
smaller. We have also highlighted that the use of an
imaginary-time quantum dynamics allows for faster opti-
mization since excited states are exponentially suppressed.

Surprisingly, the quantum relaxation, efficient if it is
slow enough, may completely fail to find the true ground
state if the annealing is too fast, even for potentials as
simple as the double-well. Such a phenomenon – absent in
the classical case – can be interpreted as a competition
between the quantum kinetic and potential energy terms of
the Hamiltonian. Its signature is a Landau – Zener avoided
crossing between the ground state and the first excited
state in the instantaneous spectrum. The role of this feature
has been highlighted for the simple case of the double-
well potential, where it is possible to trace thoroughly
the influence of the few geometrical parameters of the

Figure 9. Average residual excess length found after CA and QA for a total time \( \tau \) (in MC steps), for the \( N = 1002 \) instance
pr1002 of the TSPLIB. QA is once again faster than CA. (Adapted from [17].)
landscape (barrier height, minimum energies, curvatures, etc.) on the annealing dynamics.

The introduction of Monte Carlo (MC) techniques, both in classical and quantum cases, has allowed us to attack larger problems by using annealing, and made the technique competitive for the title of best-performing optimization algorithm in some cases. Notably, our group has verified that path-integral MC (PIMC) Quantum Annealing is more efficient than classical annealing in optimizing Ising spin glasses (a feature reminiscent of Aeppli’s experiments) and in instances of an ‘unphysical’ combinatorial optimization problems, namely the Travelling Salesman Problem.

Many roads still need to be explored, and some theoretical aspect must be better understood. First of all, although PIMC quantum annealing proved to be surprisingly efficient in optimizing some landscapes, when it is applied to other problems (e.g. the random K-SAT) it behaves dramatically worse than its classical counterpart. This effect – as counterintuitive as the failure of quantum annealing in finding the right minimum of a simple double-well potential (near to a Landau–Zener crossing) – calls for a predictive theory allowing to determine whether or not a certain landscape can be ‘quantum friendly’ for optimization.

A possible negative aspect of the PIMC approach to simulated quantum annealing is the need for a finite temperature. Other quantum MC schemes, like the Green’s function MC, work directly at zero temperature and, in this sense, they can catch more genuinely quantum effects. Research is ongoing for verifying the effectiveness of such alternative simulation methods.

In the studies described here, we grasped some relevant feature of a quite general family of quantum annealing algorithms. Nonetheless, the exciting field of quantum optimization is still hiding many secrets and surprises for the researchers to come.

Acknowledgments

This project was sponsored by MIUR through FIRB RBAU017SR004, FIRB RBAU01LX5H, COFIN2003 and PRIN2004, and by INFM (‘Iniziativa trasversale calcolo parallelo’). We especially thank Giuseppe Santoro and Erio Tosatti, who introduced us to quantum annealing.

References


Demian Battaglia obtained his MSc in Theoretical Physics in 2002 at the University of Torino (Italy) with a thesis on dynamical systems over discrete space-times. As a graduate student in the International School of Advanced Studies (SISSA) of Trieste (Italy), he worked with Professor Riccardo Zecchina at ICTP (Trieste) and Professor Giuseppe Santoro at SISSA on applications of classical and quantum statistical mechanics to combinatorial optimization and information theory. After his PhD, he moved (2005) to the University ‘René Descartes’ in Paris becoming more and more interested in the secrets hidden by the neural code.

Lorenzo Stella studied Theoretical Physics at the University of Milan (Italy) where he obtained his MSc in 2001 with a thesis on mesoscopic systems. After joining the International School of Advanced Studies (SISSA) of Trieste (Italy), he moved his interests to the fascinating world of quantum annealing. In 2005, he obtained his PhD at SISSA, with a thesis about the comparison between classical and quantum annealing. He is currently a research fellow at University College, London.