Quantum Monte Carlo annealing with multi-spin dynamics

Guglielmo Mazzola\textsuperscript{1} and Matthias Troyer\textsuperscript{1,2}

\textsuperscript{1} Theoretische Physik, ETH Zurich, 8093 Zurich, Switzerland
\textsuperscript{2} Quantum Architectures and Computation Group, Microsoft Research, Redmond, WA 98052, United States of America
E-mail: gmazzola@phys.ethz.ch

Received 3 February 2017
Accepted for publication 17 April 2017
Published 12 May 2017

Abstract. We introduce a novel simulated quantum annealing (SQA) algorithm which employs a multispin quantum fluctuation operator. At variance with the usual transverse field, short-range two-spin flip interactions are included in the driver Hamiltonian. A Quantum Monte Carlo algorithm, capable of efficiently simulating large disordered systems, is described and tested. A first application to SQA, on a random square lattice Ising spin glass, reveals that the multispin driver Hamiltonian improves upon the usual transverse field. This work paves the way for more systematic investigations using multi-spin quantum fluctuations on a broader range of problems.

Keywords: quantum Monte Carlo simulations, spin glasses
1. Introduction

Discrete combinatorial optimization problems can be encoded into minimizing the energy of classical Ising-type Hamiltonians [1] and it has been proposed that adding quantum mechanics could help in finding the ground state of these frustrated Ising system faster than any classical technique. One can define a suitable time-dependent quantum Hamiltonian that connects an initial Hamiltonian, whose ground state is easy to prepare, to the final classical Ising Hamiltonian. At sufficient low temperatures a quantum system undergoing such adiabatic relaxation reaches its ground state and solves the optimization problem. This technique is called quantum annealing (QA) [2–5].

Early numerical studies predicted QA [3, 4, 6] to be a competitive computational resource for solving spin glass problems, compared to its closely related classical counterpart, the simulated annealing (SA) algorithm [7]. However, so far no quantum speedup has been observed in experiments [8, 9]. The essential difference between these two heuristic methods relies on the type of fluctuations which drive the system away from the multiple local minima, occurring in rugged energy landscapes.

Quantum fluctuations are expected to give an advantage to quantum annealing in particular when the free energy landscape displays tall but narrow barriers. These are easier to tunnel through quantum-mechanically, compared to climbing over them by means of thermally activated escape events.
In QA the system closely follows the ground state of a time-dependent Hamiltonian $H(s)$ which at $s = 0$ is dominated by a pure quantum fluctuation part $H_Q$ whereas the final Hamiltonian encodes only the cost function $H_P$ of the combinatorial optimization problem. The Hamiltonian as a function of the rescaled time $s$ may read, in the case of simple linear annealing schedules,

$$H(s) = H_P + (1 - s) H_Q,$$

with $s \in [0, 1]$. Since random ensembles of hard problems are closely connected to spin glass models, we choose, for our tests, the problem Hamiltonian as an Ising spin glass

$$H_P = \sum_{i,j} J_{ij} \sigma_i^z \sigma_j^z,$$

defined on a graph of arbitrary connectivity. We can assume the sites $i, j$ to be coupled in the graph. $\sigma_i^z$ is a Pauli matrix, acting on spin $i$, and $J_{ij}$ is the coupling between spins $i$ and $j$, which are randomly uniformly distributed in a range $[-1, 1]$. Most QA studies so far employed a single type of the quantum term $H_Q$, a transverse-field (TF)

$$H_Q^{TF} = -\Gamma \sum_i \sigma_i^x,$$

where the $\sigma_i^x$ operator acts locally on the spin index $i$ inducing quantum fluctuations. This is simplest to implement both in physical devices, such as the D-Wave devices [8, 10–12] and in simulated quantum annealing (SQA) by means of quantum Monte Carlo (QMC) methods [6, 13–15].

However, as pointed out by [16], it is possible to choose also different types of quantum fluctuations operators $H_Q$. Following [16] we define a two body fluctuation operator with ferromagnetic interactions (FI) as

$$H_Q^{FI} = -\Gamma \sum_i \sigma_i^x - \Lambda \sum_{i,j} \sigma_i^x \sigma_j^x,$$

where $i, j$ are neighbors of coupled spins, so that the interaction is short-range. In the following we always assume that the sums in equation (4) and equation (2) are equal. However, this is not a strict requirement for the algorithm we are going to present, as the number of interactions in equation (2) can be larger than the one in equation (4).

The adiabatic theorem states that the system follows the instantaneous ground state as long as the total annealing time $t_{\text{final}} \gg 1/\Delta^2$ where $\Delta$ is the minimum energy gap from the ground state, which is encountered along the annealing run of equation (1). It has been conjectured and shown for simple models, that employing quantum fluctuations beyond the TF term could be beneficial to avoid these small gap events [16, 17], which are the bottlenecks of QA [6, 18–20]. Another argument in support of this possibility is that, in equation (4), we are effectively adding a new schedule parameter $\Lambda$ which can be also optimized to increase the QA performance.

Moreover, it has been found that quantum annealing with transverse field fails to identify all degenerate ground-state configurations, preventing a fair sampling of
equally probable states [21, 22]. Multi-spin quantum fluctuations could alleviate or possibly remove this issue.

In this paper we devise a QMC algorithm to perform SQA with this two-spin transverse ferromagnetic interaction. We note that the minus sign in front of $\Lambda$ makes this Hamiltonian stoquastic for $\Lambda > 0$. This means that it is sign-problem free and can thus be simulated by QMC methods. On the other hand, the existence of a sign problem precludes the possibility to simulate the two spin transverse antiferromagnetic ($\Lambda < 0$) interaction [23–25] within the same approach unless the graph of couplings is bipartite.

The paper is organized as follows, in section 2 we describe the QMC algorithm, which is tested against exact results for small systems in section 3. In section 4 we report a first illustrative application to the relevant problem of Ising glass SQA. We discuss the results as well as possible future developments in section 6.

2. Quantum Monte Carlo methods

2.1. Overview

Quantum Monte Carlo (QMC) methods are the only classical approaches for simulating quantum annealing on systems as large as the ones realized experimentally, which consider more than $N = 1000$ qubits. Other numerical methods such as unitary evolution [3, 4] scale exponentially with $N$ and are therefore limited to much smaller numbers of spins ($N \approx 20$).

The most widely used QMC technique in the context of QA is Path Integral Monte Carlo (PIMC). PIMC relies on the path integral formalism of quantum mechanics and samples the density matrix corresponding to the quantum Hamiltonian $H$ by means of a classical Hamiltonian $H_{cl}$ on an extended system having an additional dimension, the imaginary time direction [6, 13, 14]. The original transverse field quantum spin system is then mapped into a classical one, which can be simulated by standard Metropolis Monte Carlo.

It has recently been shown [14, 26, 27] that for tunneling through a barrier the PIMC efficiency scales as a function of the system size as a physical QA would do. This connection exists due to the fact that the tunneling rate in QMC and in the exact real-time quantum evolution scales in the same way, to leading order, as a function of the relevant parameters.

This connection has two implications. The first is that a QMC annealing simulation is relevant for investigating the behaviour of a real quantum annealer, as long as it is stoquastic and the QMC simulations thus do not suffer from a sign problem. The second is that it makes QMC-SQA a competitive tool compared to QA in real devices, by precluding a scaling advantage of stoquastic QA over SQA (see also [15]).

Starting from this mapping, other quantum inspired algorithms can be developed such as PIGS, where open boundary conditions in imaginary time are used [14] for further acceleration. In this case QMC represents only a classical optimization algorithm and the physical meaning of the simulation is lost. Similarly, it has been recently
shown that performing PIMC away from the converged physical limit, provides a more efficient algorithm compared to the continuous time one [13]. This can be also considered a quantum inspired algorithm. For this reason we focus here on implementing a discrete-time PIMC algorithm rather than a continuous time PIMC.

2.2. Local versus cluster updates

The simplest Metropolis algorithm performs local updates: one generates trial configurations by simply flipping one spin at a time. While this algorithm is ergodic, it can be very inefficient. For non-frustrated systems considerable improvements are obtained by cluster algorithms [28, 29]. They allow simultaneous flips of clusters of spins and are especially advantageous when large scale fluctuations are important, usually around phase transitions, where local update algorithms slow down.

Standard cluster algorithms [28, 29] allow the clusters to grow without restrictions. While efficient for non-frustrated systems this approach breaks down for frustrated spin systems, as spin glasses. Due to the frustration the clusters grow to fill a very large fraction of the lattice, if not the entire lattice. The clusters are then larger than the physical domains that should be flipped and the algorithm becomes inefficient [30–33]. Essentially the whole system freezes out and one flips almost all spins.

Therefore, in the context of SQA of spin glasses with transverse field a hybrid approach is usually employed, in which restricted clusters are built [13, 34]. Here, local clusters can grow on a single site only in the imaginary time dimension, where there is no frustration. In this paper we employ the same strategy with an important modification. Due to the non local nature of the two-spin interaction we cannot restrict ourselves to strictly local clusters but we allow a small extent in the spatial direction.

2.3. Cluster QMC for pure two-spin couplings

For the sake of simplicity we first start with the case $\Gamma = 0$, i.e. considering only the two-spin flip term in equation (4). The Hamiltonian $H = H_P + (1-s)H_Q$, at any time $s$ is as a special case of the anisotropic quantum XYZ model [35]. In this section we present a modified version of the loop algorithm used to simulate the thermodynamics of this model. Readers unfamiliar with the loop algorithm are referred to the review [36] for details.

Moreover we consider, for the time being, a ferromagnetic Hamiltonian with homogeneous couplings $J > 0$ in this discussion, defined on an arbitrary graph of $N$ sites connected by $B$ bonds (the disordered case $J_{ij}$ will be considered at the end):

$$H = -\left( J \sum_{i,j} \sigma_i^z \sigma_j^z + \Lambda \sum_{i,j} \sigma_i^x \sigma_j^x \right).$$

The generalization of the algorithm to spatially varying couplings, which is relevant for QA, is then straightforward.

We first identify the non-commuting pieces of this Hamiltonian to perform a Trotter breakup. This procedure is less straightforward compared to the transverse field case where the Hamiltonian splitting was a trivial splitting $H = H_P + H_Q^{TF}$. Here we need to split the Hamiltonian $H = \sum_{k=1}^{K} h_k$ into a sum of commuting bond Hamiltonians $h_k$. 

https://doi.org/10.1088/1742-5468/aa6de1
where \( i_b, j_b \) are two sites connected by the bond \( b \) and \( H_b = -(J \sigma^+_{i_b} \sigma^-_{j_b} + \Lambda \sigma^z_{i_b} \sigma^z_{j_b}) \). Each set \( \{ b \}_k \) is defined such that its elements don’t share any site, i.e. no site \( i \) appears twice (see figure 1(a)). The number \( K \) of non-commuting terms \( h_k \) depends on the graph’s connectivity. For example, in a linear chain we have \( K = 2 \) and the standard checkerboard decomposition [36], whereas for a square lattice \( K = 4 \).

The partition function then reads

\[
Z = \text{tr} e^{-\beta H} = \lim_{M \to \infty} \text{tr} \left( \prod_{k=1}^{K} e^{-\frac{\beta}{M} h_k} \right)^M.
\]

Inserting complete sets of \( \sigma^z \) eigenstates, and following [36] we write

\[
Z = \sum_{\{ s^z_i \}} W(\{ s^z_i \}) = \sum_{\{ s^z_i \}} \prod_p W_p(\{ s_p \}),
\]

where the outer summation is carried over all possible spin configuration on the extended lattice. The index \( i \) runs over the original graph sites \( i = 1, \cdots, N \), whereas \( l \) label the imaginary time coordinate \( l = 1, \cdots, MK \). The index \( p \) extends over all the shaded plaquettes of the extended lattice (see [36]). We define a plaquette as the 4-spin configuration \( p = p(b, l) = \{(i_b, l), (j_b, l), (i_b, l + 1), (j_b, l + 1)\} \). The statistical weight of each plaquette \( p = p(b, l) \) is

\[
W_p(s_p) = \langle s(i_b, l)s(j_b, l) | e^{-\Delta H} | s(i_b, l + 1)s(j_b, l + 1) \rangle,
\]

where \( \Delta = \beta/M \). \( W_p(s_p) \) is different from 1 only when the piece of Hamiltonian \( h_k \) acts on the bond \( b \) at the correct time \( l = [(\ell - 1)K + k, \ell K] \), with \( \ell = 1, \cdots, M \). This condition defines the shaded plaquettes. For each bond \( b \) we have exactly \( M \) shaded plaquettes, one for each Trotter time step \( \ell \) (see figure 1(b)). The non-trivial Hamiltonian evolution in imaginary time occurs only on these special 4-spins configurations. For the simple case of the spin chain, the shaded plaquettes construction resembles a checkerboard lattice. There are only eight non-vanishing matrix elements on the shaded plaquettes, which can be divided in four types:

\[
W(1) = \langle + + | e^{-\Delta H_b} | + + \rangle \approx 1 + \Delta J
\]

\[
W(2) = \langle + - | e^{-\Delta H_b} | + - \rangle \approx 1 - \Delta J
\]

\[
W(3) = \langle + - | e^{-\Delta H_b} | - + \rangle \approx \Delta \Lambda
\]

\[
W(4) = \langle + + | e^{-\Delta H_b} | - - \rangle \approx \Delta \Lambda,
\]

expanding the exponential for small \( \Delta \), and taking into account that, for example \( W(1) = \langle + + | e^{-\Delta H_b} | + + \rangle = \langle - - | e^{-\Delta H_b} | - - \rangle \) since the one body magnetic field operator \( \sigma_z \) is absent. Since \( H_b \) preserves the parity of the magnetization, only 8 out of the 16 possible states give a non-zero contribution. Notice that this property would not hold in the transverse field case.

https://doi.org/10.1088/1742-5468/aa6de1
The simplest Monte Carlo algorithm usually employs local updates, i.e. by proposing to flip one spin at a time. This is not a good choice for XX couplings as unconstrained single spin flips lead to forbidden plaquette configurations with zero weight. Our strategy is therefore to implement directly a cluster MC algorithm which automatically avoids sampling of forbidden configurations and minimizes the autocorrelation times. The key feature of the loop algorithm \cite{36} is that it allows nonlocal changes of spins configurations, by making only local stochastic decisions on the shaded plaquettes. We refer the reader to the original \cite{36} for an exhaustive and rigorous justification of the algorithm. Here we simply describe the outline of our optimized version for XX couplings, which, as it usually happens in most MC algorithms, can be divided in two steps: cluster formation and cluster flip.

2.3.1. Outline of the algorithm. We first restrict the growth of the cluster to a subset of bonds $B_m$ such that $\bigcup_m B_m = B$ contains all the $B$ bonds of the graph (see figure 2). For example, a possible choice could be these set of bonds which constitute all the possible smallest loops in the graph, as in section 4.2. This is necessary to keep the cluster size small in the $N \to \infty$ limit (see section 2.2).

Next we explicitly show how to combine to standard Loop algorithm \cite{36} with the restricted update scheme, therefore in the following we implicitly refer to the Loop algorithm terminology and we suggest that the reader should become familiar first with the standard version described in \cite{36}.

The algorithm proceeds as follows:

i. We randomly select the set $B_m$ of spins on which we want to build a cluster. Then we choose one bond $b_O \in B_m$ and a Trotter time-slice $\ell = 1, \cdots, M$. This

\[h_1 \]
\[h_2 \]
\[h_3 \]
\[i=1\]
\[i=2\]
\[i=3\]
\[i=4\]
\[l=1\]
\[l=2\]
\[l=3\]
\[l=4\]

Figure 1. (a) Splitting of non-commuting bond Hamiltonians $h_k$ for a simple, non-regular graph with four sites. In this specific case $K = 3$. (b) Extendend graph with shaded plaquettes construction along the imaginary time direction. The bond Hamiltonian $h_1$ acts at the time slice $l = 1 \mod K$, etc. The primitive cell repeats vertically each $K = 3$ slices. (c) Example of cluster building (see section 2.3.1). In this sketch we add 6 sites after the initial $(i_O, l_O)$, following the Loop Algorithm procedure detailed in \cite{36}. The cluster building will continue until we reach again the starting site $(i_O, l_O)$.
identifies the corresponding shaded plaquette \( p = p(b_O, l_O) \). Finally we randomly select one of the two sites, i.e. \((i_O, l_O)\), connected by the bond \( b_O \), at the imaginary time \( l_O \). This will be the first site added to the cluster.

ii. We follow the Loop Algorithm rules to add spins to the cluster. In this case we add only spins which are connected by a bond in the set \( B_m \).

iii. We propose to flip this cluster with suitable probability \( p_{\text{flip}} \), i.e. we reverse each spin \( s \rightarrow -s \) belonging to the cluster. \( p_{\text{flip}} = \min[1, \exp(\delta E K/\Delta)] \) is given by the change in energy \( \delta E \) due to the cluster flip. Notice that only the bonds outside the set \( B_m \) will be taken into account in computing \( \delta E \) since the cluster building rules take care of detailed balance for all bonds within \( B_m \). In this way we evaluate the energy gain of the cluster flip restricted to \( B_m \), which is embedded in the graph. This is the main difference compared to [36].

Notice also that if \( B_m = B \) already contains all the possible edges of the graph, then step iii. simplifies to flipping the cluster with probability 1, and we recover the single cluster formulation of the Loop Algorithm. We label this type of move as unrestricted or global update. If instead the cluster is restricted to be on a localized set of bonds, we are performing a semi-local update.

2.3.2. Breakup probabilities. The essential ingredients of the Loop Algorithm are the so-called ‘breakup weights’. In short, this set of weights \( w_{ij} \) (\( i, j = 1, 2, 3, 4 \)) determines the shape and the size of the clusters. For example, the weight \( w_{ij} \) sets the probability that a shaded plaquette of type \( i \) is changed into a plaquette type \( j \), after the accepted MC update (see equation (10)). In particular, this probability is given by \( p_{i \rightarrow j} = w_{ij}/W(i) \). For example, if we connect and flip the spins at the bottom of a type-1 plaquette, we obtain a type-4 plaquette. The flip of two spins inside a shaded plaquette always produce a different type of configuration, which is in turn a valid plaquette configuration, by construction. This is the main idea of the Loop algorithm. One needs to define every possible \( i \rightarrow j \) transition weight. Some of these weights can also be zero, this means that some particular transitions are not-allowed.

We refer the interested reader to the original Loop Algorithm [36] for details, derivations and discussions. Here we simply provide our choice for the breakup weights for future reproducibility.

While the plaquette weights \( W(i) \), with \( i = 1, 2, 3, 4 \), are fixed by the Hamiltonian (see equation (10)), there is freedom in choosing the breakup weights \( w_{ij} \), provided that the following constraints are met

\[
\sum_j w_{ij} = W(i),
\]

\[
w_{ij} > 0.
\]
to minimize this weight. The optimal weight’s set varies as a function of the annealing schedule. Let us assume that $\Delta$ is small so that $W(1), W(2) > W(3), W(4) > 0$.

Let us consider first the case $\Lambda > |J|$, i.e. at the beginning of the annealing. Under this condition it is possible to set all the freezings to zero. The non-zero weights are:

$$w_{12} = 1 - \Delta J$$
$$w_{13} = w_{14} = \Delta J$$
$$w_{34} = \Delta (\Lambda - J),$$

with $w_{ij} = w_{ji}$, $i, j = 1, 2, 3, 4$. In the later stage of the annealing we have $\Lambda < |J|$ instead, so we find

$$w_{12} = 1 - \Delta J$$
$$w_{13} = w_{14} = \Delta \Lambda$$
$$w_{11} = 2\Delta (J - \Lambda).$$

In this case is not possible to have $w_{ii} = 0$, but this is the optimal choice as pointed out also in [37]. Notice that these two sets always satisfy equation (14) and work both for ferromagnetic and anti-ferromagnetic couplings.

Figure 2. Different types of cluster update, unrestricted in panel (a) and restricted, (b). Colors (red, orange and pink) depict loops that are glued together after that the cluster formation procedure is terminated, if freezing breakups occur (see [36]). In the restricted case, the cluster is confined into a 4 bond loop in the physical lattice (green square), other choices are obviously possible. While in the unrestricted case the cluster is always flipped, in the restricted case the cluster is flipped according to a suitable Boltzmann probability (see text).
2.4. Adding the transverse field

To obtain an algorithm for the full Hamiltonian

\[ H = \sum_{i,j} J_{ij} \sigma_i^z \sigma_j^z - \Gamma \sum_i \sigma_i^x - \Lambda \sum_{i,j} \sigma_i^x \sigma_j^x, \tag{22} \]

we need to include the transverse field term \( \sigma_i^x \) to the Loop Algorithm (\( \Gamma = 0 \)) described in section 2.3. This extension is hinted in [36] and is not trivial as the transverse field term \( \sigma_i^x \) breaks the parity of the magnetization within each interaction plaquettes. Instead of enlarging the total number of allowed plaquette configurations, taking into account all these possible states, we treat this operator stochastically by adding additional single-site breakups. The transverse field operator can end a loop cluster at any spacetime point \((i, l)\), whereever \( \sigma_i^x \) acts. In our algorithm we add the possibility to stop the cluster when entering into a shaded plaquette at site \( i \), with probability given by \( p_x = \sinh (\Delta \Gamma / K_i) \), where \( K_i \) denotes the number of physical neighbours of the site \( i \) in the graph. This is the generalization to arbitrary graphs of the procedure sketched in [36] for a linear chain. In the following we give details on the actual implementation of this idea.

i. Suppose we start the cluster from position \((i_0, l_0)\).

ii. When we jump into a shaded plaquette, say at position \((i, l)\), if the plaquette type is not of type 3 or 4, we stop the cluster with probability \( p_x \).

iii. Each time we stop we keep track of this position by inserting a \( \sigma_i^x \) operator label. We put this label on the bond above (below) site \( i \), if the direction was upward (downward) in imaginary time direction. Otherwise we continue to build the cluster following the procedure described in section 2.3.1.

iv. If we stop the cluster then we restart from \((i_0, l_0)\) and proceed in the opposite direction until we stop again, notice that now we can stop either by inserting a new \( \sigma_i^x \) operator or by touching an existing one already in place.

v. The cluster flipping decision remains unchanged.

vi. Finally, remove the \( \sigma_i^x \) operator labels between segments of equal orientation. This completes one update.

We notice that this kind of update respect the detail balance conditions as discussed in [36]. The last issue concerns the existence of plaquettes having an odd number of spin up (down), which fall outside the breakup selection rule of the XX loop algorithm. These plaquettes certainly occur if \( \Gamma > 0 \). Since in our approximation we can encounter the \( \sigma_i^x \) operator only going vertically along the imaginary time direction, the decision rule can be adapted from the standard transverse field cluster algorithm [34] in this case: suppose we are at site \((i, l)\) and we are proceeding upward, if there is not a \( \sigma_i^x \)

\footnote{For simplicity we don’t break clusters on these types of plaquettes. This approximation is justified in the \( M \to \infty \) limit.}
operator already in place above \((i, l)\), then we check whether the spin at \((i, l + 1)\) is parallel to \((i, l)\), or not. In the latter case we stop the cluster.

Finally we notice that, in the \(\Lambda = 0\) case, this algorithm does not reduce to the common TF algorithm for disordered systems \([34]\), in which clusters are built only along the imaginary time direction, if a non-empty bond set \(B_m\) is considered. Indeed, the cluster can still span the entire bond region \(B_m\), due to the occurrence of the freezing plaquettes, which make the cluster non-local. We use therefore the same cluster algorithm, within the same choice of \(B_m\)’s, to compare FI and TF Hamiltonians, also considering the limiting \(\Gamma = 0\) case.

3. Phase diagram of the XX model with a transverse field

We test the accuracy of the algorithm against exact diagonalization (ED) results \([38]\), for a small uniform ferromagnetic chain and square lattices with spacial periodic boundary conditions, having Hamiltonian given by equation (22), i.e. without bond disorder. In figure 3 we plot the \(\langle \sigma_i^z \sigma_j^z \rangle\) correlation function, for nearest neighbours, as a function of the inverse temperature \(\beta\), for several choices of the parameters \(\Lambda\) and \(\Gamma\). The agreement of QMC with ED is satisfactory in the full temperature range.

Next, in figure 4, we focus on the low-temperature regime relevant for SQA and benchmark simulations obtained with different combinations of \(\Lambda\) and \(\Gamma\) values, which can be realized along a generic SQA run.

Finally in figure 5, we show, as an example, the low-temperature phase diagram of the Hamiltonian in equation (22), with \(J_{ij} = 1\), on a 2D square lattice. Generalizing the breakup weights in section (2.3.2) for an arbitrary XYZ Hamiltonian, it would be possible to study the phase diagram of sign-problem free XYZ models with transverse field on arbitrary lattice.

4. Simulated quantum annealing

4.1. Annealing schedule

We now report a first application of the algorithm to SQA, to assess the annealing sensitivity to the \(H_Q^{\text{FI}}\) driver Hamiltonian. We consider a spin glass problem Hamiltonian, defined on a \(10 \times 10\) square lattice, with periodic boundary conditions, and uniformly randomly distributed couplings in the range \([-1, 1]\). Following previous SQA studies \([13, 6]\), we plot the residual energy \(E_{\text{res}}(t_{\text{final}}) = E(t_{\text{final}}) - E_0\) as a function of different annealing times \(t_{\text{final}}\), where \(E_0\) is the exact solution \([39]\) and \(E(t_{\text{final}})\) is the ground state of the Hamiltonian found at the end of the annealing.

In order to compare the performance of different annealing strategies in the \(\Gamma - \Lambda\) parameters space, we need to rigorously define the computational effort for a QMC algorithm having unrestricted and semi-local updates. Different annealing schedules lead to different average cluster sizes, and, the larger is the cluster built, the heavier is the
Quantum Monte Carlo annealing with multi-spin dynamics

Figure 3. Nearest neighbours $\langle \sigma_i^z \sigma_j^z \rangle$ correlation function as a function of the inverse temperature $\beta$, for several choices of the parameters $\Lambda$ and $\Gamma$. The system is a ferromagnetic 8-site chain. Points label QMC results, while continuos lines represent ED results. For each temperature we use an appropriate number of Trotter slices $M$ to ensure convergence to the continuos imaginary time limit. QMC error bars are smaller than the symbol’s size.

Figure 4. Ground state nearest neighbours $\langle \sigma_i^z \sigma_j^z \rangle$ correlation function for several combinations of $\Lambda$ and $\Gamma$. We parametrize these values as $\Lambda = \lambda Z$ and $\Gamma = (1 - \lambda) Z$ and we plot $\langle \sigma_i^z \sigma_j^z \rangle$ as a function of $\lambda$, for $Z = 0.5, 1, 2$. Points label QMC results obtained with suitably converged value of $\beta J = 5 \div 20$ in order to approach the ground state limit, while continuos lines represent ED results. The left panel refers to a 8-sites ferromagnetic chain, while the right panel to a 16-sites square lattice. QMC error bars are smaller than the symbol’s size.
computational cost of each update. This cost is proportional to the cluster size $\bar{n}$ and to the number of bonds one has to check for evaluate $\delta E$ in the acceptance/rejection step. We notice that, in the algorithm, the computationally expensive operations are made on the shaded plaquettes. Therefore, the total effort is still proportional to the number of sites $N$ (or generically to the bond number $B$) and the number of Trotter slices $M$, although the total number of slices along the imaginary time direction is $M \times K$. Here we thus define the computational cost associated with each annealing run as

$$C(t_{\text{final}}) = t_{\text{final}} \bar{n}, \quad (23)$$

where $t_{\text{final}}$ is the total number of Monte Carlo updates.

We notice that the QMC algorithm is efficient, for low $\Lambda$ and in the thermodynamic limit, only if we restrict the cluster growth, at each update, to a small bond subset with $\#B_m \ll B$. In this way $\bar{n} \ll N \times M K$

Usually, with TF-SQA, a linear schedule with starting field $\Gamma_0 = \Gamma(s = 0) > \max |J_{ij}|$ is employed [13], in our notations, this would correspond to decrease linearly the transverse field parameter as $\Gamma(s) = \Gamma_0 (1 - s)$ and set $\Lambda = 0, \forall s$. In the following we will compare to the simplest annealing path for SQA with ferromagnetic interactions (FI-SQA), given by

$$\Gamma(s) = \Gamma_0 (1 - s)$$

$$\Lambda(s) = \Lambda_0 (1 - s), \quad (24)$$

namely, decreasing linearly both the control parameters. We will also compare to the choice $\Lambda(s) = \Lambda_0 (1 - s)$ and $\Gamma = 0, \forall s$, i.e. employing only the two pure two-body operator. The optimization of the annealing path to non-trivial $\Lambda(s), \Gamma(s)$ time-dependences is left for future studies.

Figure 5. Low temperature ($\beta = 20$) phase diagram of the model (22), with $J_{ij} = 1$, on a $10 \times 10$ square lattice. We plot the $\langle \sigma^z_i \sigma^z_j \rangle$ correlation, with $i, j$ nearest neighbours, as a function of $\Lambda$ and $\Gamma$ parameters, both in the range $[0, 1.9]$. The phase boundary between the ordered (below) and disordered (above) phases roughly lies along the $\Lambda + \Gamma/2 = 1$ line.
4.2. 2D spin glass results

In figure 6 we show the residual energy as function of the annealing time \( t_{\text{final}} \), renormalized in order to express the total theoretical computational effort of each run as in equation (23). Here, the QMC algorithm employs restricted cluster updates. Dashed lines with empty symbols indicate SQA with standard TF Hamiltonian (with \( \Gamma_0 = 2 \)), while continuous ones with solid symbols refer to FI-SQA, with annealing schedule given by equation (24), with \( \Gamma_0 = 1 \) and \( \Lambda_0 = 1 \). Each color (orange, blue, and red) labels different SQA setup, having different \( M \) and \( \beta \). We qualitatively checked that the relative annealing performance does not depend on the specific choices of these parameters, provided that we always start the annealing in the disordered phase. The FI Hamiltonian always performs better, for sufficiently long \( t_{\text{final}} \), than the TF, within the same \( M, \beta \) setup. Finally, we perform SQA without TF component (purple dot-dashed line). This choice gives poor performances.

Figure 6. Residual energy \( E(t_{\text{final}}) \) as a function of total annealing time \( t_{\text{final}} \), renormalized in order to express the total theoretical computational effort of each run as in equation (23). Here, the QMC algorithm employs restricted cluster updates. Dashed lines with empty symbols indicate SQA with standard TF Hamiltonian (with \( \Gamma_0 = 2 \)), while continuous ones with solid symbols refer to FI-SQA, with annealing schedule given by equation (24), with \( \Gamma_0 = 1 \) and \( \Lambda_0 = 1 \). Each color (orange, blue, and red) labels different SQA setup, having different \( M \) and \( \beta \). We qualitatively checked that the relative annealing performance does not depend on the specific choices of these parameters, provided that we always start the annealing in the disordered phase. The FI Hamiltonian always performs better, for sufficiently long \( t_{\text{final}} \), than the TF, within the same \( M, \beta \) setup. Finally, we perform SQA without TF component (purple dot-dashed line). This choice gives poor performances.
In this study we compare TF and FI-SQA as classical optimization algorithms. Therefore we renormalized the annealing time in such a way to fairly compare the theoretical computational effort of each run, as discussed above. We also define the residual energy as the minimum energy $E_{\text{min}}$ among all the possible $M$ Trotter slices.

We perform the simulations at different inverse temperatures $\beta = 20, 40, \text{and } 80$. We keep the Trotter time-step $\beta/M = 0.3125$, constant and close to convergence to the physical limit (see also figure 8).

In figure 6 we observe that, for fixed $M$ and $\beta$, TF and FI-SQA display different behaviour, as the two curves are not related by a trivial shift along the time axis. We notice that the FI driver always outperforms the standard TF, for sufficiently large annealing times, at each temperature, despite the larger complexity of the FI driver. Interestingly, the lower is temperature, the larger is the difference between the FI and TF residual energies.

Surprisingly, performing the annealing at zero transverse field, $\Gamma = 0$, results in a drastic decrease of the performance. These observations can be explained in the following way: i. For short annealing time, a large part of the residual energy can be recovered by simple one-flip moves, then the TF driver is more effective in eliminating the defects in the extended lattice. ii. The pure two body driver Hamiltonian $\sigma_i^x \sigma_j^x$ is way less efficient in this regards, as the clusters are usually bigger and always closed in the extended lattice. Therefore, the transverse-field operator is still an important ingredient for a QA driver Hamiltonian.

Finally, we also check the performance using the global update algorithm. In this case the cluster can freely traverse the extended lattice. We consider the same system and perform the same set of simulations and show the results in figure 7. We notice that the
trend is preserved, although the annealing profiles are slightly different. In particular, the pure two-body driver performs significantly worse than the mixed FI-SQA, having both $\Lambda$ and $\Gamma$ non-zero. However, FI-SQA still outperforms TF-SQA for sufficiently long annealing times, as in the previous semi-local algorithm. Interestingly, this global updates version of the algorithm displays a better efficiency compared to the previous approach. Indeed such system size can be considered still small and the global update scheme does not suffer much from a critical slowing down.

Note though the non-monotonicity as a function of the annealing time compared to the TF (see also [13]), which hints at inefficiencies of the global updates. This feature disappears when the average energy is used (see figure 8), instead of the lowest one among the Trotter slices.

5. Efficiency of quantum annealing

In this last section we perform simulations indicative of the relative efficiency between the FI and TF operators in a real QA device instead of as a classical optimization algorithm. QMC is expected to reproduce the performance of physical QA for typical tunneling problems as discussed in [14, 26]. This time we take the continuous time limit

![Figure 8. Residual energy $E(t_{\text{final}})$ as a function of total annealing time $t_{\text{final}}$ for different operators and Trotter slices $M$. Here, the QMC algorithm employs the same restricted cluster updates for the different operators TF and FI. The annealing time is not renormalized taking into account the computational effort (which is proportional to $M$) and we obtain the final energy by averaging over all the Trotter slices. Dashed lines indicate simulations with standard TF Hamiltonian (with $\Gamma_0 = 2$), while continuous ones refer to the FI operator, with annealing schedule given by equation (24), with (with $\Gamma_0 = 1$, $\Lambda_0 = 1$). We use an inverse temperature $\beta = 20$ and we check the continuous time limit convergence using $M = 64$ and 128. The FI Hamiltonian always performs better, for sufficiently long $t_{\text{final}}$, than the TF, within the same $M, \beta$ setup.](image)
by using a large enough number of time slices $M$, and we measure the annealing time simply using the Monte Carlo steps, as the computational effort related to the different update scheme is not relevant for the study of the physical machine. Moreover, following [13] we average the final energy over all the $M$ Trotter slices. In figure 8 we observe the same trend as in SQA, where the FI operator eventually outperforms the TF at later stages of the annealing.

6. Conclusions

We introduced a novel QMC algorithm to perform SQA with a transverse ferromagnetic two-spin driver Hamiltonian equation (4). This type of fluctuation extends the standard TF-SQA, adding also a two-body transverse operator of the form $\sigma_i^x \sigma_j^x$. Though this possibility has been introduced already some years ago [16], it has never been used before with QMC simulations, and therefore never applied so far on large optimization problems. We notice that, since QMC techniques are limited by the so-called sign problem, we can simulate only the ferromagnetic version of two body transverse interaction.

Our discrete-time path-integral Monte Carlo algorithm is an extension of the well established Loop Algorithm [36], with the inclusion of the transverse field operator, and implements restricted cluster updates to simulate efficiently large disordered systems.

A first application to quantum annealing, on a random square lattice, reveals that the new driver Hamiltonian improves upon the usual transverse field, though more systematic studies are required to address conclusively this question. Indeed it will be interesting to optimize the interplay of the two parameters $\Gamma(s)$ and $\Lambda(s)$ along the annealing schedule and to perform a size scaling analysis. Moreover it will be important to find other classes of problem Hamiltonians which can benefit more from this technique.

We notice that the range of applicability of the present algorithm can go beyond SQA, since it would be also possible to explore phase diagrams of XYZ sign-problem free Hamiltonians with transverse field, defined on arbitrary lattices, by performing equilibrium simulations with a converged number of Trotter time steps.

Acknowledgments

This work has been supported by the Swiss National Science Foundation through the National Competence Center in Research QSIT, by the European Research Council through ERC Advanced Grant SIMCOFE and by ODNI, IARPA via MIT Lincoln Laboratory Air Force Contract No. FA8721-05-C-0002. We acknowledge useful discussions with GE Santoro, H Katzgraber, D Herr and M Dolfi.

References

[1] Fu Y and Anderson P W 1986 Application of statistical mechanics to np-complete problems in combinatorial optimisation J. Phys. A: Math. Gen. 19 1605
[2] Farhi E, Goldstone J, Gutmann S and Sipser M 2000 Quantum computation by adiabatic evolution (arXiv:quant-ph/0001106)
Quantum Monte Carlo annealing with multi-spin dynamics

[33] Zhu Z, Ochoa A J and Katzgraber H G 2015 Efficient cluster algorithm for spin glasses in any space dimension Phys. Rev. Lett. 115 077201

[34] Rieger H and Kawashima N 1999 Application of a continuous time cluster algorithm to the two-dimensional random quantum ising ferromagnet Eur. Phys. J. B 9 233–6

[35] Takahashi M and Suzuki M 1972 One-dimensional anisotropic heisenberg model at finite temperatures Prog. Theor. Phys. 48 2187–209

[36] Evertz H G 2003 The loop algorithm Adv. Phys. 52 1–66

[37] Kawashima N 1996 Cluster algorithms for anisotropic quantum spin models J. Stat. Phys. 82 131–53

[38] Bauer B et al 2011 The alps project release 2.0: open source software for strongly correlated systems J. Stat. Mech. P05001

[39] www.informatik.uni-koeln.de/spinglass/

[40] Misra J and Gries David 1992 A constructive proof of vizing’s theorem Inf. Process. Lett. 41 131–133