Solving Quantum Spin Glasses with Off-Diagonal Expansion Quantum Monte Carlo

Itay Hen\textsuperscript{1,2} and Tameem Albash\textsuperscript{1,2}

\textsuperscript{1} Information Sciences Institute, University of Southern California, Marina del Rey, California 90292, USA
\textsuperscript{2} Department of Physics and Astronomy and Center for Quantum Information Science & Technology, University of Southern California, Los Angeles, California 90089, USA

E-mail: itayhen@isi.edu

Abstract. We study the performance of the Off-Diagonal Expansion quantum Monte Carlo algorithm on quantum spin glasses, specifically, 3-regular random antiferromagnets augmented with a transverse magnetic field. We demonstrate the advantages, sometimes by orders of magnitude, of the technique by comparing its runtimes against existing state-of-the-art schemes such as Path Integral quantum Monte Carlo and Stochastic Series Expansion.

1. Introduction

Solving for the ground state of spin glasses—disordered, frustrated Ising spin systems—is a formidable challenge. Spin glasses are prototypical classically-hard (also known as NP-hard) optimization problems \cite{1,2}, and they stymie all algorithms designed to solve them, so much so that specialized hardware has been built to simulate and study their thermal equilibrium properties \cite{3,4,5}. When such Ising spins systems are augmented with a non-commuting (quantum) component, forming quantum spin glasses, they continue to pose a significant challenge to Quantum Monte Carlo (QMC) algorithms.

QMC algorithms are known to be notoriously inefficient in describing quantum spin glasses in ‘almost classical’ parameter regimes, i.e., inside the spin-glass phase, where updates resulting from thermal fluctuations are expected to be far more dominant than those resulting from quantum fluctuations \cite{6,7,8}. Since QMC methods normally evolve via configuration updates that are based on quantum fluctuations, the acceptance rates of quantum updates, e.g., single spin (or cluster) flips in the dual Ising system, decrease dramatically in classical regimes, often causing QMC algorithms to dramatically slow down or ‘freeze’ (see, e.g., Ref \cite{9}). In these regimes, these models are known to exhibit additional structure that is not only interesting from a scientific standpoint but is also of critical importance to adiabatic approaches to quantum computation like quantum annealing \cite{10,11,12,13,14,15,16,17,18,19}, since the minima in the spin glass phase may be the relevant bottlenecks for the quantum algorithm as opposed to the critical point associated with the paramagnetic-spin glass phase transition. One reason that efficient classical thermal updates are typically hard to implement within the framework of QMC algorithms is that these algorithms do not normally converge to classical Monte Carlo algorithms in the limit where the model becomes classical.

Here, we analyze the performance of a recently developed algorithm, named the off-diagonal expansion (ODE) quantum Monte Carlo algorithm \cite{20}, devised specifically to have the algorithmic flexibility to simulate interacting many-body systems ranging from the fully-quantum to the fully-classical. The algorithm is based on a novel decomposition of the canonical quantum partition function into a sum of Boltzmann-like weights, and it converges to the usual...
decomposition of the classical partition function in the limit where the Hamiltonian of the system becomes classical. Based on this unique decomposition, the ODE algorithm aims to improve the convergence rates of simulated systems for which existing techniques are often inefficient.

We study random 3-regular MAX2SAT instances augmented with a transverse magnetic field. This class of instances corresponds to a particular choice of the Ising Hamiltonian, whereby each spin is coupled antiferromagnetically (with strength 1) with exactly three other spins picked at random. This class of instances is known to exhibit a quantum spin glass phase transition and is notoriously difficult to simulate by standard QMC techniques [9, 21], making it suitable to illustrate the strengths of the ODE algorithm.

Within the ODE approach the quantum imaginary-time dimension of the algorithm is 'elastic', i.e., it can stretch or shrink dynamically depending on the strength of the quantum part of the system — the off-diagonal portion of the Hamiltonian. In addition, the method does not introduce Trotter-type errors as in discrete-time path-integral QMC, a source of errors that normally occurs from an insufficient discretization of the imaginary-time dimension (over-discretization tends to sharply reduce the acceptance rates of the QMC updates). Moreover, in the classical limit where off-diagonal terms vanish, the algorithm naturally reduces to a classical thermal simulation. As we illustrate, the above properties allow the method to naturally overcome certain inefficiencies typically encountered by other QMC algorithms.

2. Decomposition of the partition function

We start by providing a brief description of the decomposition of the quantum partition function on which the ODE algorithm relies on, and the reader is referred to Ref. [20] for more details and a full derivation.

The canonical quantum partition function of a system described by a Hamiltonian $H$ is given by $Z = \text{tr} \left[ e^{-\beta H} \right]$. The Hamiltonian is first written in the form

$$
H = H_c - \sum_j \Gamma_j V_j
$$

where $H_c$ is the classical portion of the Hamiltonian, i.e., a diagonal operator in some known basis, which we refer to as the computational basis, and whose basis states will be denoted by $\{ |z \rangle \}$. The $\{ \Gamma_j \}$ are generally complex-valued parameters, and $\{ V_j \}$ are off-diagonal operators satisfying $[V_j, H_c] \neq 0$. In order for the decomposition of the partition function to be feasible, the off-diagonal operators are chosen to obey $V_j|z \rangle = |z' \rangle$ for every basis state $|z \rangle$, where $|z' \rangle \neq |z \rangle$ is also a basis state. We take the $\Gamma_j$ parameters to be identical, namely that $\Gamma_j = \Gamma, \forall j$, and we study the quantum random 3-regular MAX2SAT model whose Hamiltonian is

$$
H = \sum_{(i,j)} \sigma_i^x \sigma_j^x - \Gamma \sum_j \sigma_j^x,
$$

where $H_c = \sum_{(i,j)} \sigma_i^z \sigma_j^z$ is the classical Hamiltonian of 3-regular MAX2SAT whereby each spin is coupled antiferromagnetically (with strength 1) with exactly three other spins picked at random, and the off-diagonal operators are the spin-flip terms $V_j = \sigma_j^x$.

The partition function can be written as a sum over sequences $S_q = V_{i_1} V_{i_2} \ldots V_{i_q}$ of $q$ off-diagonal operators, with an associated classical ‘effective energies’ $E_{q(0 \ldots q)}$ as

$$
Z = \sum_{|z\rangle} \sum_{q} \sum_{\{S_q\}} \frac{(\beta \Gamma)^q}{q!} \langle z | S_q | z \rangle e^{-\beta E_{q(0 \ldots q)}},
$$

where $E_{q(0 \ldots q)}$ is the ‘effective classical energy’ of the configuration. The effective classical energy is calculated from the multiset of energies $\{E_0, \ldots, E_q\}$, where $E_0$ is the classical energy associated with $|z \rangle$ and $E_i$ is the classical energy associated with the computational basis state
\(|z'\rangle = \prod_{k=1}^{N} V_k |z\rangle\). To calculate \(E_{(0,...,q)}\), we can use the divided differences recursion relation (see Ref. [20]) to arrive at
\[
E_{(0,...,q)} = \bar{E} - \frac{1}{\beta} \log \frac{2q \sinh \beta \Delta E}{\beta (E_q - E_0)},
\]
(4)
where \(2\bar{E} = E_{(1,...,q)} + E_{(0,...,q-1)}\) and \(2\Delta E = E_{(1,...,q)} - E_{(0,...,q-1)}\). In the limiting case where all energies in the sequence are equal, the above relation neatly becomes \(E_{(0,...,q)} = E(0) = E_0\). The initial condition for the above recursion is simply \(e^{-\beta E(0)} = e^{-\beta E_i}\).

Since by construction the term \(\langle z|S_q|z\rangle\) evaluates to either 0 or to 1 (the operation \(S_q\) returns a basis state \(|z'\rangle\) and therefore \(\langle z|S_q|z\rangle = \langle z|z'\rangle = \delta_{z,z'}\)), the partition function can be neatly presented in its final form as a sum over only non-vanishing terms:
\[
Z = \sum_{\{S_z:|z|S_q|z\rangle = 1\}} (\beta \Gamma)^q q! e^{-\beta E_{(0,...,q)}},
\]
(5)
We interpret the terms in the sum in Eq. (5) as weights, i.e., \(Z = \sum_{\{C\}} W_C\), where the set of configurations \(\{C\}\) is all the distinct pairs \(|z\rangle, S_q\rangle\). Because of the form of \(W_C\),
\[
W_C = (\beta \Gamma)^q q! e^{-\beta E_{(0,...,q)}},
\]
(6)
we refer to it as a ‘generalized Boltzmann weight’ (or, GBW). We shall refer to \(E_{(0,...,q)}\) as the ‘effective classical energy’ of the configuration \(C\) and denote it at times for brevity simply by \(E_C\).

In order to interpret the \(W_C\) terms as actual weights, they must be non-negative for any simulated system that is not plagued by the sign problem [22]. As is evident from the above expression, even values of \(q\) always yield positive weights regardless of the sign of \(\Gamma\) for the Hamiltonian Eq. (2). Therefore, in order for the \(\langle z|S_q|z\rangle\) terms to evaluate to one rather than to zero, off-diagonal operators must always be produced and annihilated in pairs, implying that the total sign of the weight, Eq. (6), is positive.

3. The ODE algorithm
We now describe the basic ingredients of our Off-Diagonal Expansion (ODE) algorithm, based on the above partition function decomposition as it applies to the transverse-field Ising model, Eq. (2). We first establish the computational complexity associated with implementing this new algorithm, discussing in detail generic updates as well as measurements. We then present some results that allow us to fully characterize and to quantify the advantages of the algorithm over state-of-the-art QMC methods, specifically path-integral QMC and SSE.

3.1. General description of the algorithm
An ODE configuration is a pair \(C = \{z, S_q\}\) where \(z\) corresponds to a classical bit configuration and \(S_q = V_1 V_2 \cdots V_q\) is a sequence of (possibly repeated) off-diagonal operators. As was discussed above, each configuration \(C\) induces a list of states \(Z = \{|z_0\rangle, |z_1\rangle, \ldots, |z_q\rangle\} = \{|z\rangle\}\), which in turn also generates a corresponding multiset of diagonal energies \(M_C = \{E_0, E_1, \ldots, E_q\}\) of not-necessarily-distinct values (recall that \(E_i = \langle z_i|H_c|z_i\rangle\)). For systems with discretized energy values, the multiset can be stored efficiently in a ‘multiplicity table’ \(\{m_0, m_1, \ldots, m_j, \ldots\}\), where \(m_j\) is the multiplicity of the energy \(E_j\) in the multiset. Given \(M_C\), the evaluation of the effective classical energy \(E_C\) and the GBW \(W_C\) follow from the definition of the GBW, Eq. (6).

The initial configuration of the ODE algorithm is a random classical configuration \(|z\rangle\) and the empty sequence \(S_q=0\) = 1. The weight of this initial configuration is
\[
W_{C_{\text{init}}} = e^{-\beta E(z)},
\]
(7)
i.e., the classical Boltzmann weight of the initial random state \(|z\rangle\). Here the effective classical energy \(E_{C_{\text{init}}}\) is the classical energy of \(|z\rangle\).
3.2. Updates
We next describe the basic update moves for the algorithm. We consider here only generic local updates but note that updates of the global type can be specifically tailored to the system in question. An update is considered local if it changes the multiset \( M_C \) by a finite (i.e., by a system-size independent) number of terms, e.g., \( M_C \rightarrow M_C + \{ E(z_i) \} - \{ E(z_j) \} \). The basic updates are summarized in Fig. 1 and are discussed in detail below.

3.2.1. Classical moves
Classical moves are any moves that involve a manipulation of the classical state \( | z \rangle \) while leaving \( S_q \) unchanged [see Fig. 1(a)]. In a single bit-flip classical move, a spin from the classical bit-string state \( | z \rangle \) of \( C \) is picked randomly and is flipped, generating a state \( | z' \rangle \) and hence a new configuration \( C' \). Performing this change requires recalculating the energies associated with the sequence \( S_q \) leading to a new multiset \( M_{C'} \) and can become computationally intensive if \( q \) is large. Classical moves should therefore be attempted with low probabilities if \( q \) large. Simply enough, the acceptance probability for a classical move is

\[
p = \min \left( 1, \frac{W_{C'}}{W_C} \right) = \min \left( 1, e^{-\beta \Delta E} \right),
\]

where \( \Delta E = E_{C'} - E_C \) is the difference between the effective classical energies of the proposed configuration \( C' \) and current configuration \( C \).

In the absence of a quantum part to the Hamiltonian (\( \Gamma = 0 \)), not only are classical moves the only moves necessary, but they are also the only moves with a nonzero acceptance probability. In this case, the ODE algorithm automatically reduces to a classical Monte Carlo algorithm keeping the size of the imaginary-time dimension at zero (\( q = 0 \)) for the duration of the simulation.

3.2.2. Local swap
A local swap is the swapping of neighboring off-diagonal operators in the sequence \( S_q \). A random pair of adjacent off-diagonal operators in the sequence is picked and swapped [as shown in Fig. 1(b)]. If the state between \( V_k \) and \( V_{k+1} \) is \( |z\rangle \) and is \( |z'\rangle \) after the

![Figure 1. Basic update moves of the ODE algorithm. (a) Classical moves (e.g., a single bit flip), whereby only the initial state \( z \) is changed to \( z' \) leaving \( S_q \) unchanged. (b) Local swap, whereby two adjacent operators \( V_k V_{k+1} \) are interchanged changing the state between them from \( z' \) to \( z'' \). (c) Block swap, whereby two partitions of the sequence are interchanged. This changes the initial state from \( z \) to \( z' \) as well as the ordering of the sequence. (d) Pair creation/annihilation, whereby a new pair of operators is inserted or deleted.](image)
swap, then the swap involves adding an energy $E(z')$ and removing an energy $E(z)$ from the energy multiset [note that $E(z)$ and $E(z')$ may be the same]. The acceptance probability for the move is as in Eq. (8) with $M_C = M_C + \{E(z')\} - \{E(z)\}$.

3.2.3. Block-swap A block swap [Fig. 1(c)] is a local update that involves a change of the classical state $z$. Here, a random position $k$ in the sequence $S_q$ is picked such that the sequence is split into two (non-empty) parts, $S_q = S_1S_2$, with $S_1 = V_{i_1} \cdots V_{i_k}$ and $S_2 = V_{i_{k+1}} \cdots V_{i_q}$. The classical state $|z'\rangle$ at position $k$ in the sequence is given by

$$
|z'| = |z|S_1 = |z|V_{i_1} \cdots V_{i_k},
$$

where $|z\rangle$ is the classical state of the current configuration. The state $|z'\rangle$ has energy $E(z')$, and the state $|z\rangle$ has energy $E(z)$. We consider a new configuration defined by $\langle z'|S_2S_1|z\rangle$. The multiplicity table of this configuration differs from that of the current configuration by having one fewer $E(z)$ state and one additional $E(z')$ state. The weight of the new configuration is then proportional to $e^{-\beta M_C'}$ where the multiset $M_C' = M_C + \{E(z')\} - \{E(z)\}$. The acceptance probability is as in Eq. (8) with the aforementioned $M_C'$.

3.2.4. Creation/annihilation of off-diagonal operators The moves presented so far have left the size of $S_q$ unchanged. The creation/annihilation move shown in Fig. 1(d) has the effect of changing the value of $q$ by 2, i.e., $q \rightarrow q \pm 2$, which in the transverse-field Ising model corresponds to creating or destroying off-diagonal operators $\sigma_x^j$ in pairs. We implement this via the insertion or deletion of two adjacent, identical operators. With probability $p_{\text{del}}$ (e.g., $p_{\text{del}} = 1/2$) we try to annihilate an adjacent pair, and with probability $1 - p_{\text{del}}$ we try to insert a pair.

For pair insertion, we randomly pick an internal insertion point in the sequence (we denote this internal state by $|z'\rangle$) and a random $V$ to insert. This adds two new energies $E(z')$ and $E(z'')$ to the multiset, where $|z''\rangle = V|z'\rangle$. The acceptance probability for pair creation is given by

$$
p = \min \left( 1, \frac{p_{\text{del}}}{1 - p_{\text{del}}} \frac{N \beta^2 \Gamma^2}{(q + 2)(q + 3)} e^{-\beta \Delta E} \right)
$$

where as before $\Delta E = E_{C'} - E_C$ is the difference between the effective classical energies of the proposed configuration $C'$ and current configuration $C$ and $M_{C'} = M_C + \{E(z'), E(z'')\}$. For deletion, we randomly pick an internal point in the sequence. If the two operators to the side of the insertion point are not identical, no deletion is performed, and the move is rejected. If the two operators are identical, they are deleted and the relevant energies $E(z')$ and $E(z'')$ are removed from the multiplicity table. The probability of acceptance for the deletion move is given by

$$
p = \min \left( 1, \frac{1 - p_{\text{del}} q(q + 1)}{p_{\text{del}}} \frac{N \beta^2 \Gamma^2}{e^{-\beta \Delta E}} \right),
$$

where as before $\Delta E = E_{C'} - E_C$ and $M_{C'} = M_C - \{E(z'), E(z'')\}$.

The size of the imaginary time dimension $q$ comes strictly from off-diagonal terms and shrinks or grows depending on the strength of the ‘quantum component’ of the model. This property is expected to be heavily utilized in order to overcome the freezing of QMC algorithms in almost classical regimes. In these regimes, $q$ is small, and the algorithm reduces to being a classical thermal algorithm.\(^1\)

3.3. Measurements
An integral part of any QMC algorithm is the acquisition of various properties of the model such as average energy, magnetization, specific heat and correlation functions. In the ODE algorithm (as in SSE), diagonal (classical) measurements are measured differently than off-diagonal ones.\(^1\)

\(^1\) This is to be contrasted with the standard SSE formalism where one normally introduces an additional parameter $L$ in order to fix the size of imaginary time dimension for more efficient weight calculations. The fixing of the size of imaginary time may adversely affect the convergence of the algorithm if it is chosen to be too large. Here, this parameter too is spurious.
3.3.1. **Diagonal measurements**  
Diagonal operators $D$ obey $D|z\rangle = d(z)|z\rangle$ where $d(z)$ is a number that depends both on the operator and the state it acts on. Since $\langle z|DS_q|z\rangle = d(z)|z\rangle S_q|z\rangle$, for any given configuration $C = \{z\}, S_q$, there is a contribution $d = d(z)$ to the diagonal operator thermal average $\langle D \rangle$. To improve statistics, we can also consider rotations in the periodic imaginary time. To do that, we may consider ‘virtual’ block-swap moves (see Sec. 3.2.3) that rotate $S_q$ and as a result also change the classical configuration from $|z\rangle$ to $|z_i\rangle$. The contribution to the expectation value of a diagonal operator $D$ thus becomes:

$$d = \frac{1}{Z} \sum_{i=0}^{q-1} d(z_i) e^{-\beta E_{C_i}}.$$  

(12)

where $E_{C_i}$ is the effective classical energy associated with configuration $C_i$ whose multiset is $M_{C_i} = M_c + \{E(z_i)\} - \{E(z)\}$ (recall that $z_0 \equiv z$, so $M_{C_0} = M_c$). The normalization factor $Z$ above is the sum

$$Z = \sum_{j=0}^{q-1} e^{-\beta E_{C_j}} = \sum_j m_j e^{-\beta E_{C_j}}$$

(13)

over all nonzero multiplicities $m_j$. In the case where $D = H_c$ the above expression simplifies to:

$$d = \frac{1}{Z} \sum_{i=0}^{q-1} E(z_i) e^{-\beta E_{C_i}} = \frac{1}{Z} \sum_j m_j E(z_j) e^{-\beta E_{C_j}}.$$  

(14)

3.3.2. **Off-diagonal measurements**  
We next consider the case of measuring the expectation value of an off-diagonal operator $V_k$, namely, $\langle V_k \rangle$. To do this, we interpret the instantaneous configuration as follows

$$W_C = \frac{(\beta \Gamma)^q e^{-\beta E_{C'}}}{q!} (|S_q\rangle) = \left( \frac{\beta \Gamma}{q e^{-\beta \Delta E}} \right) \left[ \frac{(\beta \Gamma)^{q-1} e^{-\beta E_{C'}}}{(q-1)!} \langle z|S_{q-1}V_k|z\rangle \right],$$

(15)

where $\Delta E = E_{C'} - E_{C}$ and $C'$ is the configuration associated with the multiset $M_{C'} = M_c - \{E(z)\}$. In the above form, we can reinterpret the weight $W_C$ as contributing

$$v_k = \delta_{k,q} \frac{q e^{-\beta \Delta E}}{\beta \Gamma},$$

(16)

to $\langle V_k \rangle$.

As in the case of the diagonal measurements, one can take advantage of the periodicity in the imaginary time direction to improve statistics by rotating the sequence such that any of the elements of $S_q$ becomes the last element of the sequence (see Sec. 3.2.3), weighted accordingly by the block-swap probability. By doing so, $v_k$ becomes

$$v_k = \sum_i \frac{q}{\beta \Gamma} \frac{e^{-\beta E_{C_i}}}{\sum_{j=0}^{q-1} e^{-\beta E_{C_j}}} e^{-\beta E_{C'}} = \frac{q N_k}{\beta \Gamma} e^{-\beta E_{C'}}$$

(17)

where $M_{C_i} = M_c + \{E(z_i)\} - \{E(z)\}$, the sum $\sum_i$ is over all rotated configurations $C'$ whose $S_q$ ends with $V_k$, and $N_k$ is the number of times $V_k$ appears in the sequence $S_q$.

3.4. **Quantum-classical parallel tempering**  
As we reviewed in Sec. 2, the ODE partition function decomposition naturally reduces to the classical one when the strength of the off-diagonal terms in the Hamiltonian are sent to zero. As we show next, this allows us to naturally unify the classical Parallel Tempering (CPT) algorithm (also known as ‘exchange Monte Carlo’) [23, 24] and its quantum counterpart (QPT, see e.g., Refs [25, 26]). CPT is a refinement of the simulated annealing algorithm [27], whereby $N_T$
replicas of an $N$-spin system at inverse-temperatures $\beta_1 < \beta_2 < \ldots < \beta_N$ undergo Metropolis spin-flip updates independently of one another and in addition, replicas with neighboring temperatures regularly attempt to swap their temperatures with probabilities that satisfy detailed balance \cite{28}. In this way, each replica performs a random-walk on the temperature axis, which generally allows for quicker equilibration of the system in comparison to other techniques. Analogously in QPT, temperature is replaced by a parameter $\Gamma$ of the (quantum) Hamiltonian, e.g., the strength of the transverse magnetic field in the transverse Ising model, and each replica performs a random-walk on the $\Gamma$ axis.

Here, we shall consider replicas along a curve in the $\beta$-$\Gamma$ plane at points $\{(\beta_1, \Gamma_1), \ldots, (\beta_N, \Gamma_N)\}$, then a parallel tempering swap probability between the $i$-th and $(i+1)$-th replica is given by:

$$P = \min \left( 1, \frac{W_{\mathcal{C}_i}(\beta_{i+1}, \Gamma_{i+1})W_{\mathcal{C}_{i+1}}(\beta_i, \Gamma_i)}{W_{\mathcal{C}_i}(\beta_i, \Gamma_i)W_{\mathcal{C}_{i+1}}(\beta_{i+1}, \Gamma_{i+1})} \right),$$

where the above weight ratio is conveniently simplified to:

$$\frac{W_{\mathcal{C}_i}(\beta_{i+1}, \Gamma_{i+1})W_{\mathcal{C}_{i+1}}(\beta_i, \Gamma_i)}{W_{\mathcal{C}_i}(\beta_i, \Gamma_i)W_{\mathcal{C}_{i+1}}(\beta_{i+1}, \Gamma_{i+1})} = \left( \frac{\beta_i \Gamma_i}{\beta_{i+1} \Gamma_{i+1}} \right)^{q_i+q_{i+1}} e^{-\beta_i(E_{\mathcal{C}_{i+1}} - E_{\mathcal{C}_i})} e^{-\beta_{i+1}(E_{\mathcal{C}_i} - E_{\mathcal{C}_{i+1}})},$$

where $E_{\mathcal{C}_i}$ and $E_{\mathcal{C}_{i+1}}$ are the effective classical energies of configurations $\mathcal{C}_i$ and $\mathcal{C}_{i+1}$, respectively and $E_{\mathcal{C}_i}'$ and $E_{\mathcal{C}_{i+1}}'$ are the effective classical energies of these configurations when calculated with switched $\beta$ and $\Gamma$.

4. Performance of ODE on quantum spin glasses

Having laid the groundwork for the ODE QMC algorithm, we present in this section some results that highlight some of its properties and advantages over existing QMC techniques, specifically a discrete-time Path Integral quantum Monte Carlo (PIQMC) algorithm with cluster-updates\footnote{We use Wolff cluster updates \cite{29} along the imaginary time direction \cite{30, 31}.} and with the continuous-time SSE algorithm with cluster updates \cite{32, 9}.

To demonstrate the advantages of ODE over existing PIQMC and SSE, we carry out multiple runs of parallel tempering QMC simulations on random MAX2SAT instances augmented with a transverse field [see Eq. (2)] with various problem sizes $N$, a range of $\beta$ (inverse temperature) and $\Gamma$ (quantum strength) values.

Runtime comparisons of ODE against both PIQMC and SSE are summarized in Figs. 2–5. The figures depict the performance of ODE, PIQMC and SSE as it is reflected by the thermal average of $x$-magnetization $H_x = -\sum_{i=1}^{N} \sigma_i^x$, of the classical Hamiltonian $H_c$, and of the specific heat per spin $C = \beta^2 \left( \langle H^2 \rangle - \langle H \rangle^2 \right) / N$ (other observables exhibit similar behavior) for increasing problem size and different values of $\beta$ and $\Gamma$. Each data point in each of the panels is an average over 8 random instances.

The first three figures, Figs. 2–4, depict the results of PT simulations done along a fixed $\Gamma$ line in the $\beta$-$\Gamma$ plane, that is, different $\beta$ values. As is clearly evident from the figures, while for large temperatures ($\beta \leq 2$) the performance of all three algorithms is comparable, with decreasing temperature and increasing problem size, ODE performs up to four or more orders of magnitude faster than the other two algorithms. Specifically, a large fraction of the instances did not finish running over the 24 hour window allocated for each of the simulations, whereas ODE had converged after a few seconds.

We note that while PIQMC can perform comparably to the ODE for the evaluation of the average $H_c$, PIQMC behaves poorly for the evaluation of the average of $H_x$. Similarly, while SSE performs better than PIQMC for evaluating the average of $H_x$, it performs worse for evaluating the average of $H_c$. From this perspective, we can conclude that the performance of ODE is also consistently better for different observables, which is reflected in its superior performance for calculating the average of the specific heat in Fig. 4.
Figure 2. Thermal average of the off-diagonal Hamiltonian as obtained by ODE, PIQMC and SSE for a fixed $\Gamma = 0.1$. Results are obtained using PT simulations of different system sizes $N = 60, 96$ and $128$, each with 11 replicas at inverse temperatures, $\beta \in \{0.1, 0.2, 0.5, 1, 2, 5, 10, 20, 30, 40, 50\}$. A subset of the results are shown for $\beta = 2, 10$ and 50. Each data point is the mean value of $10^3$ bootstraps performed on the 500 measurements of the $x$-magnetization-per-spin taken in the course of simulations with different runtimes (horizontal axis), while the error bars correspond to the $2\sigma$ confidence interval generated by the bootstrap. For PIQMC we use 1280 Trotter slices. As is evident, ODE converges to the true value at times orders of magnitude faster than both SSE and PIQMC [33].

Figure 5 depicts the results of PT simulations done along a fixed $\beta$ line in the $\beta$-$\Gamma$ plane, that is, different $\Gamma$ values. Interestingly, here all three algorithms seem to perform comparably, indicating that ODE is, as expected, at its optimum in near classical regimes.

To demonstrate the performance of the ODE algorithm away from the ‘almost classical’ regime, i.e., in regimes where the quantum strength parameter $\Gamma$ is not small, we show in Fig. 6 some results obtained for $\Gamma = 0.4$ and 0.5 for different problem sizes. Here too, we find a regime where ODE converges faster to the true values by several orders of magnitude than PIQMC (not shown in this figure are SSE results, which we expect performs similarly to PIQMC in these regimes as is evident from the data presented in the preceding figures).

5. Conclusions

We have presented results for the performance of the ODE algorithm for a class of quantum spin glasses, 3-regular MAX2SAT with a transverse field, relative to PIQMC and SSE. Near the classical regime, when $\Gamma$ the strength of the transverse is small, a clear advantage can be observed over PIQMC and SSE, with at times orders of magnitude faster performance. This advantage is expected because the ODE method has an elastic quantum dimension, which can shrink appropriately in the near classical limit.
Figure 3. Thermal average of the classical Hamiltonian as obtained by ODE, PIQMC and SSE for a fixed $\Gamma = 0.1$. Results are obtained using PT simulations of different system sizes $N = 60, 96$ and 128, each with 11 replicas at inverse temperatures, $\beta \in \{0.1, 0.2, 0.5, 1, 2, 5, 10, 20, 30, 40, 50\}$. A subset of the results are shown for $\beta = 2, 10$ and 50. Each data point is the mean value of $10^3$ bootstraps performed on the 500 measurements of the $x$-magnetization-per-spin taken in the course of simulations with different runtimes (horizontal axis), while the error bars correspond to the $2\sigma$ confidence interval generated by the bootstrap. For PIQMC we use 1280 Trotter slices. As is evident, ODE converges to the true value at times orders of magnitude faster than both SSE and PIQMC [33].

We have demonstrated how the algorithm applies to the transverse-field Ising model. It would be interesting to see how it performs with respect to existing techniques on other models considered difficult to simulate. Another aspect worth studying is the existence of additional updates that are more global in nature in order to further speed up convergence. These will more likely have to be specifically tailored to the system in question. Last, methods to facilitate the evaluation of the generalized Boltzmann weights are of significance as these scale in the worst case as the square of the imaginary time dimension. More efficient methods will serve to further increase the usefulness of the ODE algorithm. We leave the resolution of these questions for future work.

Acknowledgments
Computation for the work described here was supported by the University of Southern California’s Center for High-Performance Computing (http://hpcc.usc.edu). The research is based upon work (partially) supported by the Office of the Director of National Intelligence (ODNI), Intelligence Advanced Research Projects Activity (IARPA), via the U.S. Army Research Office contract W911NF-17-C-0050. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies.
Figure 4. Thermal average of the specific heat per spin $C = \beta^2 \left( \langle H^2 \rangle - \langle H \rangle^2 \right) / N$ as obtained by ODE, PIQMC and SSE for a fixed $\Gamma = 0.1$. Results are obtained using PT simulations of different system sizes $N = 60, 96$ and 128, each with 11 replicas at inverse temperatures, $\beta \in \{0.1, 0.2, 0.5, 1, 2, 5, 10, 20, 30, 40, 50\}$. A subset of the results are shown for $\beta = 1, 2$ and 5. Each data point is the mean value of $10^3$ bootstraps performed on the 500 measurements of the $x$-magnetization-per-spin taken in the course of simulations with different runtimes (horizontal axis), while the error bars correspond to the $2\sigma$ confidence interval generated by the bootstrap. For PIQMC we use 1280 Trotter slices. As is evident, ODE converges to the true value at times orders of magnitude faster than both SSE and PIQMC [33].

or endorsements, either expressed or implied, of the ODNI, IARPA, or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright annotation thereon.

References
[1] A. P. Young, ed., *Spin Glasses and Random Fields*. Singapore: World Scientific, 1998.
[2] F. Barahona, “On the computational complexity of Ising spin glass models,” *J. Phys. A: Math. Gen.*, vol. 15, no. 10, p. 3241, 1982.
[3] F. Belletti, M. Cotallo, A. Cruz, L. A. Fernandez, A. Gordillo, A. Maiorano, F. Mantovani, E. Marinari, V. Martin-Mayor, A. Muñoz Sudupe, D. Navarro, S. Perez-Gaviro, J. J. Ruiz-Lorenzo, S. F. Schifano, D. Sciorti, A. Tarancón, R. Tripiccione, and J. L. Velasco, “Simulating spin systems on JANUS, an FPGA-based computer,” *Comp. Phys. Comm.*, vol. 178, pp. 208–216, 2008.
[4] F. Belletti, M. Guidetti, A. Maiorano, F. Mantovani, S. F. Schifano, R. Tripiccione, M. Cotallo, S. Perez-Gaviro, D. Sciorti, J. L. Velasco, A. Cruz, D. Navarro, A. Tarancón, L. A. Fernandez, V. Martin-Mayor, A. Muñoz-Sudupe, D. Yllanes, A. Gordillo-Guerrero, J. J. Ruiz-Lorenzo, E. Marinari, G. Parisi, M. Rossi, and G. Zanier, “JANUS: An FPGA-based system for high-performance scientific computing,” *Computing in Science and Engineering*, vol. 11, p. 48, 2009.
[5] M. Bait-Jesi, R. A. Baños, A. Cruz, L. A. Fernandez, J. M. Gil-Narvion, A. Gordillo-Guerrero, D. Iniguez, A. Maiorano, F. Mantovani, E. Marinari, V. Martin-Mayor, J. Monforte-Garcia, A. Muñoz Sudupe,
Figure 5. Thermal average of the off-diagonal Hamiltonian as obtained by ODE, PIQMC and SSE for a fixed $\beta = 10$. Results are obtained using PT simulations of different system sizes $N = 60$ and 96, each with 5 replicas at transverse field strengths, $\Gamma \in \{0.1, 0.2, 0.3, 0.4, 0.5\}$. A subset of the results are shown for $\Gamma = 0.3, 0.4$ and 0.5. Each data point is the mean value of $10^3$ bootstraps performed on the 500 measurements of the $x$-magnetization-per-spin taken in the course of simulations with different runtimes (horizontal axis), while the error bars correspond to the $2\sigma$ confidence interval generated by the bootstrap. For PIQMC we use 1280 Trotter slices. Here, all three algorithms perform similarly[33].

Figure 6. Thermal average of the off-diagonal Hamiltonian as obtained by ODE and PIQMC at large $\Gamma$ values. Results are obtained using PT simulations of different system sizes $N = 60, 96$ and 128, each with 7 replicas at inverse temperatures, $\beta \in \{0.1, 0.2, 0.5, 1, 2, 5, 10\}$. Results are shown for $\beta = 10$. Each data point is the mean value of $10^3$ bootstraps performed on the 500 measurements of the $x$-magnetization-per-spin taken in the course of simulations with different runtimes (horizontal axis), while the error bars correspond to the $2\sigma$ confidence interval generated by the bootstrap. For PIQMC we use 1280 Trotter slices. As is evident, even at large $\Gamma$ values ODE can converge faster than PIQMC by several orders of magnitude [33].
D. Navarro, G. Parisi, S. Perez-Gaviro, M. Pivanti, F. Ricci-Tersenghi, J. J. Ruiz-Lorenzo, S. F. Schifano, B. Seoane, A. Tarancón, R. Trippicione, and D. Yllanes, “Janus II: a new generation application-driven computer for spin-system simulations,” 

Comp. Phys. Comm., vol. 185, pp. 550–559, 2014.

[6] P. Pfeuty and R. J. Elliott, “The ising model with a transverse field. ii. ground state properties,” Journal of Physics C: Solid State Physics, vol. 4, no. 15, p. 2370, 1971.

[7] R. B. Stinchcombe, “Ising model in a transverse field. i. basic theory,” Journal of Physics C: Solid State Physics, vol. 6, no. 15, p. 2459, 1973.

[8] B. K. Chakrabarti, “Critical behavior of the ising spin-glass models in a transverse field,” Phys. Rev. B, vol. 24, pp. 4062–4064, Oct 1981.

[9] E. Farhi, D. Gosset, I. Hen, A. W. Sandvik, P. Shor, A. P. Young, and F. Zamponi, “Performance of the quantum adiabatic algorithm on random instances of two optimization problems on regular hypergraphs,” Phys. Rev. A, vol. 86, p. 052334, Nov 2012.

[10] B. Apolloni, C. Carvalho, and D. de Falco, “Quantum stochastic optimization,” Stochastic Processes and their Applications, vol. 33, pp. 233–244, 12 1989.

[11] B. Apolloni, N. Cesa-Bianchi, and D. de Falco, “A numerical implementation of quantum annealing,” in Proceedings of the Ascona/Locarno Conference, p. 97, 1988.

[12] R. L. Somorjai, “Novel approach for computing the global minimum of proteins. 1. general concepts, methods, and approximations,” The Journal of Physical Chemistry, vol. 95, pp. 4141–4146, 05 1991.

[13] P. Amara, D. Hsu, and J. E. Straub, “Global energy minimum searches using an approximate solution of the imaginary time schroedinger equation,” The Journal of Physical Chemistry, vol. 97, pp. 6715–6721, 2015/01/03 1993.

[14] A. B. Finnila, M. A. Gomez, C. Sebenik, C. Stenson, and J. D. Doll, “Quantum annealing: A new method for minimizing multidimensional functions,” Chemical Physics Letters, vol. 219, pp. 343–348, 3 1994.

[15] T. Kadowaki and H. Nishimori, “Quantum annealing in the transverse Ising model,” Phys. Rev. E, vol. 50, p. 3535, Nov. 1999.

[16] J. Brooke, D.比特科夫, T. F., Rosenbaum, and G. Aeppi, “Quantum annealing of a disordered magnet,” Science, vol. 284, no. 5415, pp. 779–781, 1999.

[17] J. Brooke, T. F. Rosenbaum, and G. Aeppi, “Tunable quantum tunnelling of magnetic domain walls,” Nature, vol. 413, pp. 610–613, Oct. 2001.

[18] E. Farhi, J. Goldstone, S. Gutmann, and M. Sipser, “Quantum computation by adiabatic evolution,” ArXiv, 2000.

[19] E. Farhi, J. Goldstone, S. Gutmann, J. Lapan, A. Lundgren, and D. Preda, “A quantum adiabatic evolution algorithm applied to random instances of an NP-Complete problem,” Science, vol. 292, pp. 472–475, Apr 2001.

[20] T. Albash, G. Wagenbreth, and I. Hen, “Off-diagonal expansion quantum monte carlo,” Phys. Rev. E, vol. 96, p. 063309, Dec 2017.

[21] C.-W. Liu, A. Polkovnikov, and A. W. Sandvik, “Quantum versus classical annealing: Insights from scaling theory and results for spin glasses on 3-regular graphs,” Phys. Rev. Lett., vol. 114, p. 147203, Apr 2015.

[22] M. N. G. Barkema, Monte Carlo Methods in Statistical Physics. Oxford University Press, 1999.

[23] K. Hukushima and K. Nemoto, “Exchange monte carlo method and application to spin glass simulations,” J. Phys. Soc. Japan, vol. 65, p. 1604, 1996.

[24] E. Marinari, “Optimized Monte Carlo methods,” in Advances in Computer Simulation (J. Kertész and I. Kondor, eds.), p. 50, Springer-Verlag, 1998. (arXiv:cond-mat/9612010).

[25] P. Sengupta, A. W. Sandvik, and D. K. Campbell, “Bond-order-wave phase and quantum phase transitions in the one-dimensional extended hubbard model,” Phys. Rev. B, vol. 66, p. 155113, Apr 2002.

[26] I. Hen and A. P. Young, “Exponential complexity of the quantum adiabatic algorithm for certain satisfiability problems,” Phys. Rev. E, vol. 84, p. 061152, Dec 2011.

[27] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, “Optimization by simulated annealing,” Science, vol. 220, pp. 671–680, May 1983.

[28] A. Sokal, “Monte carlo methods in statistical mechanics: Foundations and new algorithms,” in Functional Integration: Basics and Applications (C. DeWitt-Morette, P. Cartier, and A. Folacci, eds.), Plenum, 1997.

[29] U. Wolff, “Collective Monte Carlo updating for spin systems,” Phys. Rev. Lett., vol. 62, pp. 361–364, Jan 1989.

[30] R. Martońák, G. E. Santoro, and E. Tosatti, “Quantum annealing by the path-integral Monte Carlo method: The two-dimensional random Ising model,” Phys. Rev. B, vol. 66, p. 094203, Sep 2002.

[31] S. Boixo, T. F. Ronnow, S. V. Isakov, Z. Wang, D. Wecker, D. A. Lidar, J. M. Martinis, and M. Troyer, “Evidence for quantum annealing with more than one hundred qubits,” Nat. Phys., vol. 10, pp. 218–224, 03 2014.

[32] A. W. Sandvik, “Stochastic series expansion method for quantum Ising models with arbitrary interactions,” Phys. Rev. E, vol. 68, p. 056701, 2003.

[33] The simulations were performed on individual cores of Intel Xeon E5-2640 CPUs running at 2.60GHz.