Kinetic Monte Carlo study of accelerated optimization problem search using Bose-Einstein condensates

Kai YAN¹, Tim BYRNES², and Yoshihisa YAMAMOTO³
¹,²,³National Institute of Informatics
²Institute of Industrial Science, University of Tokyo
³E.L. Ginzton Laboratory, Stanford University

ABSTRACT

In a previous work [ArXiv:0909.2530] we proposed a method for accelerating optimization problem search using Bose-Einstein condensation (BEC). The system encodes an optimization problem into an Ising model and cools it down by the process of BEC to find its ground state spin configuration which corresponds to the solution of the problem. The system uses the final state stimulation (FSS) property of bosonic particles, an effect originating from the quantum indistinguishability of bosons, to provide speedups over the classical case. The speedup is typically $\propto N$, where $N$ is the number of bosons in the system per site. In this article we firstly review the proposed system, and give a more detailed numerical study of the equilibration time with the boson number and the number of sites $M$ in the Ising model.

We find that the equilibration time scales as $\tau \sim \exp(M)/N$ in agreement with previous arguments based on simulated annealing. A detailed description of the kinetic Monte Carlo method used for the study of the proposed system is also discussed.

KEYWORDS

Bose-Einstein condensation, optimization, kinetic Monte Carlo, quantum computing

1 Introduction

Optimization problem search usually takes huge amount of computational time to find the optimal solution. In particular, for the class of NP-complete problems [1] the time complexity grows exponentially with the problem size. Solving large-scale NP-complete problems is a nontrivial and important task with many practical applications. Quantum computers promise to offer a new approach to computation with great increases in speed over current computers for certain kinds of problems. Shor’s factoring algorithm [2] and Grover’s database search [3] algorithms are two examples where quantum computers have been shown to outperform the best classical computer algorithms. However, it has proven rather difficult to construct quantum algorithms for a general computational problem. Secondly, in the ideal case quantum computation [4], assumes that the computation is performed by manipulating qubits using perfect unitary transformations, which implicitly assumes that there is zero coupling to an external environment. In practice there is always some small coupling to the environment which results in decoherence, a major obstacle in the quest for building a scalable quantum computer. Therefore, alternative methods of computation beyond the standard quantum computation model that possess speedups beyond the classical case are of distinct practical interest.

Previously we have proposed an alternative model of computation [5] that involves a dissipative coupling to the environment, and a large number of (quantum) bosonic particles in the system. The general idea of the scheme is to cool the system from an initial high temperature state into the ground state, via the dissipative coupling to the environment. The approach has similarities with thermal annealing where the ground state is found via cooling the system, rather than performing a computation in the traditional sense. We have shown

DOI: 10.2201/NiiPi.2011.8.5
that there is a quantum speedup by using indistinguishable bosons in the system, originating from the bosonic final state stimulation (FSS) effect [6]. FSS refers to the effect where bosons cool faster by a factor of $N + 1$ in a dissipative system, when there are $N$ bosons occupying the destination state. In this article we give a more detailed presentation of the system and investigate the scaling of the system with problem size.

2 Review of the proposed system

We start by considering a representative optimization problem, MAX-CUT, which is known to be in the class of NP-complete problems [7]. Given a graph with $M$ vertices and connection links between them, the task is to group the vertices into two groups $A$ and $B$ such as to minimize the number of connections between the groups (Fig. 1). Such a computational problem can be formulated as an energy minimization problem of a given Hamiltonian [8]. By assigning each vertex to a spin $\sigma$, (i.e. the “up” state spin as the vertex being grouped in $A$ and “down” state as in $B$) the problem is equivalent to finding the minimal energy spin configuration $\{\sigma_i\}$ of an Ising model

$$H = \sum_{i,j=1}^{M} J_{ij} \sigma_i \sigma_j. \quad (1)$$

For this particular problem, $J_{ij} = 1$ if the vertices are connected and zero otherwise. For other types of optimization problems, the $J_{ij}$ can take any real numbered values such that the matrix is symmetric.

The problem one typically encounters in such a scheme is that it is very difficult to find the ground state of such a Hamiltonian because of the nature of the configurational energy landscape. In simulated annealing [9], very long annealing times are necessary to ensure that the system does not get caught in local minima. Quantum annealing [10] overcomes such problems due to local minima by introducing a quantum tunneling term but requires a slow adiabatic evolution to prevent leaks into excited states. The basic idea of this work is to take advantage of the fact that bosons tend to accumulate in the ground state of a system, and that they reach this ground state very quickly through the process of FSS, to find the ground state of (1).

The schematic device proposed in Ref. [5] is shown in Fig. 2. Each spin $\sigma_i$ in the Hamiltonian (1) is associated with a trapping site containing $N$ bosons. In each trap, the bosons can occupy one of two states, which we label by $\sigma = \pm 1$. Any particle that displays bosonic statistics may be used, such as neutral atoms used to create BECs. Atom chips may therefore be a suitable candidate for making such a device [11]. Another possibility is to use exciton-polaritons in semiconductor microcavities, which have recently observed to undergo BEC [12]–[14]. BECs are convenient for this purpose since we assume only two states on each site, which could be provided by the internal degree of freedom of the bosonic particles used. Exciton-polaritons possess a spin of $\sigma = \pm 1$ which can be injected by optical pumping with a linearly polarized laser beam. The interactions between the sites are externally controlled (Fig. 2b) such as to follow the Hamiltonian

![Fig. 1](image1.png)  A simple example of MAX-CUT problem.

![Fig. 2](image2.png) (a) The schematic device configuration. Each site of the Ising Hamiltonian is encoded as a particular site, containing $N$ bosons. The bosons can occupy one of two states $\sigma = \pm 1$, depicted as either red or blue. (b) The interaction between the sites denoted by the arrows in (a) are externally controlled by a measurement feedback loop in order to produce the desired interaction $J_{ij}$. 


where $S_i = \sum_i \sigma_i^z$ is the total spin on each site $i$, and $J_{ij}$ is the same matrix as (1) which specifies the computational problem. The interaction (2) can be created by measuring the average spin on each site and feeding this back into the system by applying a field $B_i = \sum_{j=1}^M J_{ij} S_j$ on site $i$ (see Fig. 2b). Although in general $J_{ij}$ has a large connectivity and is long-ranged, by using such a feedback method to induce the interactions there is no restriction to the kind of interactions $J_{ij}$ that can be produced in principle.

Initially each site is prepared in a random configuration of $\sigma = \pm 1$ spins such that the ensemble average gives zero spin on each site. The system is cooled in the presence of the interactions between the sites, by immersing the system in an external heat bath. The read-out of the computation is simply performed by measuring the average spin on each site after some time after the system cools down. Since the “computation” here is the cooling process itself, no complicated gate sequence needs to be employed in contrast to the standard model of quantum computation.

The computation on a BEC system starts from a random spin state in each site, and evolves with the presence of the interaction, finally the system reaches a dynamic equilibrium. The time required to reach the equilibrium thus is considered to be the computation time. The time evolution of the bosonic Ising Hamiltonian is performed by an extension of the method presented by Glauber[15] to bosons. Given the $M$-site Hamiltonian

$$ H = \sum_{i,j} M J_{ij} S_i S_j + \lambda N \sum_i S_i, \quad (3) $$

the states are labeled

$$ |k⟩ = \prod_{i=1}^M \frac{1}{\sqrt{k_i!(N - k_i)!}} (a_i^{\dagger})^{k_i} (a_i)_{N-k_i} |0⟩, \quad (4) $$

where the $k_i$ range from 0 to $N$, $a_i^{\dagger}$ is the creation operator for a boson on site $i$ in the state $\sigma$, and we have defined the vector $k = (k_1, k_2, ..., k_M)$. At $t = 0$, there is an equal probability $p_k$ of all states $|k⟩$. The probability distribution then evolves according to the equation

$$ \frac{dp_k}{dt} = \sum_{i,j} N_k \sum_{\delta k} \{ -w(k, \delta k)p_k + w(k + \delta k, -\delta k)p_{k+\delta k} \}, \quad (5) $$

where $p_k$ is the probability of the system in the state $k$. Here $\delta k = (0, ..., 0, \delta i, 0, ..., 0)$, the $w(k, \delta k)$ is a weight factor for the process $|k⟩ \rightarrow |k + \delta k⟩$, containing a transition rate factor from Fermi’s golden rule and a coefficient to ensure that the system evolves to the correct thermal equilibrium distribution, in a similar way to that discussed in [15].

We first derive the transition rate for the process $|k⟩ \rightarrow |k + \delta k⟩$. To derive the transition rate for $|\delta k⟩ > 1$, we require a generalization of Fermi’s Golden rule beyond first order in perturbation theory. Such a generalization is provided in [16], considering the case of a multi-level atom transitioning to a continuum. In our case, the bosonic site plays the role of the multi-level atom and the continuum is the phonon reservoir that the system loses energy to. We repeat the final result for a $N$th order transition here for convenience:

$$ T = \frac{2\pi}{\hbar} \rho(E_K)|V_{KN-1}|^2 \prod_{n=1}^{N-1} \frac{|V_{n-1}|^2}{\hbar^2(E_K) + S_n^2(E_K)}, \quad (6) $$

where

$$ h_{N-1}(E_K) = E_K - E_{N-1} - \delta_{N-1,K}(E_K) $$

$$ s_{N-1}(E_K) = \frac{1}{2} \gamma_{N-1,K}(E_K) $$

$$ h_{n-1}(E_K) = E_K - E_{n-2} - |V_{n-2,n-1}|^2 h_{n-1}(E_K) $$

$$ s_{n-1}(E_K) = \frac{|V_{n-2,n-1}|^2 s_{n-1}(E_K)}{h_{n-1}(E_K) + s_{n-1}(E_K)} $$

$$ \delta_{N-1,K}(E_K) = \rho(E_K)|[V_{N-1,K}|^2 \frac{dE}{dK} - E_K $$

$$ \gamma_{N-1,K}(E_K) = 2\pi \int \rho(E_K)|[V_{N-1,K}|^2 \frac{dE}{dK} \frac{dE}{dK} \quad (7) $$

and $E_K$ is the energy of the continuum state $K$, $E_n$ is the energy of the $n$th atomic energy level, $\rho(E_K)$ is the density of states for the continuum state $K$, and $V_{nm}$ is the transition matrix element between the $n$th and $m$th atomic levels. For $N = 1$ the transition rate reduces to that of Fermi’s golden rule.

Applying this formula to our case, we assume a perturbative transition on the system

$$ V \propto g \sum_j [a_j^{\dagger} a_1 + a_1^{\dagger} a_j], \quad (8) $$

and energy levels in the bosonic site that are equally spaced

$$ E_n = n\Delta E + E_K. \quad (9) $$

The matrix elements are

$$ |(k_1 + 1)|V|k⟩|^2 = g^2(k_1 + 1)(N - k_1). \quad (10) $$

If we assume that the perturbation $V$ is very weak ($g \ll \Delta E$), we may simplify the expression (6) and take

$$ h_{n-1}^2(E_K) + s_{n-1}^2(E_K) = h_{n-1}^2(E_K) $$
giving

\[ T(k \rightarrow k + \delta k_i) = \begin{cases} \frac{\alpha \xi^{\delta k - 1}}{(\delta k - 1)!^2} & (\delta k > 0) \\ \prod_{m=1}^{\delta k} (k + m)(N - k - \delta k + m) & (\delta k < 0) \end{cases} \]

where we have defined the parameters

\[ \alpha = \frac{2\pi}{\hbar} (E_k) g^2 \]

\[ \xi = \left( \frac{2}{\Delta E} \right)^2 \]

and compacted the notation such that \( \delta k_i \rightarrow \delta k \).

We take the parameters \( \alpha \) and \( \xi \) as phenomenological parameters that may be determined by experiment. For our numerical simulations, we measure all time scales relative to \( \alpha \) (or equivalently set \( \alpha = 1 \)), and choose a small value for \( \xi \), i.e. \( \xi \ll 1 \).

At thermal equilibrium, the transition rates are equal between the states \( [k] \leftrightarrow [k + \delta k] \), which ensures that \( \frac{\delta k}{\Delta t} = 0 \). Following [15], we demand that

\[ \frac{w(k, \delta k_i)}{w(k + \delta k_i, -\delta k_i)} = \frac{p_{k + \delta k_i}}{p_k}. \]

The transition rates calculated in (12) cancel, and on the RHS we can calculate from the probability distribution at thermal equilibrium

\[ \frac{p_{k + \delta k_i}}{p_k} = \exp \left[ -\delta k \beta \left( 2\Delta E + \sum_{j \neq i} J_{ji} (4k_j - 2N) \right) \right], \]

where \( \beta = \frac{1}{k_BT} \).

Putting all this together gives the coefficients as

\[ w(k, \delta k_i) = \begin{cases} \left( 1 + \gamma(\delta k) \right) \frac{\alpha \xi^{\delta k - 1}}{(\delta k - 1)!^2} & (\delta k > 0) \\ \prod_{m=1}^{\delta k} (k + m)(N - k - \delta k + m) & (\delta k < 0) \end{cases} \]

where

\[ \gamma(\delta k) = \tanh \left[ -\delta k \beta \left( \Delta N + \sum_{j \neq i} J_{ji} (2k_j - N) \right) \right]. \]

For a \( N \) boson, \( M \) site problem, the number of equations in (5) grows as \((N + 1)^M\), which quickly becomes intractable to solve directly. We therefore use a kinetic Monte Carlo (KMC) method to investigate multiple site systems with comparatively large particle numbers.

### 3 Kinetic Monte Carlo simulation

In the real BEC computing device, the bosons in each site make transitions from one state to another following a stochastic process as time advances. Their ensemble property is described by the master equation (5). For such a system with a stochastic transition behavior, KMC is an ideal method to find its dynamic nature. Here we give a detailed description of the algorithm used to generate our results.

We start the simulation from a random initial value of \( k = (k_1, k_2, ..., k_M) \) in (5), and update the system by repeating the stochastic transition process following the KMC method. In each update we calculate the transition weight \( w(k, \delta k_i) \) in (5) for all the possible transitions, and its cumulative function

\[ R_j = \sum_{i=1}^{n} \sum_{\delta k_i=-k_i}^{k_i} w(k, \delta k_i) \]

where we use the index \( j = (m - 1)N + (n + k_i) \), for \( m = 1, ..., M \) and \( n = -k_i, ..., N - k_i \) (all possible transitions), thus \( j \) ranges from 0 to \( MN \). We generate a uniform random number \( r \in (0, 1) \) to carry out the transition according to \( R_{j-1} < r < R_j \) where we simply define \( R_{-1} = 0 \). In this way the transitions with larger weights are more likely to be chosen.

To determine the time increment \( \Delta t \) between the transitions, a rejection-free approach is applied, where the time advance is not related with what transition is taken but only depends on the total transitions weight [17]. The \( \Delta t \) is assumed to be long enough that the system has no memory of how it entered to the current state, thus the transition weights are independent of what state preceded the current state. During each unit time every possible pathway has its own transition weight \( w(k, \delta k_i) \), and the total transitions weight \( W_{\text{total}} = R_{MN} \) is the sum of them. The system has the same probability of transition to another state during each increment of time, which gives a first-order process with exponential decay statistics. Thus the probability that the system remains in its current state is

\[ P_{\text{remain}}(t) = \exp(-W_{\text{total}}t). \]
Since
\[ 1 - P_{\text{remain}}(t') = \int_0^t P(t)dt, \quad (20) \]
from (19) and (20) we have the probability distribution function \( P(t) \) for the time of transition
\[ P(t) = W_{\text{total}} \exp(-W_{\text{total}}t). \quad (21) \]
The average time for transition \( t_{av} \) is the first moment of the distribution (21)
\[ t_{av} = \int_0^\infty tP(t)dt = \frac{1}{W_{\text{total}}}. \quad (22) \]
Therefore the time increment \( \Delta t \) randomly drawn according to the distribution (21) is calculated according to
\[ \Delta t = \frac{-1}{W_{\text{total}}} \ln(r) \quad (23) \]
where \( r \in (0, 1] \) is a randomly generated number and \(-\ln(r)\) forms an exponentially distributed random number.

The entire picture of the update process can be seen as a Markov chain, and the transition dynamics corresponds to a Markov walk. For one trajectory we set the simulator to repeat the transition until the system evolves to a certain running time, and by repeating a large number of the trajectories we acquire the ensemble result equivalent to the numerical solution of (5).

For each site \( i \), in the program we divide the running time into a certain number of divisions with equal lengths of time. Each division represents the \( \langle k_i \rangle \) in that time slice, and is calculated as
\[ \langle k_i(t) \rangle = \langle k_i(u\Delta T) \rangle = \sum_{x=1}^{N_{\text{tra}}(u)} k_i(t,x) \Delta T'(t,x) \frac{\Delta T_{\text{total}}}{N_{\text{tra}}}, \quad (24) \]
where \( t \in (u\Delta T, (u+1)\Delta T) \), \( u \) is the index of the time divisions, the summation is over all simulation trajectories, \( k_i(t,x) \) is the state of site \( i \) at time \( t \) of trajectory \( x \), \( \Delta T \) is the length of the time division, \( \Delta T'(t,x) \) is the time partial that in time \( t \) of trajectory \( x \) the state \( k_i \) was kept inside the division \( \Delta T \), and \( N_{\text{tra}} \) is the total number of the trajectories (see Fig. 3). A complete flow chart of the simulation is given in Fig. 4.

The simulator is implemented by a C program running on a Windows PC. For various small systems we have verified the KMC simulation results to be identical with numerically evolved results calculated in Mathematica.

An example of the output from the simulator is given in Fig. 5. The time evolution curve is fitted by a func-

---

**Fig. 3** Recording the states by time divisions in a KMC simulation.

**Fig. 4** KMC simulation flow chart.
Fig. 5 (a) Ensemble result by KMC simulation for site 1 of a 3 site 3 particle system. (b) One trajectory of the KMC simulation for the same system showing transitions between the levels on site 1.

\[ f(t) = a \exp(-t/t_0) + k_\text{equil} \]

where \( k_\text{equil} \) is the equilibrium value of \( \prod_{i=1}^{M} k_i \), and the equilibrium time \( \tau \) is defined as the time required for the fitted function to reach \( 0.9k_\text{equil} \).

We use the simulator to investigate systems in the parameter space of \((\beta, N, M)\) which are the inverse temperature, boson number, and site number respectively.

4 Simulations and results

First to find the dependency of equilibrium time \( \tau \) to the boson number \( N \), we analyzed a four site BEC system with Hamiltonian

\[ H = \sum_{\langle ij \rangle} J_{ij} S_i S_j + \lambda N \sum_{i} S_i \]  \( (25) \)

with \( J_{ij} = -10 \) and \( \lambda = -1 \). The Hamiltonian has a local minimum state \( \uparrow \uparrow \uparrow \uparrow \), and a global minimum state \( \downarrow \downarrow \downarrow \downarrow \). Fig. 6 is obtained by finding the equilibrium time \( \tau \) for site \( i \) and plotting this against the error for site \( i \) defined as

\[ \epsilon_i = 1 - \frac{1}{Z} \sum_{\{S, \sigma_i > 0\}} \exp[-H/k_BT] \]  \( (26) \)

where \( Z \) is the partition function, the summation is over all configurations with the same sign of spin as the ground state \( \sigma_i = -1 \) of the original Ising model problem. Fig. 6a shows the curves approach zero equilibration time as \( \propto 1/N \) for large \( N \). In a typical system implemented by exciton-polaritons, the \( N \) can be \( > 10^5 \). Therefore potentially very large speedups are possible using the bosonic systems with large particle numbers. Fig. 6b shows the average time between spin flips, which is defined to be a transition where any one of the spins \( S_i \) in the system changes its sign. The results show that as the boson number \( N \) is increased, the average time between spin flips decreases in proportion to \( N \). We attribute this as the reason why the equilibration times in Fig. 6a decrease with \( N \), even in the presence of local minima in the problem.

To find how the equilibrium time \( \tau \) changes with the number of sites \( M \), we analyzed the same configuration \( J_{ij} = -J = -10, \lambda = -1 \) for \( M = 2, 3, 4, 5, 6 \) systems. The problem for different \( M \) are equivalent in the following sense. For any \( M \) and \( N \), the Hamiltonian can be written as

Fig. 6 (a) \( 1/N \) dependence of the equilibration. (b) Average time between spin flips \( \tau_{\text{flip}} \) for the 4 site bosonic Ising model. All calculations use \( \alpha = 1 \) and \( \xi = 0.001 \).
Kinetic Monte Carlo study of accelerated optimization problem search using Bose-Einstein condensates

Fig. 7 (a) Transition paths combination from local minima to global minimum for a $M = 4$, $N = 1$ system. (b) The energy landscape for a $M = 4$, $N = 1$ system with fixed $J_{ij} = -10$ and $\lambda = -1$.

Fig. 8 Equilibrium time for $M = 2, 3, 4, 5, 6$ systems, all systems use the same configuration $J_{ij} = -10, \lambda = -1$.

5 Summary and conclusions

We have performed a detailed kinetic Monte Carlo study of the equilibration time required for a Ising model problem including a local minimum for various site numbers $M$ and boson numbers $N$. Although the system sizes that we consider in this report are much smaller than in a realistic setting, we do not foresee any fundamentally different effects to occur for larger system sizes, thus we expect that similar speedups should be possible with a larger number of lattice sites and boson numbers. From the KMC results, the order of the time complexity to solve Ising model problem on BEC system can be written in a form as

$$\tau \sim \exp(M) \frac{N}{M}$$

where $\tau$ is the time, $M$ is the number of sites, and $N$ is the particle number. The acceleration provided by the $N$ is only linear, since $N$ can in practice be a very large number (e.g. $> 10^5$) this scheme is promising towards a purpose built device for solving optimization problems that can be written in an Ising form.

The scheme described here can be combined with an
annealing schedule where the temperature can be gradually reduced towards $k_B T \to 0$. In practice this is more favorable than the constant temperature method that was used in the current simulations. Since the origin of the speedups is due to the increased spin flip times of the bosons on each site, we do not foresee any fundamental difference between using an annealing schedule to the constant temperature approach calculated here. In some preliminary results we have already observed speedups when combined with an annealing schedule, however, it not clear currently what kind of schedule is optimal for the bosonic case. We thus leave this component as future work.

The computational scheme described in this report is clearly different to the standard quantum computational model, where information is stored in a superposition state of quantum bits. The “quantum” nature of this computation occurs due to indistinguishability of the bosons, which is the origin of speedup due to FSS. In the sense that there is no coherence between the sites, as shown in (28) there is no difference to the classical case in terms of the scaling with the system size. A different scaling behavior may be possible if coherence is introduced between the sites, combining the bosonic speedup with quantum parallelism. Possible extensions along this line of thinking are currently being investigated. Another advantage of the present scheme is that it is generally applicable to a large class of problems, that can be formulated as a Hamiltonian to be optimized. Due to the current experimental difficulties of creating a quantum computer, examining alternative models of quantum computation such as the present scheme may open up new possibilities towards the realization of a practical device.

Acknowledgments
This work is supported by the Special Coordination Funds for Promoting Science and Technology, Navy/SPAWAR Grant N66001-09-1-2024, MEXT, and the JSPS through its FIRST program.

References
Kinetic Monte Carlo study of accelerated optimization problem search using Bose-Einstein condensates


Kai YAN
Kai YAN received a B.S. degree in Information and Communication Engineering in 2008 and a M.S. in Information Science and Technology in 2010 from the University of Tokyo. He is currently pursuing a Ph.D. degree at the University of Tokyo, where he does theoretical research in the field of Quantum Information under the supervision of Professor Yamamoto.

Tim BYRNES
Tim BYRNES completed his Ph.D. from the University of New South Wales in 2003. In 2004, he moved to Japan to commence a postdoctoral fellowship at the University of Tokyo and the National Institute of Informatics. He is currently assistant professor at the National Institute of Informatics. His current interests are in the fields of quantum simulation, quantum computation, and condensed matter physics.

Yoshihisa YAMAMOTO
Prof. Yoshihisa YAMAMOTO is working at E.L. Ginzton Laboratory of Stanford University, Stanford, CA, USA and National Institute of Informatics, Tokyo, Japan. His group is primarily concerned with studying quantum dynamics of spin and exciton in semiconductor nanoscale systems. His experimental project combines ultrafast optical spectroscopy with semiconductor microcavities with artificial atoms (quantum dots or impurities) at low temperatures and high magnetic fields. Main focus of his current effort is physical implementation of quantum information processing systems and includes the following research programs: i) Optically controlled semiconductor spin qubits for quantum information processing. ii) Dynamic condensation of exciton-polaritons for quantum computation and simulation.