

# Optimization by a quantum reinforcement algorithm

A. Ramezanpour\*

*Department of Physics, University of Neyshabur, Neyshabur, 9319774400, Iran  
and Leiden Academic Centre for Drug Research, Faculty of Mathematics and Natural Sciences, Leiden University,  
Leiden, The Netherlands*

(Received 11 June 2017; revised manuscript received 4 August 2017; published 3 November 2017)

A reinforcement algorithm solves a classical optimization problem by introducing a feedback to the system, which slowly changes the energy landscape and converges the algorithm to an optimal solution in the configuration space. Here, we use this strategy to concentrate (localize) the wave function of a quantum particle, which explores the configuration space of the problem, preferentially on an optimal configuration. We examine the method by solving numerically the equations governing the evolution of the system, which are similar to the nonlinear Schrödinger equations, for small problem sizes. In particular, we observe that reinforcement increases the minimal energy gap of the system in a quantum annealing algorithm. Our numerical simulations and the latter observation show that such kind of quantum feedback might be helpful in solving a computationally hard optimization problem by a quantum reinforcement algorithm.

DOI: [10.1103/PhysRevA.96.052307](https://doi.org/10.1103/PhysRevA.96.052307)

## I. INTRODUCTION

Reinforcement is a very useful technique in machine learning and optimization algorithms for the study of computationally hard optimization problems [1,2]. The main idea is based on rewarding good decisions or modifying the energy landscape in a way that leads the algorithm to an optimal solution. This is usually done by introducing an appropriate feedback to the system, which depends on the information provided by the algorithm, to guide the optimization process. In this paper, we study a quantum reinforcement algorithm, which employs a continuous-time quantum random walk to explore the configuration space of an optimization problem. We show that such kind of quantum feedback can converge the quantum particle towards a solution by a preferential localization of the wave function in the solution space.

Consider the problem of finding a solution to a classical optimization problem, identified by a probability distribution over the configuration space of the problem variables. We assume that the probability distribution is nonzero only for (a possibly large number of) configurations in the subspace of solutions. Then a decimation algorithm to find a solution works by fixing the value of a randomly chosen variable according to the marginal probability of that variable. The algorithm continues until the value of every variable is fixed. The marginal probabilities at each step are obtained by an approximate sampling algorithm, e.g., Monte Carlo, conditioned on the values of the already decimated variables. Instead of fixing the variables one by one, a reinforcement optimization algorithm modifies smoothly the joint probability distribution of the variables by changing slowly the values of some external local fields acting on the variables [3]. These local fields use the estimated marginal probabilities of the variables to concentrate the joint probability distribution more and more on a single configuration in the solution space.

On the other hand, a quantum algorithm exploits the computational power of a quantum system to solve a computationally difficult problem [4–7]. Quantum random

walks [8–13] and adiabatic quantum computation [14,15] are important examples of quantum approaches to universal computations [16,17]. Specifically, we should mention recent efforts in constructing effective shortcuts to adiabaticity [18–20], nonunitary evolution of quantum random walks and non-Hermitian quantum annealing [21,22], and investigations of quantum annealing with nonstoquastic Hamiltonians [23–25]. Another related study is the quantum reinforcement learning algorithm [26], which is a quantum implementation of the reinforcement learning algorithm.

The wave function of a quantum particle in the complex energy landscape of an optimization problem can undergo a localization transition, which may limit the efficiency of a quantum annealing algorithm [27,28]. Here, we propose a quantum reinforcement algorithm, which works by concentrating the wave function preferentially on the subspace of optimal configurations. The algorithm exploits the information provided by the instantaneous wave function of the system, or expectation values of some local observables, to steer the evolution of the quantum system. In addition, we show that such a quantum reinforcement can increase the minimum energy gap that the system encounters in a quantum annealing algorithm.

It is known that a nonlinear quantum mechanic can be exploited by a quantum computer to solve a computationally hard problem in a polynomial time [29]; this does not mean that quantum mechanics is nonlinear in nature, or any nonlinearity in the time evolution of the quantum system is computationally beneficial. Here, we show that a kind of nonlinear quantum evolution inspired from the classical reinforcement algorithms can be used to increase the energy gap and speedup the computation compared with the conventional quantum annealing algorithm.

## II. MAIN DEFINITIONS

We consider the classical optimization problem of minimizing an energy function  $E(\sigma)$  of  $N$  binary spins  $\sigma_i = \pm 1$ . To be specific, as the benchmark we take a (fully connected) random spin model, with  $E(\sigma) = -\sum_{i<j} J_{ij}\sigma_i\sigma_j$ . The couplings  $J_{ij}$  are independent Gaussian random variables of mean zero and variance  $1/N$ . The scaling is chosen to have an extensive

\*aramezanpour@gmail.com

energy of order  $N$ . This model is known as the Sherrington-Kirkpatrick (SK) model [30],

We shall use a continuous-time quantum random walk to explore the space of spin configurations  $\sigma = \{\sigma_1, \dots, \sigma_N\}$ . The space is a hypercube of  $2^N$  sites corresponding to the total number of spin configurations. The Hamiltonian of the particle in the energy landscape of the classical optimization problem is given by

$$H = \sum_{\sigma} E(\sigma) |\sigma\rangle \langle \sigma| - \sum_{\sigma} \sum_{i=1}^N \Gamma (|\sigma^{-i}\rangle \langle \sigma| + |\sigma\rangle \langle \sigma^{-i}|). \quad (1)$$

The parameter  $\Gamma$  determines the strength of tunneling from  $|\sigma\rangle$  to a neighboring state  $|\sigma^{-i}\rangle$ . Here  $|\sigma^{-i}\rangle$  denotes the spin state, which is different from  $|\sigma\rangle$  only at site  $i$ . In terms of the quantum spin variables (Pauli matrices), the above Hamiltonian reads as  $H = -\sum_{i<j} J_{ij} \sigma_i^z \sigma_j^z - \sum_i \Gamma \sigma_i^x$ . Additionally, the basis states  $|\sigma\rangle$  are the  $N$ -spin states with definite  $\sigma_i^z$  values, that is  $\sigma_i^z |\sigma\rangle = \sigma_i |\sigma\rangle$ . Starting from an initial state  $|\psi(0)\rangle$ , time evolution of the system is governed by the Schrödinger equation,  $i \frac{d}{dt} |\psi(t)\rangle = H |\psi(t)\rangle$  with  $\hbar = 1$ .

### III. QUANTUM REINFORCEMENT ALGORITHM

The goal here is to find a solution to the classical optimization problem by following the time evolution of a quantum system. A quantum annealing (QA) algorithm [14] starts from the ground state of  $H_x \equiv -\sum_i \Gamma \sigma_i^x$  and changes slowly the Hamiltonian to  $H_c \equiv -\sum_{i<j} J_{ij} \sigma_i^z \sigma_j^z$ . The adiabatic theorem then ensures that in the absence of level crossing, the system follows the instantaneous ground state of the time-dependent Hamiltonian  $H_{QA}(t) = s(t)H_c + [1 - s(t)]H_x$ . The annealing parameter  $s(t)$  changes slowly from zero at  $t = 0$ , to one at  $t = T$ . For instance, in a linear annealing schedule  $s(t) = t/T$ .

In a quantum reinforcement (QR) algorithm, we add a reinforcement term to the Hamiltonian, which favors the spin states of higher probability. More precisely, the Hamiltonian is  $H_{QR}(t) = H_c + H_x + H_r(t)$ , where the reinforcement term reads as follows:

$$H_r(t) \equiv -r(t) \sum_{\sigma} |\psi(\sigma; t)|^2 |\sigma\rangle \langle \sigma|. \quad (2)$$

The reinforcement parameter  $r(t)$  is zero at the beginning and grows slowly in magnitude with time. In a linear reinforcement schedule we take  $r(t) = (t/T)2^N r_0$ . In other words, as the time passes, the on-site energy at state  $|\sigma\rangle$  decreases with an amount that is proportional to the probability of finding the walker at that site  $|\psi(\sigma; t)|^2$ . This probability could be exponentially small at the beginning of the process. That is why here we scale the reinforcement parameter with  $2^N$ . To have an extensive Hamiltonian for  $N \rightarrow \infty$ , one can also change the scaling with time such that  $r(t) \propto N$  at the end of the process, where the wave function is of order one.

For comparison with the QA algorithm, we also study a reinforced quantum annealing (rQA), where

$$H_{rQA}(t) \equiv s(t)[H_c + H_r(t)] + [1 - s(t)]H_x. \quad (3)$$

This allows us to examine the effect of reinforcement on the behavior of the quantum annealing algorithm. The reinforcement parameter here is a constant  $r(t) = Nr_0$ . So, the time

dependence of  $H_r(t)$  is determined by  $\psi(\sigma; t)$ , which can safely be replaced by the instantaneous ground state of the system for an adiabatic process.

To obtain a local version of the above algorithms, we first replace the  $|\psi(\sigma; t)|^2$  with  $\ln |\psi(\sigma; t)|^2$ , which is an increasing function of the probability distribution. Note that we can always write  $|\psi(\sigma; t)|^2 = \exp(\sum_i K_i \sigma_i + \sum_{i<j} K_{ij} \sigma_i \sigma_j + \dots)/Z$ , taking into account all the possible multispin interactions;  $Z$  is the normalization constant. The coupling parameters  $K_i, K_{ij}, \dots$  in principle can be determined from the expectation values  $\langle \sigma_i^z \rangle, \langle \sigma_i^z \sigma_j^z \rangle, \dots$ . A local quantum reinforcement (lQR) algorithm then is obtained by approximating the wave function with a product state,

$$H_r^{\text{local}}(t) \equiv -r(t) \sum_{\sigma} \sum_i K_i \sigma_i |\sigma\rangle \langle \sigma|. \quad (4)$$

The reinforcement fields  $K_i$  depend on the average spin values  $m_i = \sum_{\sigma} \sigma_i |\psi(\sigma; t)|^2$  through  $K_i = \frac{1}{2} \ln[(1 + m_i)/(1 - m_i)]$ . Here we increase the reinforcement parameter with time as  $r(t) = r_0 t$ . More accurate approximations of the wave function and the local quantum reinforcement algorithm can be obtained by considering the two-spin and the higher-order interactions in the expansion.

### IV. RESULTS AND DISCUSSION

Figure 1 shows the success probability  $\sum_{\sigma_{\text{opt}}} |\langle \sigma_{\text{opt}} | \psi(t) \rangle|^2$  of the above algorithms for the SK model. The initial state  $|\psi(0)\rangle$  is the equal superposition of spin states  $|x\rangle = \frac{1}{\sqrt{2^N}} \sum_{\sigma} |\sigma\rangle$ , or the all-positive state  $|+\rangle = |+\dots+\rangle$ . In each case we tried different values of the parameters to obtain roughly the best performances. As expected for a quantum random walk, the algorithms are sensitive to the initial state of the system [11]. We see from the figure that the QR algorithms can localize a considerable fraction of the wave function on the optimal spin configurations. Nevertheless, the performance of these algorithms degrades by increasing the number of spins (for  $3 < N < 9$ ). We know that the success probability of the QA algorithm decreases exponentially with  $N$  because of the exponentially small energy gaps of the Hamiltonian [27]. On the other hand, we know from Ref. [29] that nonlinear quantum evolution could be helpful; therefore, it would be interesting to see how the success probability of the QR algorithm scales with the number of spins for larger systems.

Figure 1 also shows the percentile values of the success probability for some independent realizations of the problem; we see that there are always a finite fraction of the problem instances for which the success probability of the QR algorithm is nearly zero; we can indeed enhance the success probability for these instances by slightly changing the algorithm parameters. In Fig. 2(a) we compare the success probability of the QR algorithm with that of the QA algorithm for a number of independent problem instances. We observe that the QR algorithm displays large sample to sample fluctuations, with very good or very bad performances compared to the QA algorithm.

The good point with the QR algorithms is that we do not need the ground state of the Hamiltonian to initialize these algorithms. Therefore, one can restart the algorithm at any stage of the evolution with an arbitrary wave function, or a spin

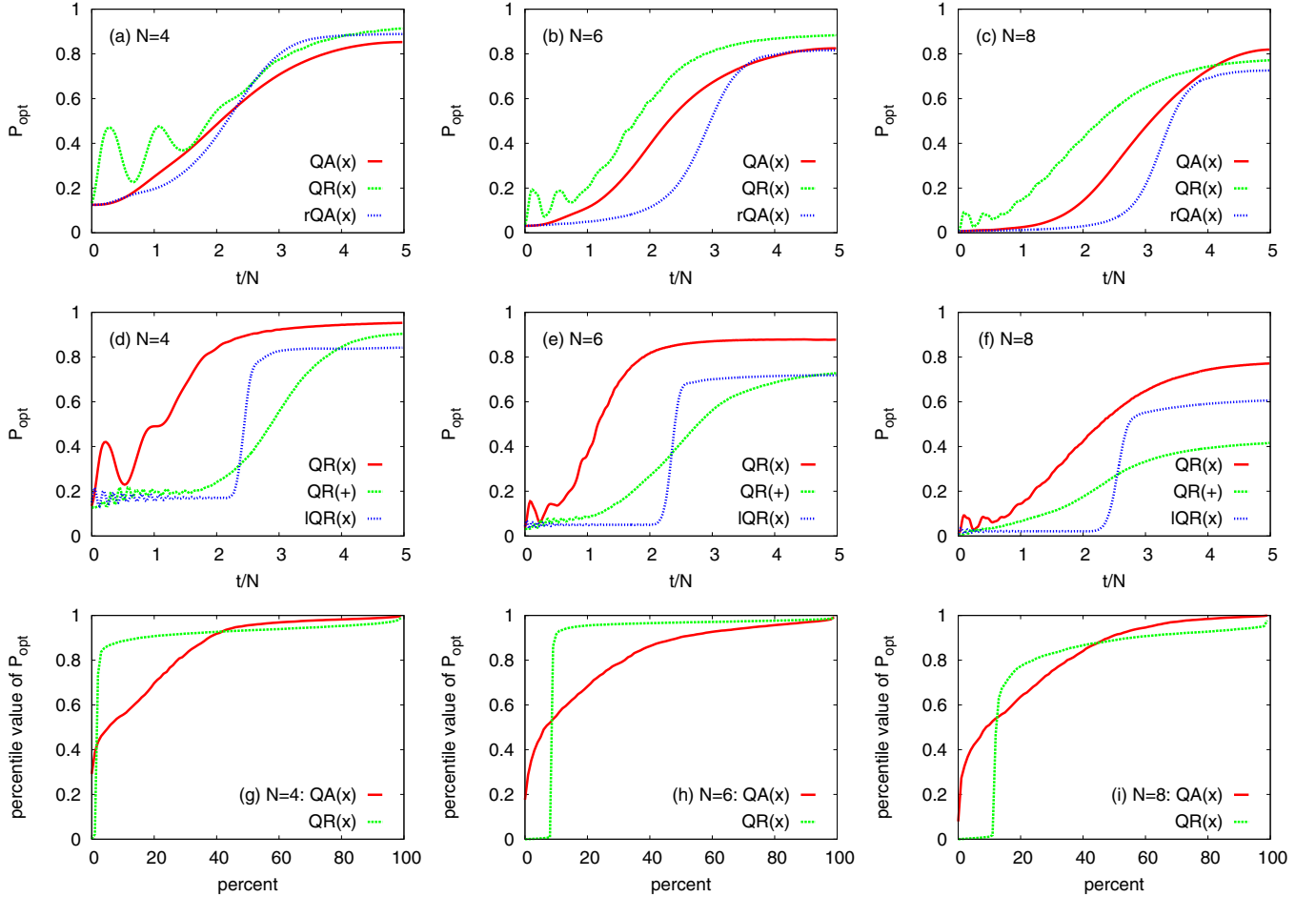


FIG. 1. Success probability  $P_{\text{opt}}$  of the algorithms for different number of spins ( $N$ ) in the SK model. (a)–(c) show the results obtained by the quantum annealing (QA), quantum reinforcement (QR), and reinforced quantum annealing (rQA) algorithms. (d)–(f) compare the results obtained by the QR algorithm for different initial conditions and the local quantum reinforcement (IQR) algorithm. (g)–(i) display the percentile values of the success probability at the end of the process ( $t = T$ ). The initial states are indicated by ( $\times$ ) for the equal superposition of all spin configurations, and (+) for the all-positive spin configuration. The data are results of 2000 independent realizations of the random spin model. The statistical errors are about 0.01. For the QA( $\times$ ) we used  $\Gamma = 0.3$  (a), (b), (g), (h) and  $\Gamma = 0.5$  (c), (i). For the QR( $\times$ ) we used  $\Gamma = 0.8, r_0 = 1$  (a), (g),  $\Gamma = 1.2, r_0 = 1$  (b), (h),  $\Gamma = 1.3, r_0 = 0.5$  (c), (f), (i),  $\Gamma = 1, r_0 = 2.5$  (d), and  $\Gamma = 1.5, r_0 = 2.5$  (e). For the rQA( $\times$ ) we used  $\Gamma = 0.5, r_0 = 1$  (a),  $\Gamma = 0.9, r_0 = 1$  (b), and  $\Gamma = 1.2, r_0 = 1$  (c). For the QR(+) we used  $\Gamma = 3, r_0 = 3.5$  (d),  $\Gamma = 2.5, r_0 = 2.5$  (e), and  $\Gamma = 2.5, r_0 = 2.0$  (f). For the IQR( $\times$ ) we used  $\Gamma = 3, r_0 = 1$  (d),  $\Gamma = 3.5, r_0 = 0.5$  (e), and  $\Gamma = 2.5, r_0 = 0.2$  (f).

configuration, which is sampled from the wave function of the system at that moment (see also Refs. [31,32]). Figure 2(b) gives the success probability of a two-stage QR algorithm where the final wave function of the first stage is taken for the initial state in the second stage of the algorithm. In this study we used simple reinforcement schedules [ $r(t) \propto t$  and  $\Gamma(t) = \text{const.}$ ], which is not necessarily the best way of exploiting the reinforcement; in general, the reinforcement parameter is expected to grow at the beginning and the tunneling may diminish at the end of the process. Figure 2(b) shows also the success probability of the one-stage QR algorithm for the case in which the reinforcement parameter increases with time at the beginning of the process and then goes to zero along with the hopping parameter  $\Gamma$ . This is to demonstrate that we do not need an exponentially large reinforcement when the wave function is of order one.

To see what happens close to a level crossing, we consider a two-level system with energy function  $E(\sigma) = h\sigma$ , where  $\sigma =$

$\pm 1$ . For the initial state we take  $|\psi(0)\rangle = (|-\rangle + |+\rangle)/\sqrt{2}$ . In the QR algorithm, the evolution is governed by the following Hamiltonian:

$$H_{QR}^{(2)}(t) = \begin{pmatrix} h - r(t)|\psi(+;t)|^2 & -\Gamma \\ -\Gamma & -h - r(t)|\psi(-;t)|^2 \end{pmatrix}. \quad (5)$$

Figure 3 shows the time dependence of the wave function and the effective (reinforced) energies  $E(\sigma; t) \equiv h\sigma - r(t)|\psi(\sigma; t)|^2$ . We see that the energy landscape is favoring more and more the ground state of the system as the time passes. Notice also the oscillations in the wave function and the effective energies; these are reminiscent of the oscillations that are observed in the amplitude amplification algorithm in search problems [33].

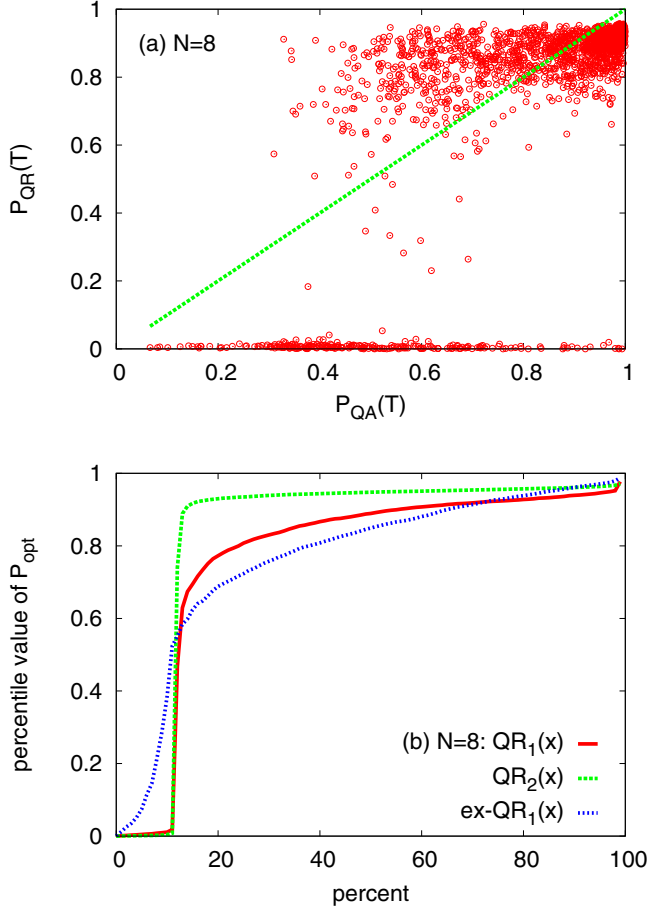


FIG. 2. Success probabilities  $P_{QR}, P_{QA}$  of the QR and QA algorithms in the SK model. (a)  $P_{QR}(T)$  vs  $P_{QA}(T)$  at the end of the process  $t = T$  for 2000 independent realizations of the problem. The parameters and initial conditions are similar to the ones in Fig. 1(c). (b) Comparing the percentile values of the success probability in the one- and two-stage quantum reinforcement algorithms ( $QR_1, QR_2$ ) starting from the equal superposition of all the spin states. The parameters in the two stages are  $r_1(t) = 0.5(t/T)2^N, \Gamma_1 = 1.3$  and  $r_2(t) = r_1(t) + 0.15, \Gamma_2 = \Gamma_1$  with  $T = 5N$  for each stage. Here ex- $QR_1$  denotes the one-stage QR algorithm with parameters  $r(t) = 0.5(t/T)(1 - t/T)2^N, \Gamma(t) = 2(1 - t/T)$ , which go to zero at the end of the process. The data are results of 2000 independent realizations of the random spin model.

Next, we consider the two-level system with a reinforced Hamiltonian

$$H_R^{(2)} = \begin{pmatrix} h - r|\psi_0(+)|^2 & -\Gamma \\ -\Gamma & -h - r|\psi_0(-)|^2 \end{pmatrix}, \quad (6)$$

where  $\psi_0(\sigma)$  is the ground state. In Fig. 4(a) we see the energy gap of the system for different values of the parameters. The eigenvalues and eigenstates of the Hamiltonian have been obtained numerically by an iterative algorithm. In Fig. 4(b), we compare the energy gap of the above reinforced Hamiltonian with that of a local reinforced Hamiltonian, where  $|\psi_0(\sigma)|^2$  is replaced with  $K\sigma$ . As before, the coupling is given by  $K = \frac{1}{2} \ln[(1+m)/(1-m)]$ , where  $m = |\psi_0(+)|^2 - |\psi_0(-)|^2$  is the magnetization. In the reinforced quantum annealing (rQA), the Hamiltonian of the two-level system reads as

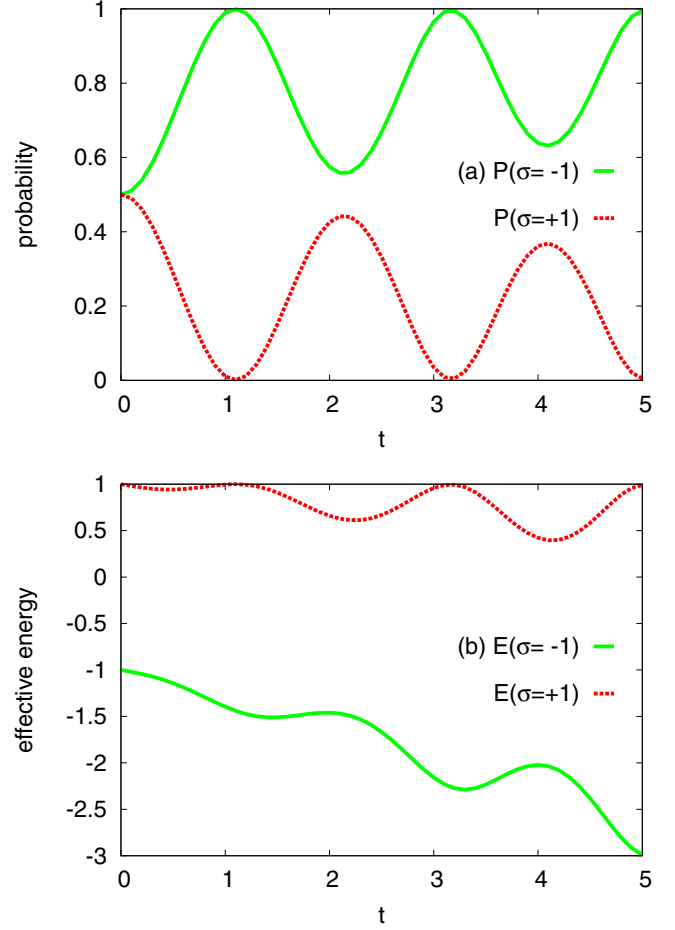


FIG. 3. Time evolution of the two-level system in the quantum reinforcement algorithm. (a) The probabilities  $|\psi(-;t)|^2$  and  $|\psi(+;t)|^2$ , and (b) the effective energy  $E(\sigma;t) = h\sigma - r(t)|\psi(\sigma;t)|^2$ . The parameters are  $h = \Gamma = 1$ , and  $r(t) = 2(t/T)$  with  $T = 5$ .

follows:

$$H_{rQA}^{(2)}(t) = \begin{pmatrix} (t/T)[h - r|\psi_0(+;t)|^2] & -(1-t/T)\Gamma \\ -(1-t/T)\Gamma & (t/T)[-h - r|\psi_0(-;t)|^2] \end{pmatrix}. \quad (7)$$

Figure 4(c) displays the time dependence of the energy gap for the quantum annealing and the (local) reinforced quantum annealing algorithms. As the figure shows, in both the cases the reinforcement tends to increase the energy gap of the two-level system.

In Fig. 5 we check the above observation for larger random spin systems. Here, for the classical problem we take the SK model with random fields,  $E(\sigma) = -\sum_{i<j} J_{ij}\sigma_i\sigma_j - \sum_i h_i\sigma_i$ . The additional interactions with the external fields ensure that the ground state is not degenerate. The fields  $h_i$  are independent Gaussian random variables of mean zero and variance one. In the numerical simulations we assume the system follows the instantaneous ground state of the Hamiltonian  $H_{rQA}(t)$ . As the figure shows, the reinforcement increases the energy gap and can be useful as another heuristic

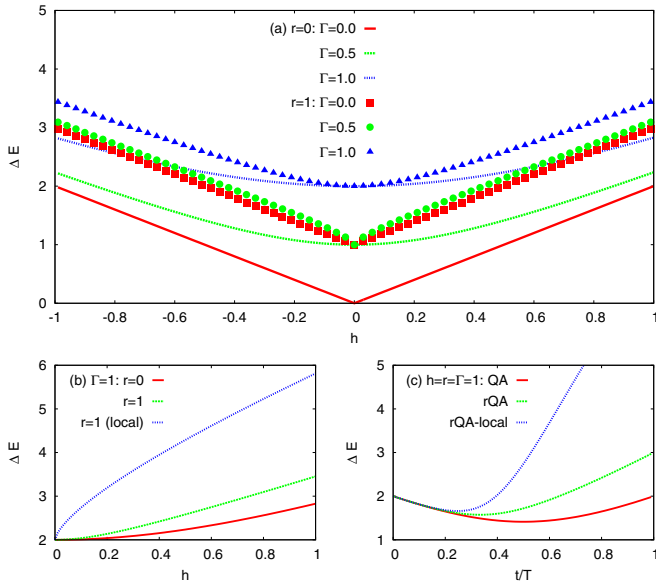


FIG. 4. (a) The energy gap of the two-level system  $\Delta E(h)$  for different values of  $\Gamma$  and  $r$  in the reinforced Hamiltonian  $H_R^{(2)}$ . (b) Energy gap of the local reinforced Hamiltonian for given values of  $\Gamma$  and  $r$ . (c) Time dependence of the energy gap  $\Delta E(t)$  for given values of  $h, \Gamma$ , and  $r$  in the quantum annealing (QA), reinforced quantum annealing (rQA), and the local version of reinforced quantum annealing (rQA-local).

algorithm to improve the efficiency of the conventional quantum annealing algorithm.

V. CONCLUSION

In summary, our numerical simulations of random spin systems show that quantum reinforcement algorithms might be useful in solving a computationally expensive optimization problem. Clearly, more studies are required to see how this strategy works in larger problem sizes. In particular, the local version of the quantum reinforcement algorithm (introduced in Sec. III) can be studied by a quantum Monte Carlo algorithm for simulation of an open quantum system at sufficiently small temperatures, to examine the method for larger systems. Another challenge lies in the experimental realization of such quantum feedbacks in practice. Recent advances in quantum control theory [34,35], e.g., the concept of continuous measurement of a quantum system, could be helpful in this direction. A naive approach is to approximate the nonlinear evolution of the quantum system by a sequence of estimations of the quantum state followed by linear evolutions of the quantum state [36]. The reinforcement terms in the Hamiltonian are updated only in the estimation stage, the time evolution of the system is then controlled by the updated Hamiltonian in the evolution stage. Specifically, in the case of the local QR algorithm, we need to engineer a local Hamiltonian with couplings that can be estimated from a weak measurement of  $N$  commuting spin variables, perhaps via interaction with some ancillary spins.

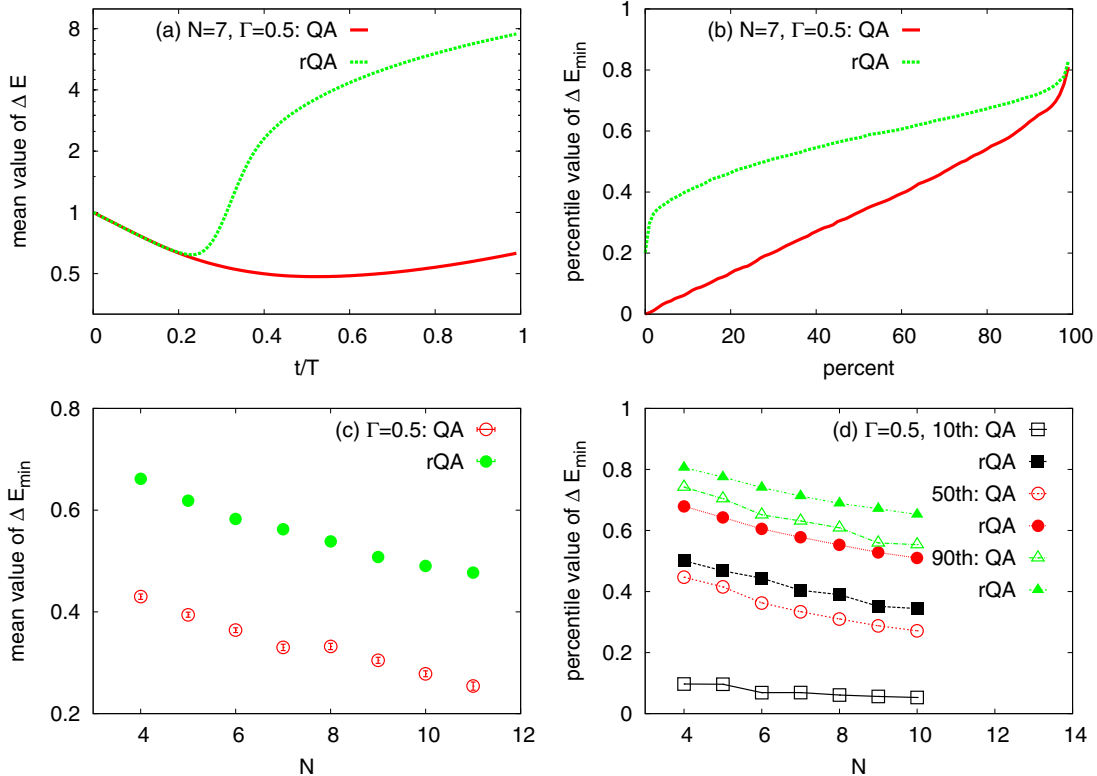


FIG. 5. The energy gap of the SK model in presence of random fields in the quantum annealing (QA) and reinforced quantum annealing (rQA) with  $r_0 = 1$ . (a) The average energy gap  $\Delta E(t)$  vs the evolution time, and (b) the percentile values of the minimum energy gap  $\Delta E_{\min}$ , for a given number of spins  $N$  and  $\Gamma$ . (c) The average of the minimal energy gap encountered in the annealing process, and (d) the percentile values of  $\Delta E_{\min}$ , vs the number of spins. The data are results of 2000 independent realizations of the random model.

- [1] Richard S. Sutton and Andrew G. Barto, *Introduction to Reinforcement Learning* (MIT Press, Cambridge, 1998), Vol. 135.
- [2] M. Mezard and A. Montanari, *Information, Physics, and Computation* (Oxford University Press, Oxford, 2009).
- [3] A. Braunstein and R. Zecchina, Learning by Message Passing in Networks of Discrete Synapses, *Phys. Rev. Lett.* **96**, 030201 (2006).
- [4] Michael A. Nielsen and I. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2002).
- [5] A. Yu Kitaev, A. Shen, and Mikhail N. Vyalyi, *Classical and Quantum Computation* (American Mathematical Society, Providence, 2002), Vol. 47.
- [6] P. W. Shor, Algorithms for quantum computation: Discrete logarithms and factoring, in *Proceedings of the 35th Annual Symposium on Foundations of Computer Science* (IEEE, Piscataway, 1994).
- [7] L. K. Grover, Quantum mechanics Helps in Searching for a Needle in a Haystack, *Phys. Rev. Lett.* **79**, 325 (1997).
- [8] Y. Aharonov, L. Davidovich, and N. Zagury, Quantum random walks, *Phys. Rev. A* **48**, 1687 (1993).
- [9] E. Farhi and S. Gutmann, Quantum computation and decision trees, *Phys. Rev. A* **58**, 915 (1998).
- [10] N. Shenvi, J. Kempe, and K. Birgitta Whaley, Quantum random-walk search algorithm, *Phys. Rev. A* **67**, 052307 (2003).
- [11] J. Kempe, Quantum random walks: An introductory overview, *Contemp. Phys.* **44**, 307 (2003).
- [12] A. Ambainis, Quantum walks and their algorithmic applications, *Int. J. Quantum. Inform.* **1**, 507 (2003).
- [13] M. Szegedy, Quantum speed-up of Markov chain based algorithms, in *Proceedings of the 45th Annual IEEE Symposium on Foundations of Computer Science* (IEEE, Piscataway, 2004).
- [14] E. Farhi *et al.*, A quantum adiabatic evolution algorithm applied to random instances of an NP-complete problem, *Science* **292**, 472 (2001).
- [15] W. M. Kaminsky and S. Lloyd, “Scalable Architecture for Adiabatic Quantum Computing of NP-Hard Problems.” *Quantum Computing and Quantum Bits in Mesoscopic Systems* (Springer, New York, 2004), pp. 229–236.
- [16] A. M. Childs, Universal Computation by Quantum Walk, *Phys. Rev. Lett.* **102**, 180501 (2009).
- [17] D. Aharonov *et al.*, Adiabatic quantum computation is equivalent to standard quantum computation, *SIAM Rev.* **50**, 755 (2008).
- [18] M. V. Berry, Transitionless quantum driving, *J. Phys. A: Math. Theor.* **42**, 365303 (2009).
- [19] A. T. Rezakhani *et al.*, Quantum Adiabatic Brachistochrone, *Phys. Rev. Lett.* **103**, 080502 (2009).
- [20] A. del Campo, Shortcuts to Adiabaticity by Counterdiabatic Driving, *Phys. Rev. Lett.* **111**, 100502 (2013).
- [21] M. Mohseni *et al.*, Environment-assisted quantum walks in photosynthetic energy transfer, *J. Chem. Phys.* **129**, 11B603 (2008).
- [22] A. I. Nesterov and G. P. Berman, Quantum search using non-Hermitian adiabatic evolution, *Phys. Rev. A* **86**, 052316 (2012).
- [23] S. V. Isakov *et al.*, Understanding Quantum Tunneling through Quantum Monte Carlo Simulations, *Phys. Rev. Lett.* **117**, 180402 (2016).
- [24] L. Hormozi *et al.*, Nonstoquastic Hamiltonians and quantum annealing of an Ising spin glass, *Phys. Rev. B* **95**, 184416 (2017).
- [25] Y. Susa, J. F. Jadebeck, and H. Nishimori, Relation between quantum fluctuations and the performance enhancement of quantum annealing in a nonstoquastic Hamiltonian, *Phys. Rev. A* **95**, 042321 (2017).
- [26] D. Dong *et al.*, Quantum reinforcement learning, *IEEE T. Syst. Man. Cyb.* **38**, 1207 (2008).
- [27] B. Altshuler, H. Krovi, and J. Roland, Anderson localization makes adiabatic quantum optimization fail, *Proc. Natl. Acad. Sci. USA* **107**, 12446 (2010).
- [28] C. L. Baldwin, C. R. Laumann, A. Pal, and A. Scardicchio, The many-body localized phase of the quantum random energy model, *Phys. Rev. B* **93**, 024202 (2016).
- [29] D. S. Abrams and S. Lloyd, Nonlinear Quantum Mechanics Implies Polynomial-Time Solution for NP-Complete and # P Problems, *Phys. Rev. Lett.* **81**, 3992 (1998).
- [30] D. Sherrington and S. Kirkpatrick, Solvable Model of a Spin-Glass, *Phys. Rev. Lett.* **35**, 1792 (1975).
- [31] A. Perdomo-Ortiz, Salvador E. Venegas-Andraca, and A. Aspuru-Guzik, A study of heuristic guesses for adiabatic quantum computation, *Quant. Info. Proc.* **10**, 33 (2011).
- [32] N. Chancellor, Modernizing quantum annealing using local searches, *New J. Phys.* **19**, 023024 (2017).
- [33] G. Brassard *et al.*, Quantum Amplitude Amplification and Estimation, *Contemp. Math.* **305**, 53 (2002).
- [34] A. C. Doherty and K. Jacobs, Feedback control of quantum systems using continuous state estimation, *Phys. Rev. A* **60**, 2700 (1999).
- [35] G. A. Smith *et al.*, Efficient Quantum-State Estimation by Continuous Weak Measurement and Dynamical Control, *Phys. Rev. Lett.* **97**, 180403 (2006).
- [36] H. M. Wiseman and G. J. Milburn, *Quantum Measurement and Control* (Cambridge University Press, Cambridge, 2009).