

Annealing by simulating the coherent Ising machine

EGOR S. TIUNOV,^{1,2} ALEXANDER E. ULANOV,¹ AND A. I. LVOVSKY^{1,3,4,*}

¹*Russian Quantum Center, 100 Novaya St., Skolkovo, Moscow, 143025, Russia*

²*Moscow Institute of Physics and Technology, Dolgoprudny, 141700, Russia*

³*Department of Physics, University of Oxford, Oxford, OX1 3PU, UK*

⁴*P. N. Lebedev Physics Institute, Leninskiy prospect 53, Moscow, 119991, Russia*

**Alex.Lvovsky@physics.ox.ac.uk*

Abstract: The coherent Ising machine (CIM) enables efficient sampling of low-lying energy states of the Ising Hamiltonian with all-to-all connectivity by encoding the spins in the amplitudes of pulsed modes in an optical parametric oscillator (OPO). The interaction between the pulses is realized by means of measurement-based optoelectronic feedforward, which enhances the gain for lower-energy spin configurations. We present an efficient method of simulating the CIM on a classical computer that outperforms the CIM itself, as well as the noisy mean-field annealer in terms of both the quality of the samples and the computational speed. It is furthermore advantageous with respect to the CIM in that it can handle Ising Hamiltonians with arbitrary real-valued node coupling strengths. These results illuminate the nature of the faster performance exhibited by the CIM and may give rise to a new class of quantum-inspired algorithms of classical annealing that can successfully compete with existing methods.

© 2019 Optical Society of America under the terms of the [OSA Open Access Publishing Agreement](#)

1. Introduction

The Ising model, originally developed for the description of phase transitions in magnetic materials [1], is a pillar of many branches in science. Its applications range from quantum field theory [2] and quantum gravity [3] to economics [4] and machine learning [5]. The model is defined by the Hamiltonian

$$H = -\frac{1}{2} \sum_{i,j} J_{ij} \sigma_i \sigma_j \quad (1)$$

where $J_{ij} = J_{ji}$, $J_{ii} = 0$. The spin variable σ_i can be a quantum operator or a number, continuous or discrete. In this paper we are interested in the classical discrete case, $\sigma_i = \pm 1$. The goal is to find the configurations of spins which realize or approach the global minimum of the Ising Hamiltonian (1). This is an NP-hard problem, meaning that the number of operations needed to find the exact solution grows exponentially with the size of the spin system [6]. However, there exists a variety of classical algorithms to find an approximate solution [7–10].

Recently, the Ising problem has been approached using noisy intermediate-scale quantum (NISQ) devices. One example is the D-Wave quantum annealer based on superconducting qubits. However, it has a shortcoming associated with low connectivity: each physical qubit in the D-Wave chimera graph architecture is connected to only eight others. Therefore each logical spin must be embedded in multiple physical qubits to model the fully connected Hamiltonian (1), resulting in a significant overhead [11].

Another example of NISQ technology is the coherent Ising machine (CIM) — an optoelectronic device developed in 2016 [12, 13]. Experiments seem to demonstrate that CIM outperforms not only classical simulated annealing algorithms [14], but also the D-Wave annealer [15]; however, the latter result is being disputed by the D-Wave team [16].

It has been argued that the fast performance of the CIM is of quantum nature, arising thanks to the nonclassical nature of optical squeezing [17]. This view, however, challenges the universally accepted belief that, in order to exhibit quantum speedup, a computational device would require large-scale entanglement and thoroughly minimized interaction with the environment, as CIM does not possess either of these features. A reasonable alternative hypothesis would be to ascribe the speedup to the analog optical processing that takes place in the CIM. Indeed, it is known that such processing, even in the absence of quantum effects, can be used to parallelize computational operations [18].

To address these questions, a number of approaches to computer simulation of the CIM and related physics have been developed [15, 19–23]. In particular, a recent study [21] made an explicit comparison of a mean-field digital annealer and observed advantage of their algorithm with respect to the CIM in terms of both the annealing time and quality of samples obtained. In the present work, we develop a new simulator, which we call SimCIM because it is largely based on the actual physics of the CIM. We compare its performance with the *bona fide* CIM and the annealer of [21], and find the results of this comparison to be generally in favor of our algorithm.

2. The coherent Ising machine

The primary element of the CIM architecture (Fig. 1) is an optical parametric oscillator (OPO) made up of a fiber loop and a degenerate parametric amplifier (single-mode squeezer). The OPO is pumped above threshold in a pulsed manner. The length of the loop is matched to an integer number of intervals between pump pulses, which results in well-identified pulsed modes circulating inside the OPO. After multiple roundtrips, these modes are amplified from the vacuum state to a microscopic amplitude. Due to the phase-sensitive nature of the squeezer, the phases of the resulting modes tend to either $\varphi_i = 0$ or π .

This symmetry is broken by a measurement-based pairwise interaction between pulses, which is arranged as follows. Upon exiting the amplifier, a part of the optical energy is deflected to a homodyne detector which measures the position quadrature of the pulse. The measurement result X_j is stored. Each pulse entering the amplifier is subjected to a phase-space displacement along the position axis. The magnitude of the displacement applied to the i th pulse is proportional to

$$\sum_j J_{ij} X_j, \quad (2)$$

where X_j are the results obtained from the measurements on other pulses.

Thanks to this displacement, the sequence of phases obtained by the pulses at the end of the amplification cycle approaches the solution of the Ising problem, with the phases $\varphi_i = 0$ and π encoding the spin values of $\sigma_i = \pm 1$, respectively. Intuitively, this can be understood as follows. Treating the spins as continuous variables and the Ising Hamiltonian as the potential energy, one finds the gradient of this energy with respect to σ_i as $F = -\nabla_j H = \frac{1}{2} \sum_j J_{ij} \sigma_j$. The physical interpretation of this gradient is the force that pushes the spin configuration towards lower energy. By applying that force in the form of displacement to the pulses circulating inside the OPO, we encourage their phases to align according to the minimum energy configuration.

3. CIM simulator

In SimCIM, we characterize each pulse by its complex amplitude $a_i = \frac{1}{\sqrt{2}}(x_i + ip_i)$, where x_i and p_i are the canonical phase space quadratures. The per-roundtrip change in the amplitude can be written as

$$\Delta a_i = w a_i^* - \gamma a_i - s |a_i|^2 a_i + \zeta \sum_j J_{ij} x_j + f_i / \sqrt{2}. \quad (3)$$

In the above equation, the first term corresponds to the parametric gain, second to linear loss, third to nonlinear loss, fourth to the displacement and fifth to the noise. The nonlinear loss is

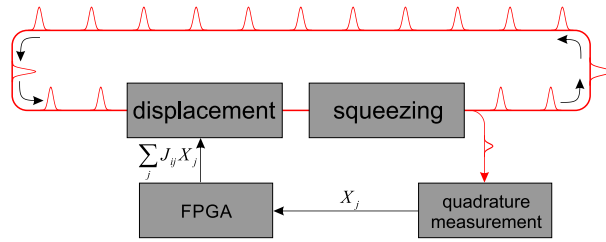


Fig. 1. CIM setup. Each pulse undergoes optical squeezing, linear and non-linear loss, as well as displacement. FPGA: field-programmable gate array.

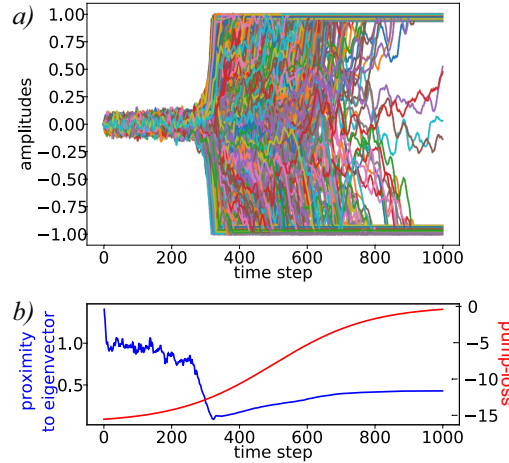


Fig. 2. a) Evolution of the components of the “spin” vector $\vec{x} = \{x_i\}$. b) Time dependence of the pump-loss factor ν and the quantity $||\mathcal{N}(\vec{x}) - \mathcal{N}(\hat{J}\vec{x})||$ [the symbol $\mathcal{N}(\cdot)$ denoting normalization] which shows the proximity of $\vec{x}$ to the eigenvector of $\hat{J}$ with the highest eigenvalue.

associated with the second-harmonic generation, i.e. the transfer of energy from the signal back to the pump of the parametric amplifier. This effect becomes significant only for macroscopic signal amplitudes, but is necessary in the treatment in order to account for the saturation of the OPO. For the displacement term in Eq. (3), we are writing $\sum_j J_{ij} x_j$ instead of Eq. (2) as if this term were calculated from the actual quadratures x_j of the circulating pulses rather than the quadratures X_j measured on a tapped part of the optical energy. The additional quantum noises associated with this replacement are absorbed into the noise term f_i together with the quantum noises arising from the losses.

We can now rewrite Eq. (3) in terms of its real and imaginary components:

$$\begin{aligned}\Delta x_i &= w x_i - \gamma x_i - s(x_i^2 + p_i^2)x_i + \zeta \sum_j J_{ij} x_j + \text{Re} f_i; \\ \Delta p_i &= -w p_i - \gamma p_i - s(x_i^2 + p_i^2)p_i + \text{Im} f_i.\end{aligned}\quad (4)$$

In SimCIM, we omit the nonlinear loss term. This simplifies the calculation by allowing us to drop the second equation in Eq. (4) and restrict the analysis to real numbers:

$$\Delta x_i = \nu x_i + \zeta \sum_j J_{ij} x_j + f_i, \quad (5)$$

where $\nu = w - \gamma$ represents both the parametric gain (controlled by the pump value) and the linear loss. To account for the saturation, we require that the amplitude $|x_i|$ not exceed a certain fixed value x_{sat} . That is, we update the amplitude in each iteration according to $x_i \leftarrow \phi(x_i + \Delta x_i)$, where

$$\phi(x) = \begin{cases} x & \text{for } |x| \leq x_{\text{sat}}; \\ x_{\text{sat}} & \text{otherwise} \end{cases} \quad (6)$$

is the “activation function”. This approach is reminiscent to the Hopfield-Tank simulated annealer [24], but differs from it in that the activation function is applied directly to the “neuron” x_i rather than its increment Δx_i . A further difference is that this function is not differentiable.

We implement this iterative procedure on the GeForce 1080 consumer videoprocessor utilizing the Tensorflow and PyTorch frameworks to take advantage of the parallel computation capability offered by this hardware. To accelerate the convergence, we use the momentum method [25] with the momentum parameter of $\beta = 0.9$. We increase the pump-loss parameter ν according to the hyperbolic tangent law (Fig. 2). When it reaches a certain threshold level, the “spins” x_i start to grow and eventually reach the limiting values of $\pm x_{\text{sat}}$. While some spins reach these values quickly, others oscillate between positive and negative almost until the end of the run.

This dynamics can be understood if we rewrite Eq. (5) in terms of the eigenvectors of the Ising matrix $\hat{J}$:

$$\begin{cases} \Delta e_i = (\nu + \zeta \Lambda_{ii}) e_i + f'_i; \\ |\sum_j R_{ji} e_j| \leq x_{\text{sat}}, \end{cases} \quad (7)$$

where $\vec{e} = \hat{R} \vec{x}$ and $\hat{\Lambda} = \hat{R} \hat{J} \hat{R}^T$, with $\hat{R}$ being the orthogonal matrix that diagonalizes $\hat{J}$. In other words, if the pulse amplitudes comprise an eigenvector of $\hat{J}$, the evolution through the OPO will preserve that eigenvector as long as its amplitude is sufficiently small. In the initial stages of the evolution, the pumping is below the threshold; the “spin” values fluctuate randomly around zero because of the noise term f_i . As the pump parameter ν ramps up, the value of $\nu + \zeta \Lambda_{ii}$ becomes positive for some i , leading to exponential growth of the corresponding eigenvector’s amplitude. However, this growth will only continue until one of the eigenvector components reaches x_{sat} , at which point the second part of Eq. (7) comes into play. After that, $\vec{x}$ is no longer an eigenvector of $\hat{J}$. As more and more components of $\vec{x}$ stabilize at $\pm x_{\text{sat}}$, the effective matrix governing the dynamics of the remaining components changes, resulting in their seemingly chaotic oscillation between positive and negative values.

4. Results

We present our simulation results in the same terms as in CIM papers [12, 13], namely the max-cut value problem [27]. In this problem, each edge of a certain graph has a value associated with it. The goal is to divide the nodes of a certain graph into two subsets such that the total value associated with the edges connecting them is maximized. This problem is equivalent to the Ising problem, as the cut value is given by

$$\text{cut} = -\frac{1}{2} \sum_{i < j} J_{ij} (1 - \sigma_i \sigma_j), \quad (8)$$

where the spin value $\sigma_i = \pm 1$ defines which subset the i th node belongs to. The cut value is maximized if the Ising energy is minimized.

We compare our algorithm with the *bona fide* CIM and the noisy mean field annealing (NMFA) algorithm [8, 21]. In NMFA, the spin in each iteration is driven towards its self-consistent mean-field value $\tanh[\sum_j J_{ij} x_j]$. To allow direct comparison with the CIM of [12], we optimize each algorithm to run for 1000 iterations and run it 100 times, constructing the histogram of the cut value Eq. (8).

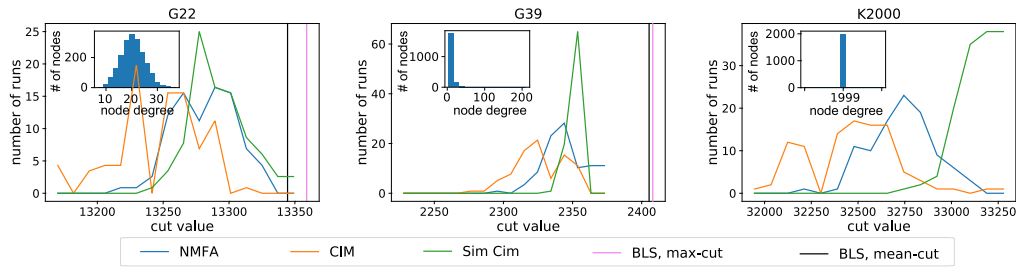


Fig. 3. Histograms for the graphs G22, G39 [26] as well as the fully connected K2000 graph. For each graph the dependence between the number of runs and cut value is presented for the NMFA and SimCim algorithms and for CIM experimental results [13]. For G22 and G39 the results of CIM and algorithms also compared to BLS algorithm which gives the best known cuts.

We first run the simulations on the same set of 2000-node Ising graphs as [12]: relatively sparse graphs G22 and G39 from the G-Set [26] and the fully connected graph K2000. The parameters of both SimCIM (the time-dependent gain-loss parameter ν , overall feedforward factor ζ and the noise f_i) and NMFA have been optimized for each graph to obtain the best performance. The results (Fig. 3) show superior performance of SimCIM with respect to the *bona fide* CIM as well as visible improvement with respect to NMFA, with the advantage being particularly strong for the denser graph K2000. Additionally, we compare the performance with the known results of the breakout local search (BLS) [10], a classical algorithm which is known to perform well on the Ising annealing problem, albeit at a comparatively low speed. We see that the mean cut value obtained in the 100 SimCIM runs lies within 98% of the best BLS result.

A similar level of performance was observed on 800-node GSet graphs G1–G10. SimCIM with the same set of parameters reached the BLS ground state on 8 out of 10 graphs with the mean probability of $\sim 25\%$. For the remaining two graphs (G2 and G9), the maximum cut value reached lay within 99% of the BLS maximum.

A single run of the CIM of [12] with 1000 roundtrips takes 5 ms. On the GeForce 1080 consumer videoprocessor, 100 runs of 1000-iteration SimCIM, launched in parallel, took 400 ms. The corresponding figure for our implementation of NMFA was 550 ms. Therefore, for practically relevant cases in which the statistics from multiple runs of the optimizer are of interest, CIM, SimCIM and NMFA are comparable in terms of the computational speed.

The *bona fide* CIM must compute new displacement amplitudes Eq. (2) at a rate that is faster than the pump pulse repetition rate (1 ns in [12]). In order to ensure that the FPGA keeps up with this rate, the designers of CIM restrict the allowed values of coupling terms in the Ising graph to $J_{ij} = \pm 1$ or 0. Simulators, on the other hand, are not limited by this condition.

To demonstrate this capability, we applied NMFA and SimCIM to a set of 100 instances of 800-node graphs whose edge values are generated randomly according to a Gaussian distribution with the unit dispersion, running each algorithm 100 times for each graph. The same set of parameters has been used for the entire set. The results for one randomly chosen graph are shown in Fig. 4. We found SimCIM to deliver consistently higher cut values, with the cut value averaged over all runs and graphs equal to 5938 for NMFA and 6001 for SimCIM. If we consider the maximum cut value obtained by either algorithm in the 100 runs for each graph, averaged over all graphs, the two methods deliver similar values of 6022 for NMFA and 6021 for SimCIM. The high quality of these results are not surprising, as it is known that the finding of the ground state is easier for all-to-all connected graphs than for sparse ones [28]. Remarkably, however, SimCIM benefits from this to a larger extent than NMFA.

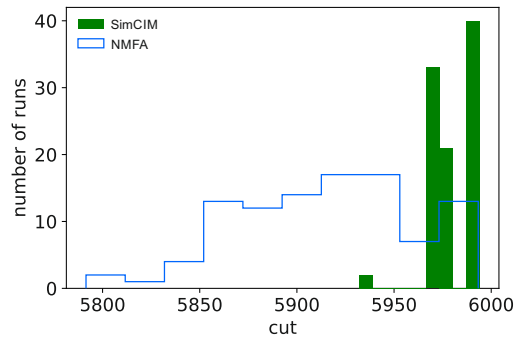


Fig. 4. Results for a random graph of dimension 800 with real-valued couplings J_{ij} that are normally distributed with zero mean and variance equal to one.

5. Conclusion

Our results show that coherent Ising machines can be outperformed by a simulation algorithm running on a classical computer, with the added advantage of the coupling values J_{ij} not being limited to 0 and ± 1 . In this sense, our results contrast those of Haribara *et al.* [14] who found that classical annealing algorithms significantly underperform in comparison with the CIM in terms of the optimization time (this reference does not report any comparison in terms of the annealing quality). We believe this difference to be partially due to the hardware used ([14] employed a CPU and a many-core processor whereas we used a GPU), and partially thanks to our algorithm specially designed to simulate and compete with the CIM.

There appears to be no added value either in the nonclassical nature of the optical states produced, or in the partially analog information processing within the CIM. This is evident from the comparative performance analysis of CIM and SimCIM, and can be understood from Eqs. (4) and (5) describing the simulator. The most computationally expensive part in each simulation step is the calculation of the feedforward term Eq. (2), which corresponds to the multiplication of a matrix and a vector. Both in SimCIM and the *bona fide* CIM, this calculation is performed via a digital processor (GPU vs. FPGA, respectively). Hence it is not surprising that the analog optical processing in the CIM (which computes other terms in the simulation equation) does not bring about visible performance benefits.

Asking ourselves, what is the feature responsible for the faster performance of the CIM compared to simulated annealing and taboo search algorithms, we are compelled to conclude that the CIM's primary asset is the different, albeit classical, underlying physics. A great potential of this quantum-inspired approach is warranted by the vast applicability of the NP-hard problem of annealing in the Ising system [9].

The potential applications of the Ising annealer can be classified into two groups, as summarized in a recent white paper by the D-Wave team [29]. First, a variety of optimization problems, such as the traveling salesman problem and generalizations thereof [30–32], financial portfolio optimization [33], protein folding [34], as well as constraint satisfaction problems, e.g. factorization [35] and satisfiability [36, 37], can be reduced to the Ising optimization. Second, sampling of low-lying energy states is a useful tool for machine learning, specifically for the training of Boltzmann machines, which are in turn the primary component of deep belief neural networks. A related application of the SimCIM and the restricted Boltzmann machine based thereupon is the simulation of quantum condensed matter systems and phase transitions therein [38], as well as quantum state and process tomography [39].

As mentioned previously, CIM's alleged superiority in comparison with the D-Wave annealer [15] has been challenged by McGeogh *et al.* [16]. This paper criticizes the performance metric

used in [15] and furthermore argues that the comparison was made on Ising graphs that are known to be easy to solve for classical machines. This criticism has, in turn, been responded to in the second version of [15]. In summary, it remains an open question on which of the above problems our simulator will prove competitive with respect to both classical heuristics and NISQ devices.

Further improvement to the SimCIM can be obtained by using it in combination with classical taboo search algorithms [40] and by on-the-fly optimization of the simulation parameters based on the real time observation of individual spin trajectories and the features of the Ising graph. Leleu *et al.* used such an approach to eliminate stable equilibria associated with the local minima of the Ising Hamiltonian and therefore enhance the likelihood of reaching the global minimum. For some GSet graphs, they demonstrated cut values above those obtained by BLS [23].

Another interesting potential research direction would be the experimental implementation of the CIM-like machine so that the feedforward term is not calculated digitally but found through optical interference of individual modes (which would need to be synchronized in time). This can be implemented in either a free-space [41] or integrated [42] fashion. The computational speedup would then be obtained thanks to the natural parallelism of analog optical processing. An important observation we make from Fig. 2 is that the gain-loss parameter can be kept negative throughout the simulator run. This means that the parametric gain element may in fact not be necessary in the optical implementation of the CIM (or its successors) provided that the per-roundtrip loss is maintained sufficiently small.

Funding

Russian Foundation for Basic Research (RFBR) (18-37-20033).

Acknowledgments

We thank Alireza Marandi for stimulating discussions, Kirill Kalinin for help in calculations, and Natalia Berloff for making us aware of Hopfield networks.

References

1. R. J. Baxter, *Exactly solved models in statistical mechanics* (Dover Publications, 2014).
2. J. B. Zuber and C. Itzykson, "Quantum field theory and the two-dimensional ising model," *Phys. Rev. D* **15**, 2875–2884 (1977).
3. J. Ambjorn, K. N. Anagnostopoulos, R. Loll, and I. Pushkina, "Shaken, but not stirred: Potts model coupled to quantum gravity," *Nucl. Phys.* **B807**, 251–264 (2009).
4. D. Sornette, "Physics and financial economics (1776-2014): puzzles, ising and agent-based models," *Reports on Prog. Phys.* **77**, 062001 (2014).
5. J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd, "Quantum machine learning," *Nature* **549**, 195 (2017).
6. F. Barahona, "On the computational complexity of ising spin glass models," *J. Phys. A: Math. Gen.* **15**, 3241 (1982).
7. S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, "Optimization by simulated annealing," *Science* **220**, 671–680 (1983).
8. G. Bilbro, R. Mann, T. K. Miller, W. E. Snyder, D. E. Van den Bout, and M. White, "Optimization by mean field annealing," in *Advances in Neural Information Processing Systems* (Neural Information Processing Systems Foundation, Inc., 1989) 91–98.
9. K. A. Smith, "Neural networks for combinatorial optimization: a review of more than a decade of research," *INFORMS J. on Comput.* **11**, 15–34 (1999).
10. U. Benlic and J.-K. Hao, "Breakout local search for the max-cut problem," *Eng. Appl. Artif. Intell.* **26**, 1162 (2013).
11. P. I. Bunyk, E. M. Hoskinson, M. W. Johnson, E. Tolkacheva, F. Altomare, A. J. Berkley, R. Harris, J. P. Hilton, T. Lanting, A. J. Przybysz, and J. Whittaker, "Architectural considerations in the design of a superconducting quantum annealing processor," *IEEE Transactions on Appl. Supercond.* **24**, 1–10 (2014).
12. Takahiro Inagaki, Yoshitaka Haribara, Koji Igarashi, Tomohiro Sonobe, Shuhei Tamate, Toshimori Honjo, Alireza Marandi, Peter L. McMahon, Takeshi Umeki, Koji Enbutsu, Osamu Tadanaga, Hirokazu Takenouchi, Kazuyuki Aihara, Ken-ichi Kawarabayashi, Kyo Inoue, Shoko Utsunomiya, and Hiroki Takesue, "A coherent ising machine for 2000-node optimization problems," *Science*, 354(6312):603–606, 2016.
13. Peter L. McMahon, Alireza Marandi, Yoshitaka Haribara, Ryan Hamerly, Carsten Langrock, Shuhei Tamate, Takahiro Inagaki, Hiroki Takesue, Shoko Utsunomiya, Kazuyuki Aihara, Robert L. Byer, M. M. Fejer, Hideo Mabuchi, and

- Yoshihisa Yamamoto, "A fully programmable 100-spin coherent ising machine with all-to-all connections," *Science*, 354(6312):614–617, 2016.
14. Y. Haribara, H. Ishikawa, S. Utsunomiya, K. Aihara, and Y. Yamamoto, "Performance evaluation of coherent ising machines against classical neural networks," *Quantum Sci. Technol.* **2**, 044002 (2017).
 15. R. Hamerly, T. Inagaki, P. L. McMahon, D. Venturelli, A. Marandi, T. Onodera, E. Ng, C. Langrock, K. Inaba, T. Honjo, K. Enbutsu, T. Umeki, R. Kasahara, S. Utsunomiya, S. Kako, K.-i. Kawarabayashi, R. L. Byer, M. M. Fejer, H. Mabuchi, D. Englund, E. Rieffel, H. Takesue, and Y. Yamamoto, "Experimental investigation of performance differences between Coherent Ising Machines and a quantum annealer," arXiv:1805.05217 (2018).
 16. C. C. McGeoch, W. Bernoudy, and J. King, "Comment on 'scaling advantages of all-to-all connectivity in physical annealers: the coherent Ising machine vs d-wave 2000q'," arXiv:1807.00089 (2018).
 17. Y. Yamamoto, K. Aihara, T. Leleu, K.-i. Kawarabayashi, S. Kako, M. Fejer, K. Inoue, and H. Takesue, "Coherent ising machines—optical neural networks operating at the quantum limit," *npj Quantum Inf.* **3**, 49 (2017).
 18. H. J. Caulfield and S. Dolev, "Why future supercomputing requires optics," *Nat. Photonics* **4**, 261 (2010).
 19. Z. Wang, A. Marandi, K. Wen, R. L. Byer, and Y. Yamamoto, "Coherent ising machine based on degenerate optical parametric oscillators," *Phys. Rev. A* **88**, 063853 (2013).
 20. W. R. Clements, J. J. Renema, Y. H. Wen, H. M. Chrzanowski, W. S. Kolthammer, and I. A. Walmsley, "Gaussian optical ising machines," *Phys. Rev. A* **96**, 043850 (2017).
 21. A. D. King, W. Bernoudy, J. King, A. J. Berkley, and T. Lanting, "Emulating the coherent ising machine with a mean-field algorithm," arXiv:1806.08422 (2018).
 22. K. P. Kalinin and N. G. Berloff, "Global optimization of spin hamiltonians with gain-dissipative systems," arXiv:1807.00699 (2018).
 23. T. Leleu, Y. Yamamoto, P. L. McMahon, and K. Aihara, "Destabilization of local minima in analog spin systems by correction of amplitude heterogeneity," *Phys. Rev. Lett.* **122**, 040607 (2019).
 24. J. J. Hopfield and D. W. Tank, "Computing with neural circuits: A model," *Science* **233**, 625–633 (1986).
 25. N. Qian, "On the momentum term in gradient descent learning algorithms," *Neural networks* **12**, 145–151 (1999).
 26. "Index of Gset," (Stanford University, 2003). <https://web.stanford.edu/yyye/yyye/Gset/>
 27. R. Karp, "Reducibility among combinatorial problems," in *Complexity of Computer Computations*, R. Miller and J. Thatcher, eds. (Plenum Press, 1972), pp. 85–103.
 28. C. Banderier, H.-K. Hwang, V. Ravelomana, and V. Zacharovas, "Average case analysis of np-complete problems: Maximum independent set and exhaustive search algorithms," *Proc. AofA* 2009 .
 29. "D-wave problem-solving handbook," (2018). https://docs.dwavesys.com/docs/latest/doc_handbook.html
 30. F. Neukart, G. Compostella, C. Seidel, D. von Dollen, S. Yarkoni, and B. Parney, "Traffic flow optimization using a quantum annealer," *Front. ICT* **4**, 29 (2017).
 31. S. Feld, C. Roch, T. Gabor, C. Seidel, F. Neukart, I. Galter, W. Mauerer, and C. Linnhoff-Popien, "A hybrid solution method for the capacitated vehicle routing problem using a quantum annealer," arXiv:1811.07403 (2018).
 32. M. Ohzeki, A. Miki, M. J. Miyama, and M. Terabe, "Control of automated guided vehicles without collision by quantum annealer and digital devices," arXiv:1812.01532 (2018).
 33. G. Rosenberg, P. Haghnegahdar, P. Goddard, P. Carr, K. Wu, and M. L. De Prado, "Solving the optimal trading trajectory problem using a quantum annealer," *IEEE J. Sel. Top. Signal Process.* **10**, 1053–1060 (2016).
 34. A. Perdomo-Ortiz, N. Dickson, M. Drew-Brook, G. Rose, and A. Aspuru-Guzik, "Finding low-energy conformations of lattice protein models by quantum annealing," *Sci. Reports* **2**, 571 (2012).
 35. E. R. Anschuetz, J. P. Olson, A. Aspuru-Guzik, and Y. Cao, "Variational quantum factoring," arXiv:1808.08927 (2018).
 36. A. Douglass, A. D. King, and J. Raymond, "Constructing sat filters with a quantum annealer," in *International Conference on Theory and Applications of Satisfiability Testing*, (Springer, 2015), pp. 104–120.
 37. Z. Bian, F. Chudak, W. Macready, A. Roy, R. Sebastiani, and S. Varotti, "Solving sat and maxsat with a quantum annealer: Foundations, encodings, and preliminary results," arXiv:1811.02524 (2018).
 38. G. Carleo and M. Troyer, "Solving the quantum many-body problem with artificial neural networks," *Science* **355**, 602–606 (2017).
 39. G. Torlai, G. Mazzola, J. Carrasquilla, M. Troyer, R. Melko, and G. Carleo, "Neural-network quantum state tomography," *Nat. Phys.* **14**, 447 (2018).
 40. E. Blanzieri and D. Pastorello, "Quantum annealing tabu search for qubo optimization," arXiv:1810.09342 (2018).
 41. A. Marandi, Z. Wang, K. Takata, R. L. Byer, and Y. Yamamoto, "Network of time-multiplexed optical parametric oscillators as a coherent ising machine," *Nat. Photonics* **8**, 937 (2014).
 42. Charles Roques-Carmes, Yichen Shen, Cristian Zanoci, Mihika Prabhu, Fadi Atieh, Li Jing, Tena Dubcek, Vladimir Ceperic, John D Joannopoulos, Dirk Englund, and Marin Soljacic. "Photonic recurrent Ising sampler," arXiv:1811.02705 (2018).