PROCEEDINGS OF SPIE

SPIEDigitalLibrary.org/conference-proceedings-of-spie

Synchronously-pumped OPO coherent Ising machine: benchmarking and prospects

Hamerly, Ryan, Inagaki, Takahiro, McMahon, Peter, Venturelli, Davide, Marandi, Alireza, et al.

Ryan Hamerly, Takahiro Inagaki, Peter L. McMahon, Davide Venturelli, Alireza Marandi, Dirk R. Englund, Yoshihisa Yamamoto, "Synchronously-pumped OPO coherent Ising machine: benchmarking and prospects," Proc. SPIE 11299, AI and Optical Data Sciences, 112990J (24 February 2020); doi: 10.1117/12.2547046

SPIE.

Event: SPIE OPTO, 2020, San Francisco, California, United States

Synchronously-pumped OPO coherent Ising machine: benchmarking and prospects

Ryan Hamerly[1,2,3], Takahiro Inagaki[4], Peter L. McMahon[5,9,2], Davide Venturelli[6,7], Alireza Marandi[8,9], Dirk R. Englund[1], and Yoshihisa Yamamoto[9,3]

[1]Research Laboratory of Electronics, MIT, Cambridge, MA 02139
[2]National Institute of Informatics, Tokyo 101-8403, Japan
[3]NTT Research, East Palo Alto, CA 94303
[4]NTT Basic Research Laboratories, Atsugi, Kanagawa 243-0198, Japan
[5]School of Applied and Engineering Physics, Cornell University, Ithaca, NY 14853
[6]NASA Ames Research Center Quantum AI Laboratory (QuAIL), Moffett Field, CA 94035
[7]USRA RIACS, Mountain View, CA 94035
[8]California Institute of Technology, Pasadena, CA 91125, USA
[9]Edward L. Ginzton Laboratory, Stanford University, Stanford, CA 94305

ABSTRACT

The coherent Ising machine (CIM) is a network of optical parametric oscillators (OPOs) that solves for the ground state of Ising problems through OPO bifurcation dynamics. Here, we present experimental results comparing the performance of the CIM to quantum annealers (QAs) on two classes of NP-hard optimization problems: ground-state calculation of the Sherrington-Kirkpatrick (SK) model and MAX-CUT. While the two machines perform comparably on sparsely-connected problems such as cubic MAX-CUT, on problems with dense connectivity, the QA shows an exponential performance penalty relative to CIMs. We attribute this to the embedding overhead required to map dense problems onto the sparse hardware architecture of the QA, a problem that can be overcome in photonic architectures such as the CIM.

Keywords: Optical parametric oscillator, phase transition, quantum annealing, Ising model, combinatorial optimization

1. INTRODUCTION

Many important tasks in modern society rely on the efficient solution of hard optimization problems such as scheduling, route planning, portfolio optimization, circuit layout, and drug discovery. These problems belong to the NP-hard complexity class,[1] for which exact solutions require exponential time. As a result, significant computational resources are required to solve such problems, even if only approximate solutions are required. Therefore, there is significant interest in accelerating the solution of such problems special-purpose machines, including ASICs,[2] *physical annealers* that map the optimization directly to the underlying physical dynamics,[3] and quantum annealers.[4] Most such special-purpose machines have been targeted at the Ising problem, a canonical NP-hard problem to which all NP problems can be reduced in polynomial time.[5–7] The Ising problem consists of finding spin configuration $\vec{\sigma} \in \{-1, 1\}^N$ that minimizes the Ising energy:

$$H = \frac{1}{2}\sum_{ij} J_{ij}\sigma_i\sigma_j + \sum_i h_i\sigma_i \tag{1}$$

While most special-purpose solvers utilize electronics, interest in photonic solvers is rapidly growing. Photonics offers unique advantages for information processing, including high data rates,[8] low latency, low power consumption,[9] elimination of the interconnect bottleneck,[10] and the ability to perform linear algebra operations with passive optics.[11,12] Motivated by this, we have proposed an *optical* annealer called the coherent Ising machine

Send correspondence to R.H.: rhamerly@mit.edu

AI and Optical Data Sciences, edited by Bahram Jalali, Ken-ichi Kitayama,
Proc. of SPIE Vol. 11299, 112990J · © 2020 SPIE · CCC
code: 0277-786X/20/$21 · doi: 10.1117/12.2547046

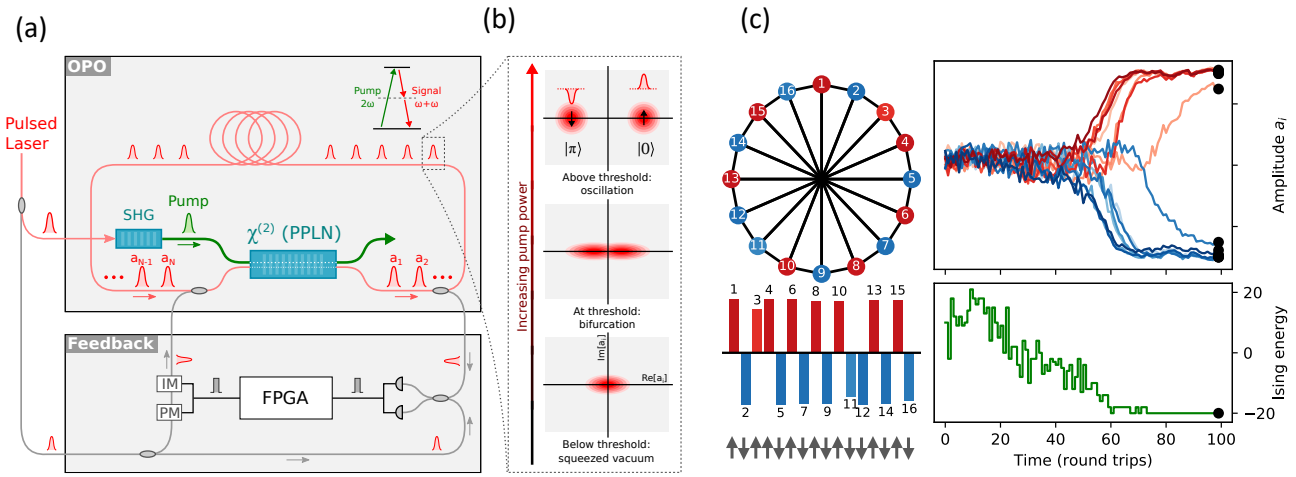

Figure 1. (a) Schematic of CIM consisting of time-multiplexed OPO and measurement-feedback apparatus. (b) Bifurcation in OPO state during transition from below-threshold squeezed state to (bistable) above-threshold coherent state. (c) Solution of antiferromagnetic Ising problem on the Möbius ladder with the CIM.[17]

(CIM) that maps the Ising problem onto the dynamics of a network of degenerate optical parametric oscillators (OPOs).[13,14] A degenerate OPO is an optical cavity with a $\chi^{(2)}$ nonlinear medium, pumped at the second harmonic of the cavity resonance.[15] If a pulsed pump is used, multiple identical OPOs can be time-multiplexed in a single cavity through a technique of synchronous pumping,[16] as shown in Fig. 1(a). Coupling between the OPOs is realized using a measurement-feedback apparatus (bottom of Fig. 1(a)), consisting of a homodyne detector, an FPGA, and a modulator. This enables progammable Ising problems with arbitrary J_{ij} of sizes up to $N = 2000$ to be studied on the CIM.[17,18]

The CIM solves Ising problems through the gain-dissipative bifurcation dynamics of OPOs.[14,19–22] Fundamentally, this is based on the OPO bifurcation from a below-threshold (quantum) squeezed state to an above-threshold (classical) coherent state (Fig. 1(b)). Due to the phase-sensitive nature of OPO gain only two states with phases $\{0, \pi\}$ can occur above threshold: the Ising spin σ_i is encoded in this phase. The optical annealing process is illustrated in Fig. 1(c), where the OPO bifurcation occurs simultaneously with the inter-OPO coupling, which pushes the OPOs to lower the collective Ising energy of the system. As a result, when the pump is sufficiently far above threshold that all OPO amplitudes have settled to values with roughly equal amplitudes, the resulting state (for this problem) is the Ising ground state.

2. BENCHMARKING CIM AND QA

This work presents a performance comparison between the CIM and the D-Wave 2000Q quantum annealer (DW2Q).[23] Inspired by the adiabatic principle,[24] quantum annealing[4,25] is an established technique for quantum-enhanced optimization and may offer the possibility for a limited speedup on certain classes of problems, although obtaining conclusive evidence for such a speedup has proved challenging.[26–28] However, existing quantum annealers suffer from limited connectivity, so that actual problems must be *embedded* into the solver architecture's native graph before they can be solved.[29,30] Embedded problems are often much larger than the originals and can include many constraints, leading to a potentially significant degradation in annealer performance.

In light of these limitations, it is insightful to consider performance benchmarks of general problem classes, chosen independently of the annealer's architecture. Here, we compare the performance of the CIM and DW2Q on three classes of NP-hard Ising problems: fully-connected Sherrington-Kirkpatrick (SK) spin glasses,[31] MAX-CUT on dense graphs, and MAX-CUT on regular sparse graphs.[5] Problems with all-to-all or dense connectivity are usually embedded with precomputed clique embeddings (Fig. 2(a)), which which map each logical qubit to an L-shaped ferromagnetic chain on the Chimera.[32] Embedding introduces an additional degree of freedom

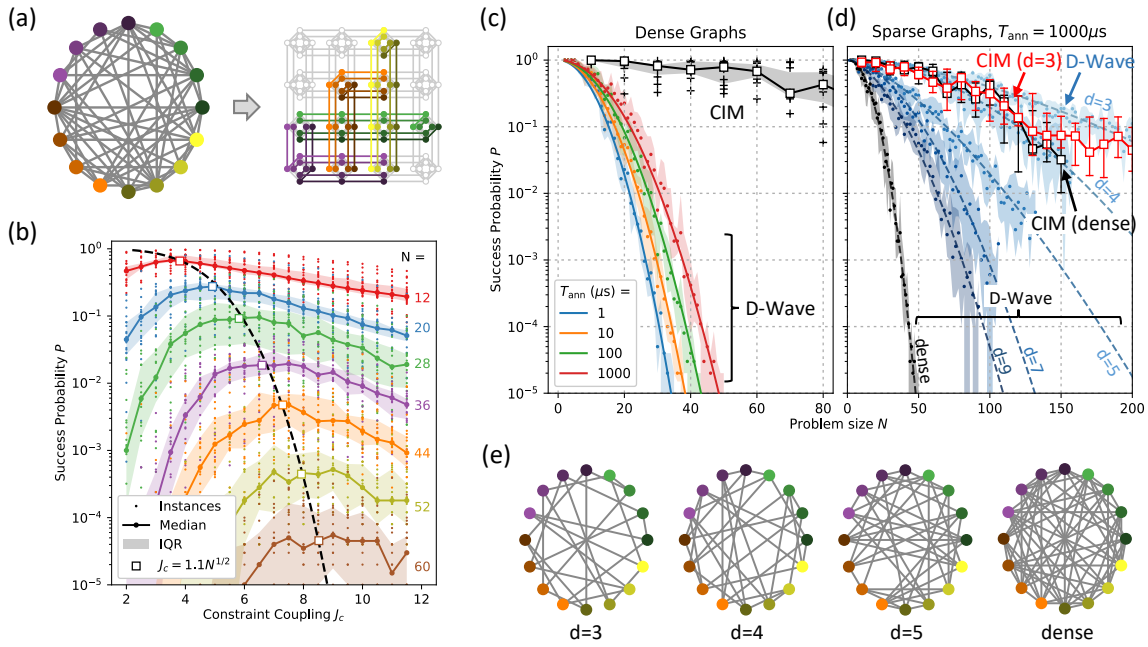

Figure 2. (a) Optimized quantum annealing of embedded SK problems on DW2Q: clique embedding to map a densely-connected problem onto a Chimera graph. (b) SK success probability as a function of problem size N and embedding constraint J_c. (c) MAX-CUT success probability for dense (edge-density-$\frac{1}{2}$) graphs. (d) Regular sparse graphs with $d = 3, 4, 5, 7, 9$ edges per vertex. (e) Representative sample of $N = 16$ sparse and dense graphs.

	Stanford CIM	NTT CIM	D-Wave 2000Q
Spins	OPO	OPO	JJ qubits
Size	100	2000	2048
Ising model	J_{ij}	J_{ij}	J_{ij} and Zeeman
Couplings	Full (5,000)	Full (2×10^6)	Chimera (12,000)
J_{ij} range	Continuous	$\{-1, 0, +1\}$	Continuous
Largest clique	$N = 100$	$N = 2000$	$N = 64$
Anneal time	1600 μs	5000 μs	1–2000 μs

Table 1. Specifications of the annealing machines used in this study.

J_c, which dictates the strength of the constraint couplings relative to the logical couplings. To optimize the performance of the quantum annealer, it is essential to find the optimal J_c, which is size- and problem-dependent (Fig. 2(b)). The optimal J_c is believed to be related to the emergence of a spin-glass state of the embedded problem

2.1 MAX-CUT and SK Performance

Fig. 2(c) plots the D-Wave success probability for MAX-CUT as a function of problem size N (the performance on SK problems is similar, see Ref.[23]). Here, the embedding parameter J_c has been optimized as a function of problem size following the procedure in Fig. 2(b). Even so, the success probability falls off super-exponentially with N, the data fitting roughly to the curve $P = e^{-(N/N_{\rm dw})^2}$, where $N_{\rm dw}$ is a constant than increases slowly, roughly logarithmically, with the annealing time $T_{\rm ann}$. By contrast, the CIM success probability decreases much more gradually with N, leading to a several-orders-of-magnitude difference in success probability for moderate-size problems $N \geq 50$.

On the other hand, embedding heuristics[30] can embed sparse graphs using far fewer physical qubits; this

reduced embedding overhead suggests that the DW2Q should perform better on sparse problems. As Fig. 2(d) shows, the DW2Q outperforms the CIM for cubic ($d = 3$) and $d = 4$ graphs, while for graphs of higher density, the CIM shows a strong scaling advantage. Unlike DW2Q, the CIM's performance does not appear to depend on the graph edge density. This is consistent with the hypothesis that the embedding overhead, which is much greater for dense graphs than sparse graphs, is a primary factor limiting the performance of physical annealers.

2.2 Time-to-Solution and Optimal Annealing Time

While success probability provides intuition about an annealer's performance, a more rigorous performance benchmark is the *time to solution* $T_{\text{soln}} = T_{\text{ann}}\lceil \log(0.01)/\log(1-P)\rceil$, defined as the time required to reach the Ising ground state with 99% probability. When calculating time to solution, it is important that all free parameters in the annealer be optimized. For the DW2Q, this includes the constraint term J_c as well as the annealing time T_{ann}. Finding the optimal annealing time is particularly important, as failing to properly optimize this parameter can lead to an illusion of speedup at small problem sizes if the annealing time is sufficiently large.[27]

The CIMs used here are experimental systems with fixed annealing time. To study the effect of varying T_{ann}, we performed semiclassical simulations of the CIM using c-number stochastic differential equations (c-SDEs). The c-SDE model, which resembles noisy mean-field annealing,[21] can accurately predict the CIM's performance for the problems studied here.[17,20] Fig. 3(a) shows the success probability and time to solution of the CIM (c-SDE simulations) for dense MAX-CUT problems. An asymptotic fit $P = \exp(-O(N))$ for large problems is observed at fixed annealing time, and the time to solution at optimal annealing times appears to follow a $T_{\text{soln}} \propto \exp(O(N^{1/2}))$ curve. This is in contrast with the DW2Q (Fig. 3(b)), where the optimal time-to-solution goes as $T_{\text{soln}} \propto \exp(O(N))$. By problem size $N = 50$, the difference in time-to-solution is $10^6\times$; extrapolated to $N = 64$ (the largest size supported by clique embeddings in the DW2Q), it exceeds $10^9\times$. A similar scaling in time to solution is observed for SK problems. For cubic MAX-CUT problems, consistent with Fig. 2(d), the DW2Q outperforms the CIM in absolute terms at all tested problem sizes, but the CIM has a scaling advantage suggesting a performance edge for larger problems.

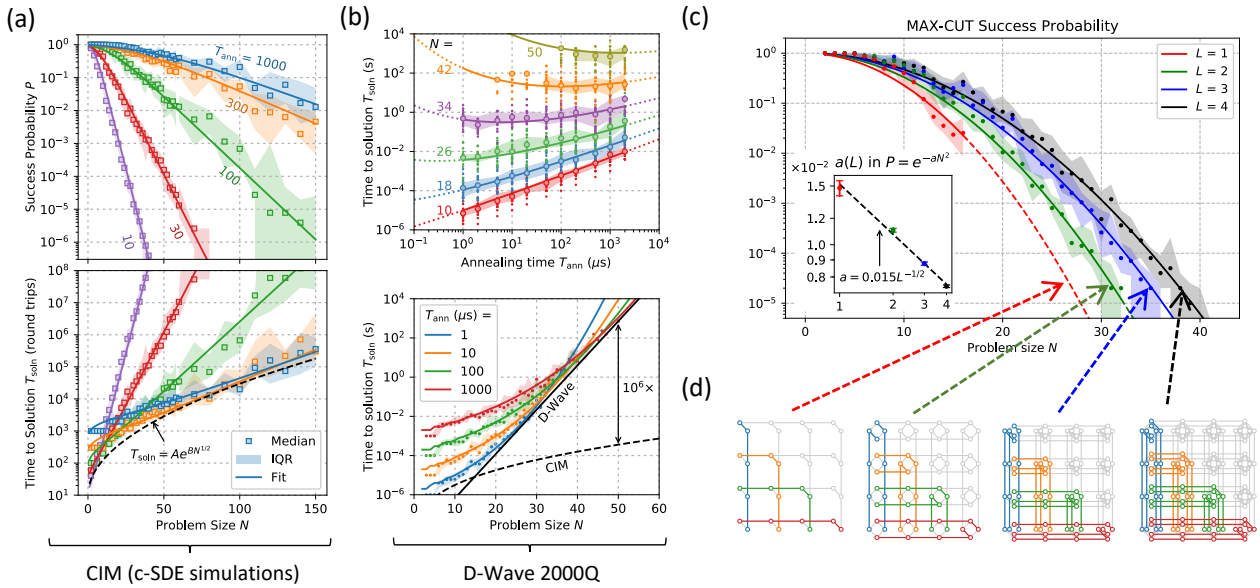

Figure 3. (a) Time-to-solution analysis for MAX-CUT. C-SDE simulations of CIM giving success probability and time to solution. (b) Time to solution for DW2Q and comparison with CIM. (c) DW2Q success probability for dense MAX-CUT run on DW2Q embedded onto subgraphs of the Chimera. Data fit to $P = e^{-a(L)N^2}$, where the constant a depends on Chimera degree L. Annealing time is 20μs. (d) Chimera subgraphs with $L = 1$–4 and their associated clique embeddings.

2.3 Effect of Annealer Sparsity

To illustrate the effect of embedding overhead limiting the performance of quantum annealers on dense problems, we varied the connectivity degree L of the Chimera graph in the DW2Q. A Chimera graph is a 2D lattice, with the $2L$ qubits at each lattice point arranged in a biclique. This arrangement supports up to $L+2$ connections per qubit, so L is a rough proxy for the connectivity of the hardware graph. It is also a proxy for embedding overhead: for clique embeddings, an N-qubit fully-connected graph maps to approximately N^2/L physical qubits, indicating the embedding overhead decreases with larger L.

Fig. 3(c) shows DW2Q results for MAX-CUT on dense graphs embedded into Chimera graphs with $L = 1$–4. Consistent with the results above, the success probability fits to a square-exponential $P = \exp(-O(N^2))$. However, the coefficient in this exponential depends on L, leading to a significant performance penalty for the sparser Chimera graphs at large problem sizes. This is consistent with the intuition that embedding a problem into a sparse native hardware graph will lead to a degradation of success probability in addition to requiring a larger number of physical qubits.

3. CONCLUSION

With its measurement-feedback coupling, the CIM has native all-to-all connectivity and does not require minor embedding. This provides it with a significant advantage over the DW2Q (and other annealers with sparse hardware connectivity) on MAX-CUT and SK problems defined on dense graphs, where the embedding overhead is costly. However, the DW2Q's lead on cubic MAX-CUT problems suggests that, if efficient embeddings are employed, annealers such as the DW2Q can maintain an edge. The importance of embedding overhead sketched here motivates recent efforts to improve the performance of physical annealers with alternate embedding strategies[33] or increased hardware connectivity.[34–36]

Quantum annealing offers the promise of quantum speedup, but this must be balanced against the overhead required to embed the problem into the annealer's hardware. As we have shown, this *embedding overhead* can cause exponential slow-down for general Ising problems. By contrast, optical annealing with the measurement-feedback CIM is intrinsically all-to-all and better adapted to problems with dense connectivity. Moreover, the CIM is based on a very extensible concept of gain-dissipative optimization. Modifications to the annealing dynamics may significantly improve the success probability for large problem instances,[37–39] while operating in a regime with stronger nonlinearities[40,41] or non-Gaussian measurements[42] may lead to a non-semiclassical regime where quantum effects play a central role and quantum speedup is possible.

ACKNOWLEDGMENTS

The authors acknowledge Salvatore Mandrà for useful discussions and parallel-tempering simulation results, and Daniel Lidar, Andrew King, and Catherine McGeoch for helpful correspondence. This research was funded by the JST ImPACT Program of the Cabinet Office of Japan. Additional support was provided by: (R.H.) IC Postdoctoral Fellowship at MIT through U.S. DOE / ODNI; (P.L.M.) Stanford Nano/Quantum Fellowship; (D.V.) NASA Academic Mission Services, no. NNA16BD14C; (D.E.) U.S. ARO at ISN / MIT (no. W911NF-18-2-0048) and SRC-NSF E2CDA.

REFERENCES

[1] Rudich, S. and Wigderson, A., [*Computational Complexity Theory*], vol. 10, American Mathematical Soc. (2004).

[2] Yoshimura, C., Yamaoka, M., Hayashi, M., Okuyama, T., Aoki, H., Kawarabayashi, K.-i., and Mizuno, H., "Uncertain behaviours of integrated circuits improve computational performance," *Scientific Reports* **5** (2015).

[3] Sutton, B., Camsari, K. Y., Behin-Aein, B., and Datta, S., "Intrinsic optimization using stochastic nanomagnets," *Scientific Reports* **7**(44370) (2017).

[4] Kadowaki, T. and Nishimori, H., "Quantum annealing in the transverse Ising model," *Physical Review E* **58**(5), 5355 (1998).

[5] Karp, R. M., "Reducibility among combinatorial problems," in [*Complexity of Computer Computations*], 85–103, Springer (1972).
[6] Barahona, F., "On the computational complexity of Ising spin glass models," *Journal of Physics A: Mathematical and General* **15**(10), 3241 (1982).
[7] Lucas, A., "Ising formulations of many NP problems," *Frontiers in Physics* **2**, 5 (2014).
[8] Wang, C., Zhang, M., Chen, X., Bertrand, M., Shams-Ansari, A., Chandrasekhar, S., Winzer, P., and Lončar, M., "Integrated lithium niobate electro-optic modulators operating at CMOS-compatible voltages," *Nature* **562**(7725), 101 (2018).
[9] Nozaki, K., Matsuo, S., Fujii, T., Takeda, K., Shinya, A., Kuramochi, E., and Notomi, M., "Femtofarad optoelectronic integration demonstrating energy-saving signal conversion and nonlinear functions," *Nature Photonics* **13**, 454–459 (2019).
[10] Miller, D. A., "Attojoule optoelectronics for low-energy information processing and communications," *Journal of Lightwave Technology* **35**(3), 346–396 (2017).
[11] Shen, Y., Harris, N. C., Skirlo, S., Prabhu, M., Baehr-Jones, T., Hochberg, M., Sun, X., Zhao, S., Larochelle, H., Englund, D., et al., "Deep learning with coherent nanophotonic circuits," *Nature Photonics* **11**(7), 441 (2017).
[12] Hamerly, R., Bernstein, L., Sludds, A., Soljačić, M., and Englund, D., "Large-scale optical neural networks based on photoelectric multiplication," *Physical Review X* **9**(2), 021032 (2019).
[13] Utsunomiya, S., Takata, K., and Yamamoto, Y., "Mapping of Ising models onto injection-locked laser systems," *Optics Express* **19**(19), 18091–18108 (2011).
[14] Wang, Z., Marandi, A., Wen, K., Byer, R. L., and Yamamoto, Y., "Coherent Ising machine based on degenerate optical parametric oscillators," *Physical Review A* **88**(6), 063853 (2013).
[15] Boyd, S. and Vandenberghe, L., [*Convex Optimization*], Cambridge University Press (2004).
[16] Marandi, A., Wang, Z., Takata, K., Byer, R. L., and Yamamoto, Y., "Network of time-multiplexed optical parametric oscillators as a coherent Ising machine," *Nature Photonics* **8**(12), 937–942 (2014).
[17] McMahon, P. L., Marandi, A., Haribara, Y., Hamerly, R., Langrock, C., Tamate, S., Inagaki, T., Takesue, H., Utsunomiya, S., Aihara, K., et al., "A fully-programmable 100-spin coherent Ising machine with all-to-all connections," *Science* **354**(6312), 614–617 (2016).
[18] Inagaki, T., Haribara, Y., Igarashi, K., Sonobe, T., Tamate, S., Honjo, T., Marandi, A., McMahon, P. L., Umeki, T., Enbutsu, K., et al., "A coherent Ising machine for 2000-node optimization problems," *Science* **354**(6312), 603–606 (2016).
[19] Haribara, Y., Utsunomiya, S., and Yamamoto, Y., "Computational principle and performance evaluation of coherent Ising machine based on degenerate optical parametric oscillator network," *Entropy* **18**(4), 151 (2016).
[20] Hamerly, R., Inaba, K., Inagaki, T., Takesue, H., Yamamoto, Y., and Mabuchi, H., "Topological defect formation in 1D and 2D spin chains realized by network of optical parametric oscillators," *International Journal of Modern Physics B* **30**(25), 1630014 (2016).
[21] King, A. D., Bernoudy, W., King, J., Berkley, A. J., and Lanting, T., "Emulating the coherent Ising machine with a mean-field algorithm," *arXiv preprint arXiv:1806.08422* (2018).
[22] Tiunov, E. S., Ulanov, A. E., and Lvovsky, A., "Annealing by simulating the coherent Ising machine," *Optics Express* **27**(7), 10288–10295 (2019).
[23] Hamerly, R., Inagaki, T., McMahon, P. L., Venturelli, D., Marandi, A., Onodera, T., Ng, E., Langrock, C., Inaba, K., Honjo, T., et al., "Experimental investigation of performance differences between coherent Ising machines and a quantum annealer," *Science Advances* **5**(5), eaau0823 (2019).
[24] Farhi, E., Goldstone, J., Gutmann, S., Lapan, J., Lundgren, A., and Preda, D., "A quantum adiabatic evolution algorithm applied to random instances of an NP-complete problem," *Science* **292**(5516), 472–475 (2001).
[25] Boixo, S., Rønnow, T. F., Isakov, S. V., Wang, Z., Wecker, D., Lidar, D. A., Martinis, J. M., and Troyer, M., "Evidence for quantum annealing with more than one hundred qubits," *Nature Physics* **10**(3), 218–224 (2014).

[26] Somma, R. D., Nagaj, D., and Kieferová, M., "Quantum speedup by quantum annealing," *Physical Review Letters* **109**(5), 050501 (2012).

[27] Rønnow, T. F., Wang, Z., Job, J., Boixo, S., Isakov, S. V., Wecker, D., Martinis, J. M., Lidar, D. A., and Troyer, M., "Defining and detecting quantum speedup," *Science* **345**(6195), 420–424 (2014).

[28] Albash, T. and Lidar, D. A., "Demonstration of a scaling advantage for a quantum annealer over simulated annealing," *Physical Review X* **8**(3), 031016 (2018).

[29] Choi, V., "Minor-embedding in adiabatic quantum computation: I. the parameter setting problem," *Quantum Information Processing* **7**(5), 193–209 (2008).

[30] Cai, J., Macready, W. G., and Roy, A., "A practical heuristic for finding graph minors," *arXiv preprint arXiv:1406.2741* (2014).

[31] Sherrington, D. and Kirkpatrick, S., "Solvable model of a spin-glass," *Physical Review Letters* **35**(26), 1792 (1975).

[32] Boothby, T., King, A. D., and Roy, A., "Fast clique minor generation in Chimera qubit connectivity graphs," *Quantum Information Processing* **15**(1), 495–508 (2016).

[33] Takemoto, T., Mertig, N., Hayashi, M., Susa-Tanaka, S., Teramoto, H., Nakamura, A., Takigawa, I., Minato, S.-i., Komatsuzaki, T., and Yamaoka, M., "FPGA-based QBoost with large-scale annealing processor and accelerated hyperparameter search," in [*2018 International Conference on ReConFigurable Computing and FPGAs (ReConFig)*], 1–8, IEEE (2018).

[34] Katzgraber, H. G. and Novotny, M., "A small-world search for quantum speedup: How small-world interactions can lead to improved quantum annealer designs," *arXiv preprint arXiv:1805.09510* (2018).

[35] Dattani, N., Szalay, S., and Chancellor, N., "Pegasus: The second connectivity graph for large-scale quantum annealing hardware," *arXiv preprint arXiv:1901.07636* (2019).

[36] Onodera, T., Ng, E., and McMahon, P. L., "A quantum annealer with fully programmable all-to-all coupling via Floquet engineering," *arXiv preprint arXiv:1907.05483* (2019).

[37] Leleu, T., Yamamoto, Y., McMahon, P. L., and Aihara, K., "Destabilization of local minima in analog spin systems by correction of amplitude heterogeneity," *Physical Review Letters* **122**(4), 040607 (2019).

[38] Ercsey-Ravasz, M. and Toroczkai, Z., "Optimization hardness as transient chaos in an analog approach to constraint satisfaction," *Nature Physics* **7**(12), 966 (2011).

[39] Kalinin, K. P. and Berloff, N. G., "Global optimization of spin hamiltonians with gain-dissipative systems," *Scientific Reports* **8**(1), 17791 (2018).

[40] Jankowski, M., Langrock, C., Desiatov, B., Marandi, A., Wang, C., Zhang, M., Phillips, C. R., Lončar, M., and Fejer, M., "Ultrabroadband nonlinear optics in nanophotonic periodically poled lithium niobate waveguides," *arXiv preprint arXiv:1909.08806* (2019).

[41] Onodera, T., Ng, E., Lörch, N., Yamamura, A., Hamerly, R., McMahon, P. L., Marandi, A., and Mabuchi, H., "Nonlinear quantum behavior of ultrashort-pulse optical parametric oscillators," *arXiv preprint arXiv:1811.10583* (2018).

[42] Yanagimoto, R., McMahon, P. L., Ng, E., Onodera, T., and Mabuchi, H., "Embedding entanglement generation within a measurement-feedback coherent Ising machine," *arXiv preprint arXiv:1906.04902* (2019).