npj | Quantum Information

www.nature.com/npjqi

REVIEW ARTICLE OPEN

Coherent Ising machines—optical neural networks operating at the quantum limit

Yoshihisa Yamamoto[1,2], Kazuyuki Aihara[3], Timothee Leleu[3], Ken-ichi Kawarabayashi[4], Satoshi Kako[1,4], Martin Fejer[2], Kyo Inoue[5] and Hiroki Takesue[6]

In this article, we will introduce the basic concept and the quantum feature of a novel computing system, coherent Ising machines, and describe their theoretical and experimental performance. We start with the discussion how to construct such physical devices as the quantum analog of classical neuron and synapse, and end with the performance comparison against various classical neural networks implemented in CPU and supercomputers.

npj Quantum Information (2017)3:49 ; doi:10.1038/s41534-017-0048-9

INTRODUCTION

Optimization problems with discrete and continuous variables, some of which belong to NP-hard or NP-complete class in complexity theory, are ubiquitous in numerous important areas, including operations and scheduling, drug discovery, wireless communications, finance, integrated circuit design, compressed sensing, and machine learning. Despite rapid advances in both algorithms and digital computer technology, even modest sized NP-hard/NP-complete problems that arise in practice may be very difficult to solve on modern digital computers. One alternative to current von-Neumann-computer-based methods for solving such problems is a neural network realized with analog electronic circuits and other systems.[1–4] Another approach of contemporary interest is the adiabatic quantum computation (AQC)[5] and quantum annealing (QA).[6,7] Sophisticated AQC/QA devices are already under development,[8–11] but providing dense connectivity between qubits remains a major challenge[12] with serious implications for the efficiency of AQC/QA systems.[13]

Networks of degenerate optical parametric oscillators (DOPOs) are an alternative physical system, with an unconventional operating mechanism,[14–17] for solving the Ising problem,[18–20] and by extension many other combinatorial optimization problems.[21] Formally, the N-spin Ising problem is to find the configuration of spins $\sigma_i \in \{-1, +1\}$ $(i = 1, \ldots, N)$ that minimizes the energy function $H = -\sum_{1\le i<j\le N} J_{ij}\sigma_i\sigma_j - \sum_{1\le i\le N} h_i\sigma_i$, where the particular problem instance being solved is specified by the $N \times N$ matrix J (with elements J_{ij}) and the N-dimensional vector h (with elements h_i).

In DOPO Ising machines, the up-spin and down-spin are represented by the positive and negative in-phase amplitudes, i.e. $\sigma_i = X_i/|X_i|$. Then, the energy function H is mapped to the effective loss, photon decay rate, of the DOPO network by implementing the mutual coupling proportional to J_{ij} between two DOPOs and injecting the dc field proportional to h_i to individual DOPO.[14] The ground state of the Ising Hamiltonian corresponds to an oscillation mode with the minimum network loss. At the pump rate below an oscillation threshold, every DOPO is in a squeezed vacuum state which is interpreted as linear superposition of $\sigma_i = +1$ (positive amplitude or up-spin) and $\sigma_i = -1$ (negative amplitude or down-spin) so that the probability amplitudes for 2^N spin configurations are all identical, as shown in Fig. 1a. At the threshold pump rate, the ground state with a minimum loss reaches a threshold first and oscillates as a single stable mode, which triggers the stimulated emission of photons and cross-gain saturation for all the other modes, as shown in Fig. 1b. Thus, the solution of the Ising problem can be found by simply detecting the single oscillation mode.[14]

The input states into all DOPOs are vacuum states, $|0\rangle_1|0\rangle_2, \cdots |0\rangle_N$. The essential quantum operations are the collective (not individual) symmetry breaking at DOPO threshold, which translates all DOPOs to either positive amplitude $|\alpha\rangle$ or negative amplitude $|-\alpha\rangle$ coherent state in a correlated way. Finally, the computational results are read out by optical homodyne detectors. A block diagram of the CIM operation is shown in Fig. 1c.

We have realized such a system with a scalable architecture that uses a measurement-feedback circuit in place of optical delay lines, which were used in earlier experiments.[15,22,23] Our measurement-feedback-based DOPO Ising machine allows connections between any spin and any other spin, and is fully programmable.[24,25] Fully connected superconducting Ising machines have been also proposed recently.[26,27]

CLASSICAL VS. QUANTUM NEURAL NETWORKS

Nonlinear devices that have a characteristic input–output relation shown in Fig. 2a represent a typical gain function of neurons in classical neural networks (CNN).[1,2] When an input signal level is weak, the neuron linearly amplifies this input signal and thus compensates for unavoidable linear loss in the network. When the input signal level exceeds a certain threshold value a_{th}, however,

[1]Japan Science and Technology Agency, K's Gobancho 7, Gobancho, Chiyoda-ku, Tokyo 102-0076, Japan; [2]E. L. Ginzton Laboratory, Stanford University, Stanford, CA 94305, USA; [3]Institute of Industrial Science, The University of Tokyo, 4-6-1 Komaba, Meguro-ku, Tokyo 153-8505, Japan; [4]National Institute of Informatics, 2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo 101-8403, Japan; [5]Division of Electrical, Electronic and Information Engineering, Osaka University, 2-1 Yamadaoka, Suita, Osaka 565-0871, Japan and [6]NTT Basic Research Laboratories NTT Corporation, 3-1 Morinosato Wakamiya, Atsugi, Kanagawa 243-0198, Japan
Correspondence: Yoshihisa Yamamoto (yyamamoto@stanford.edu)

Received: 30 July 2017 Revised: 31 October 2017 Accepted: 3 November 2017
Published online: 15 December 2017

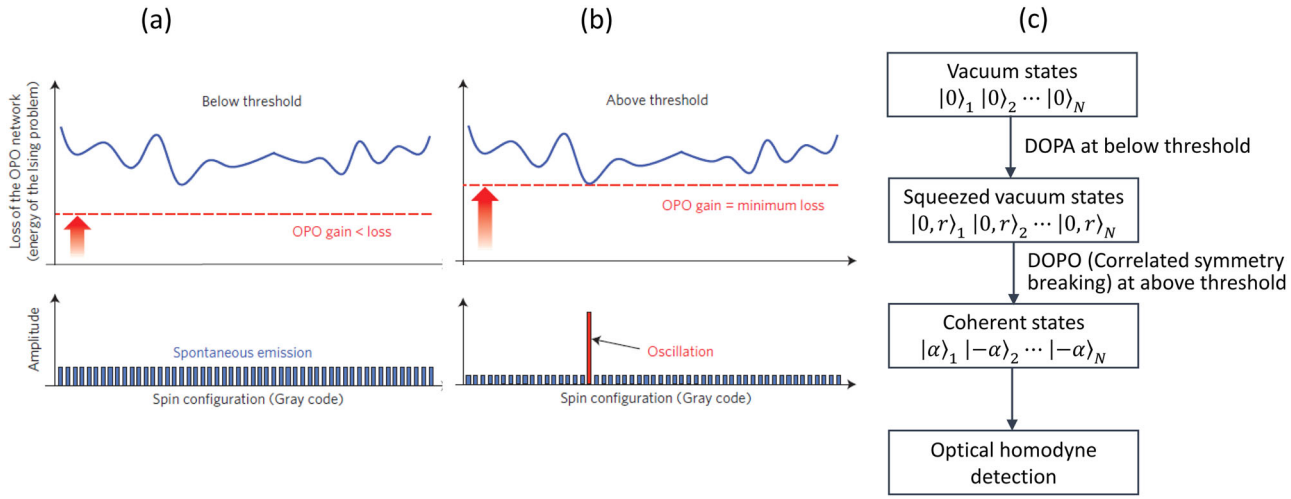

Fig. 1 Principle and sequence of operations of coherent Ising machines (CIM). **a**, **b** Principle of coherent Ising machines (CIM). When a DOPO network gain increases to reach the minimum loss rate of a ground state, a single-mode oscillation occurs at the ground state. **c** A block diagram of the CIM operation

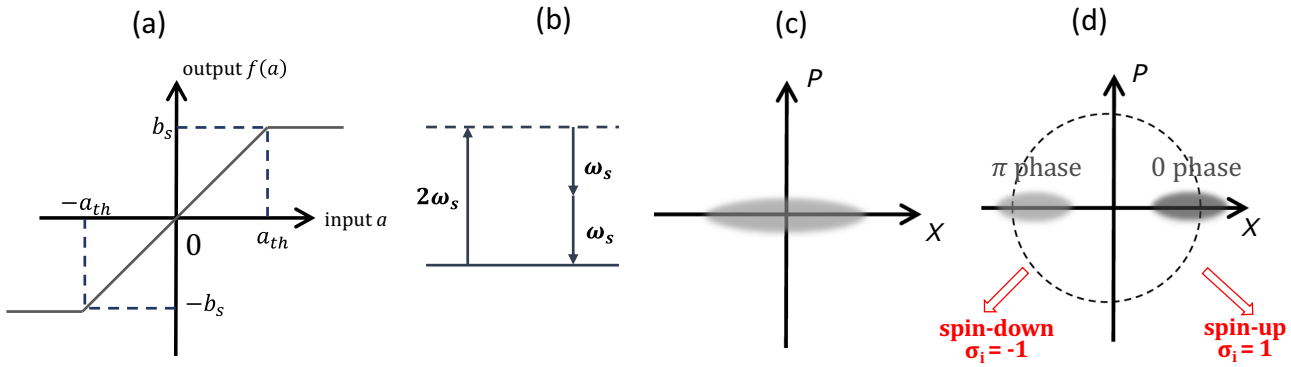

Fig. 2 Basic properties of quantum neurons. **a** The nonlinear input–output relation of a neuron in classical neural networks. **b** A degenerate optical parametric process. Quantum noise distributions of DOPA at below the threshold (**c**) and DOPO at above the threshold (**d**)

the output signal level is clamped at a constant value b_s. This nonlinear input–output relation is essential for CNN to find a stable operating point which manifests a solution for a given mathematical problem. In this model, each neuronal state is expressed by a continuous variable a_i, which obeys a continuous time evolution governed by

$$\frac{\mathrm{d}}{\mathrm{d}t}a_i = -a_i + f(a_i) - \frac{\partial V}{\partial a_i} + g_i, \tag{1}$$

where the first term of R.H.S. of Eq. (1) expresses the linear loss and the second term describes self-feedback with the nonlinear gain function f shown in Fig. 2a. Note that a time is normalized by the decay rate of a neuron excitation. The third term represents the mutual coupling among neurons and the interaction potential V implements a given mathematical problem. We emphasize that a simultaneous mutual coupling between neurons without introducing undesired instability or oscillation is made possible by the gradient descent character of the third term. The nonlinear gain function f is often implemented as a coupling coefficient for the third term of mutual coupling instead of the second term of self-feedback.[1,2] An essential feature of the CNN is, however, independent of the location of the nonlinear gain function. Finally, the fourth term, g_i, expresses a noise driving force associated with the decay and the nonlinear gain for a neuron excitation.

The classical picture behind Eq. (1) is that each neuron has an unpredictable but definite value a_i at a given time even though the network is randomly driven by internal and/or external noise. This is a correct physical picture of a neural network when thermal noise is much larger than quantum zero-point noise i.e. $k_BT \gg \hbar\omega$. However, if the opposite is true, i.e. $\hbar\omega \gg k_BT$, the interpretation of Eq. (1) must be altered. It should be considered as a c-number stochastic differential equation (CSDE), which is derived from and fully equivalent to the density operator master equation for the quantum neural network.[16,17] Now, the variable a_i does not represent any physical quantity of the neuron but rather it is merely an eigenvalue (α or β) of the coherent states used in the quasi-probability density to expand the density operator, such as a positive $P(\alpha,\beta)$ representation[16] or Wigner representation.[17] In order to compute the expectation value of any physical quantity of the neuron, we must first reconstruct the density operator by an ensemble of many trajectories of a_i generated by the CSDE. Whenever $\hbar\omega \gg k_BT$, we must take this quantum approach to reach a correct answer.

A quantum neuron (or quantum neural network) is characterized by the following three properties and in this way, it is distinct from a classical counterpart:

1. A quantum neuron is prepared in a superposition state of different neural excitations so that a quantum parallel search can be implemented.
2. A network of quantum neurons makes a decision to reach a final computational result through correlated and collective symmetry breaking at a critical point of phase transition.
3. A network of quantum neurons amplifies the above quantum solution to a classical signal via bosonic final state stimulation.

As shown in the subsequent sections, we will show the above three properties are indeed realized by DOPO and their networks.

The most important advantage of the DOPO-based quantum neural network compared to its classical counterpart is that each neuron (DOPO) is prepared in linear superposition of different in-phase amplitude eigenstates, i.e. squeezed vacuum state and so a quantum parallel search can be implemented to the whole optimization process. This is the crucial computational resource for CIM.

QUANTUM NEURONS

Degenerate optical parametric amplifiers and oscillators

Any phase-sensitive amplifier (PSA) has an input–output relation similar to that shown in Fig. 2a due to its inherent gain saturation effect. We focus here on the degenerate optical parametric amplifier (DOPA) based on the two photon emission process shown in Fig. 2b. A particular device is composed of a second-order nonlinear crystal placed in an optical cavity. The nonlinear crystal absorbs one pump photon at a frequency $2\omega_s$ and simultaneously emits two signal photons at a frequency ω_s. A pertinent interaction Hamiltonian[28] is expressed by

$$H = \hbar\kappa\left(\hat{a}_s^{\dagger 2}\hat{a}_p + \hat{a}_s^2\hat{a}_p^\dagger\right). \quad (2)$$

Here $\hat{a}_s(\hat{a}_s^\dagger)$ and $\hat{a}_p(\hat{a}_p^\dagger)$ are the annihilation (creation) operators for the signal and pump fields, and κ is a parametric coupling constant. The simultaneously generated two signal photons are a conjugate pair in a sense that they have a positive correlation in the in-phase amplitudes $\hat{X}(=(\hat{a}_s+\hat{a}_s^\dagger)/2)$ and a negative correlation in the quadrature-phase amplitudes $\hat{P}(=(\hat{a}_s-\hat{a}_s^\dagger)/2i)$. As a result of such positive and negative correlations in $\hat{X}$ and $\hat{P}$, constructive and destructive interference happens between the two signal photons along the X-axis and P-axis, respectively. As a result of such quantum interference, if a DOPA is input by an external vacuum state (zero-point fluctuation), the zero-point fluctuation is amplified and deamplified along the X-axis and P-axis, respectively, as shown in Fig. 2c. The resulting state is called a squeezed vacuum state, which is a minimum uncertainty wave packet and satisfies the Heisenberg uncertainty principle with equality, $\langle\Delta\hat{X}^2\rangle\langle\Delta\hat{P}^2\rangle = 1/16$, just as the vacuum state.

The phase-sensitive amplification/deamplification mentioned above is not a rare phenomenon in nature. One classical example is a swing driven by a person as shown in Fig. 3a, where a person (corresponding to the pump in DOPA) completes a full cycle (up–down–up) while a swing (corresponding to the signal in DOPA) makes only a half cycle (left-to-right). Note that the phase of the pump and that of the signal should be mutually locked in order to realize an amplitude amplification process, as shown in Fig. 3b. This corresponds to the situation that the in-phase amplitude $\hat{X}$ is amplified. You can easily imagine from your childhood memory what you did in order to stop the swing when you must go home. You stood up at the center of the swing

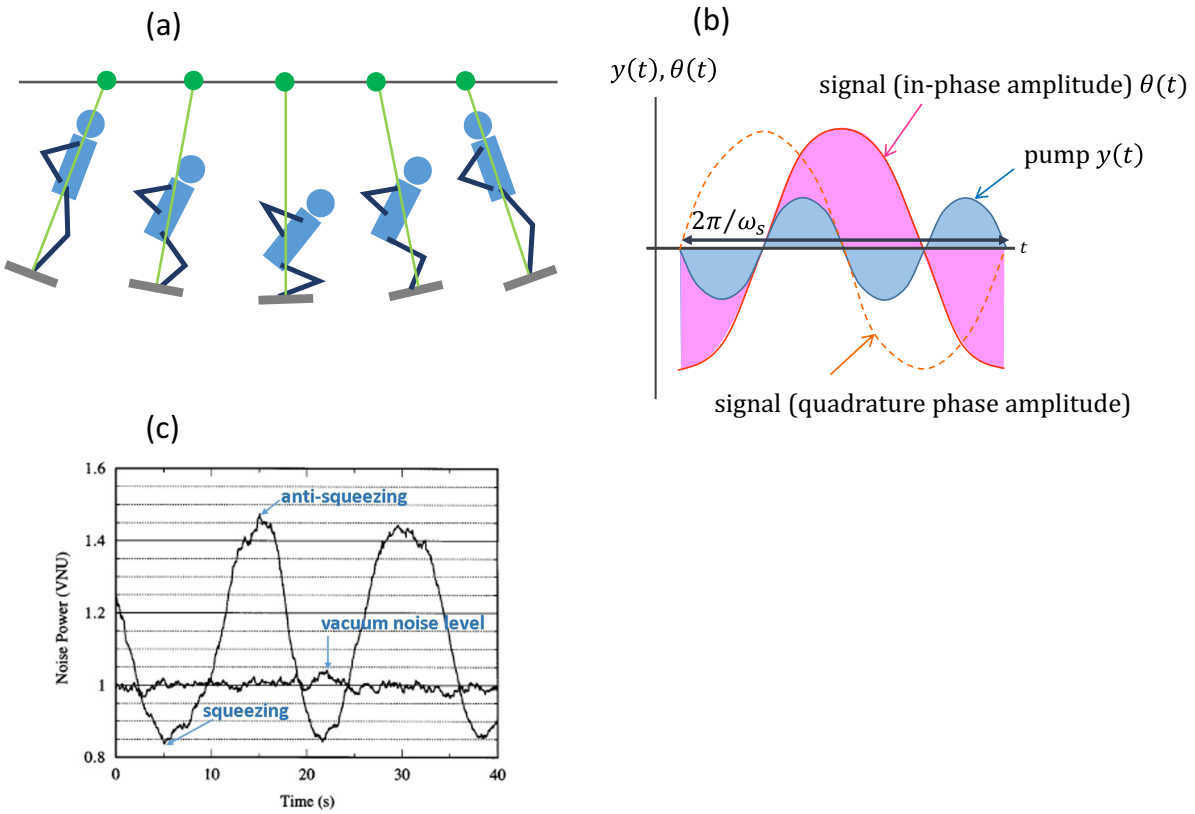

Fig. 3 A phase sensitive amplifier. **a** A phase-sensitive amplification/deamplification process of a swing, in which a person makes a full cycle (up–down–up) but a swing makes only half cycle (left-to-right). **b** The corresponding phase relation between the person's drive $y(t)$ and the swing oscillation $\theta(t)$. **c** The in-phase and quadrature-phase amplitude noise of the DOPA at below threshold[30]

period and crouched at both ends. Then, the swing (signal) amplitude is attenuated. This corresponds to the situation that the quadrature-phase amplitude $\hat{P}$ is deamplified.

The OPO has a long history of development at Stanford University.[29] In the coherent Ising machine (CIM), a periodically polled lithium-niobate (PPLN) waveguide device is used as such a PSA for signal pulses. The first experimental demonstration of deamplification (squeezing) and amplification (anti-squeezing) for optical pulses in vacuum states using this particular device was reported in 1995.[30] Figure 3c shows the suppressed quadrature-phase amplitude noise (squeezing) and enhanced in-phase amplitude noise (anti-squeezing) by the PPLN waveguide DOPA. It is a general property of the squeezing devices that the degree of squeezing (vacuum noise deamplification) is reduced by experimental system's imperfections, in particular, a linear optical loss, but that the degree of anti-squeezing (vacuum noise amplification) is not reduced so much by linear optical loss. Suppose a PSA output has a 20 dB of squeezing, $4\langle\Delta\hat{P}^2_{\rm in}\rangle = 10^{-2}$, for the quadrature-phase amplitude and 20 dB of anti-squeezing, $4\langle\Delta\hat{X}^2_{\rm in}\rangle = 10^2$, for the in-phase amplitude. An optical linear loss L degrades the degree of squeezing and anti-squeezing according to

$$\langle\Delta\hat{X}^2_{\rm out}\rangle = L\langle\Delta\hat{X}^2_{\rm in}\rangle + (1-L)\frac{1}{4},\ \langle\Delta\hat{P}^2_{\rm out}\rangle = L\langle\Delta\hat{P}^2_{\rm in}\rangle + (1-L)\frac{1}{4}. \tag{3}$$

If there is a 3 dB of linear loss after the PSA, the degree of squeezing is only ~3 dB while the degree of anti-squeezing is still ~17 dB. Note the Ising spin is represented by the in-phase amplitude rather than the quadrature-phase amplitude. Therefore, the quantum parallel search in the CIM is based on the anti-squeezed quantum noise of the in-phase amplitude rather than the squeezed quantum noise of the quadrature-phase amplitude, so that the operation of CIM is inherently robust against an optical loss.

The input–output relation of the DOPA for an increasing signal power above the saturation level approximately reproduces the nonlinear response function $f(a)$ shown in Fig. 2a.[31] A physical mechanism behind this gain saturation is the depletion of the pump power, which triggers a reverse energy flow, i.e. two signal photons are absorbed simultaneously in the nonlinear crystal to generate one pump photon. When the input signal power becomes sufficiently large, the reverse energy flow from the signal to the pump is switched on and the linear amplification process must stop. In our case of constructing quantum neurons, a related but slightly different nonlinear effect is employed to amplify a solution state at a microscopic level to that at a macroscopic level, which we will explain next.

If the parametric amplifier gain exceeds the signal decay rate from the cavity, the system can sustain a steady-state field by creating a finite average amplitude along the X-direction. The center position of a quantum wave packet in the (X,P) phase space is no more $X = P = 0$ but there emerges a non-zero X value, as shown in Fig. 2d. This is the oscillation threshold point. The device pumped at above the oscillation threshold is called a DOPO. The probability for creating a positive X value and negative X value is randomly selected by 50–50% probability, which is known as spontaneous symmetry breaking and is an ubiquitous feature for any second-order phase transition phenomena.[32] In the language of dynamical systems theory, the DOPO threshold represents the supercritical pitchfork bifurcation. The two stable states with positive and negative X values correspond to the firing and non-firing states of a classical neuron or up-spin and down-spin states of an Ising problem. A similar mechanism, dissipative quantum bifurcation machine, was discussed recently in the context of Josephson junction qubits.[33–35] When the input signal level to the DOPA is varied while the pump rate is fixed at below the oscillation threshold, the output signal level is first linearly amplified and then eventually reaches the oscillation state with a clamped amplitude due to gain saturation. In this way, the nonlinear input/output relation shown in Fig. 2a is realized in the transition from DOPA to DOPO in our quantum neurons.

Why are DOPOs referred to as quantum neurons and how are they different from classical neurons? We will present the answers to these important questions in the next section.

Linear superposition

A set of photon number eigenstates $|n\rangle$ can expand an arbitrary state of the field as an orthonormal set. The wavefunction of a squeezed vacuum state (see Fig. 2c) can be mathematically constructed as a superposition of photon number eigenstates with even eigenvalues:

$$|\psi_{\rm sv}\rangle = c_0|0\rangle + c_2|2\rangle + c_4|4\rangle + \cdots. \tag{4}$$

A simple physical reasoning behind Eq. (4) is explained as follows: The photon number of an intense pump field exhibits large quantum uncertainty, typically on the order of $\langle\Delta\hat{n}^2\rangle \sim \langle\hat{n}\rangle$ for a coherent pump field. Therefore, we cannot extract which-path information, even in principle, whether the number of absorbed pump photons in the nonlinear crystal is zero, one, two … at a given time. Because one pump photon is converted into two signal photons, those cases correspond to the output signal field in $|0\rangle$, $|2\rangle$, $|4\rangle$, $\cdots$ photon number eigenstates. Because of the lack of which-path information, the proper expression for the quantum state of the output signal field of DOPA must be a superposition of those states. Except for an irrelevant phase factor, the probability amplitudes c_0,c_2,c_4,$\cdots$ have an identical phase, for instance, positive real numbers.[36] In the (X, P) phase space, the different photon number eigenstates constructively interfere with each other along the X-axis, while they destructively interfere with each other along the P-axis.

Alternatively, a squeezed vacuum state can be mathematically constructed as a superposition of quadrature amplitude eigenstates $|X\rangle$, which are the eigenstates of the Hermitian operator $\hat{X}$:

$$|\psi_{\rm sv}\rangle = \int_{-\infty}^{\infty} C(X)|X\rangle {\rm d}X. \tag{5}$$

It is not difficult to show that the different in-phase amplitude eigenstates constructively interfere with each other in the small P-regions centered at $P = 0$, while they destructively interfere with each other in the large P-regions. Therefore, more anti-squeezing (enhanced quantum noise) along the X-axis realizes more squeezing (reduced quantum noise) along the P-axis.

It is worth pointing out that a statistical mixture of in-phase amplitude eigenstates

$$\hat{\rho}_{\rm mix} = \int_{-\infty}^{\infty} P(X)|X\rangle\langle X|{\rm d}X \tag{6}$$

instead of Eq. (5), covers the same range of the eigenvalues X but requires an infinite amount of energy to realize, since the localization of the state along the P-axis near $P = 0$ is now impossible due to the lack of phase coherence between different $|X\rangle$ eigenstates. The infinite uncertainty along the P-axis means an infinite energy. A squeezed vacuum state, represented by Eq. (5) and produced by the DOPA, allows a quantum parallel search with a finite energy, while a classical mixed state, represented by Eq. (6), does not allow such a quantum parallel search.

A particularly unique quantum feature of the DOPA/DOPO is that the aforementioned superposition survives not only below the threshold but also above the threshold to some extent if the cavity loss is small. At well above the oscillation threshold, the DOPO produces either 0-phase or π-phase coherent field as shown in Fig. 2d. However, a DOPO at just above the oscillation threshold maintains the superposition of 0-phase and π-phase states, in spite of finite photon leakage from the resonator.[37–39] This is

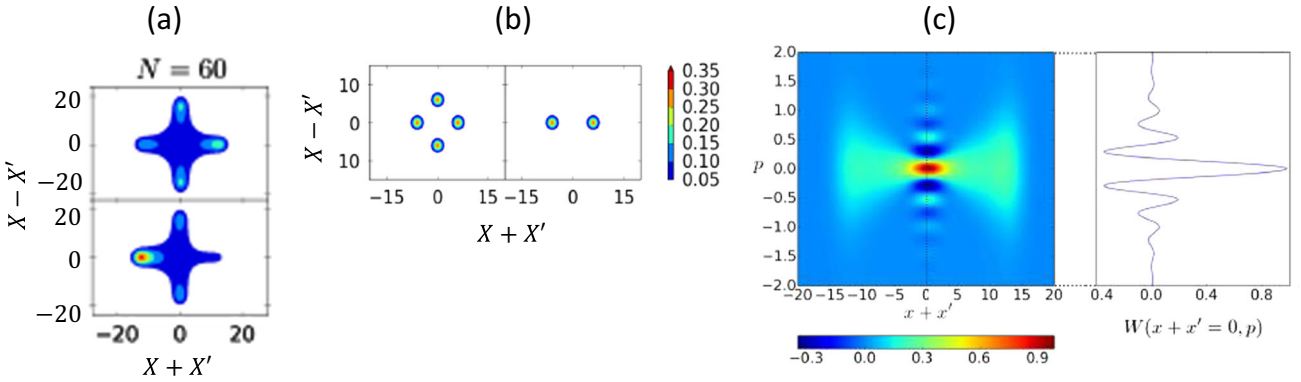

Fig. 4 Quantum states of DOPO fields at the threshold. **a** The density matrix elements $\langle X|\hat{\rho}|X'\rangle$ of the two out-of-phase coupled DOPO pulses in a high-Q ring cavity at the threshold pump rate. The in-phase amplitude probability distribution $\langle X|\hat{\rho}|X\rangle$ and the quantum coherence $\langle X|\hat{\rho}|-X\rangle$ of the two DOPO pulses are given by a horizontal cut at X–$X'=0$ and a vertical cut at X+$X'=0$, respectively. The round trip loss is 0.1%. **b** The density matrix elements $\langle X|\hat{\rho}|X'\rangle$ for a superposition state, $1/\sqrt{2}(|\alpha\rangle+|-\alpha\rangle)$ and a mixed state, $\hat{\rho}=1/2(|\alpha\rangle\langle\alpha|+|-\alpha\rangle\langle-\alpha|)$. (**c**) The corresponding Wigner functions to Fig. 4a[40]

possible because a hypothetical measurement performed for a leaked signal field cannot identify which phase the DOPO selects due to the enhanced quantum noise along the X-axis and also small separation between the two states. The anti-squeezed in-phase amplitude noise realizes a so-called "quantum erasure" for this hypothetical which-path measurement.

The evidence for the above statement is shown in Fig. 4a.[40] Here we consider a CIM consisting of synchronously pumped DOPO pulses circulating in a fiber ring cavity.[24,25] The density matrix elements $\langle X|\hat{\rho}|X'\rangle$ in the (X+X', X–X') phase space for two out-of-phase (anti-ferromagnetically) coupled DOPOs are plotted, where the pump rate is just above the threshold. In Fig. 4a, the in-phase amplitude probability distribution $\langle X|\hat{\rho}|X\rangle$, which is given as the horizontal slice at X–$X'=0$, and the quantum coherence $\langle X|\hat{\rho}|-X\rangle$ between the two in-phase amplitude eigenstates $|X\rangle$ and $|-X\rangle$, which is given as the vertical slice at X+$X'=0$, clearly show macroscopically separated coherent states $|\alpha\rangle$ and $|-\alpha\rangle$ maintain quantum coherence.[40] Figure 4b compares the density matrix elements $\langle X|\hat{\rho}|X'\rangle$ for a superposition state, $1/\sqrt{2}(|\alpha\rangle+|-\alpha\rangle)$, and a mixed state, $\hat{\rho}=1/2(|\alpha\rangle\langle\alpha|+|-\alpha\rangle\langle-\alpha|)$, in the same ($X$+$X'$, X–X') phase space, respectively. We can conclude from Fig. 4a that the two DOPOs are approximately in Schrödinger's cat states and yet their centers of gravity are negatively correlated due to out-of-phase coupling. That is, the DOPO1 (above) is biased toward a positive amplitude, while the DOPO2 (below) is biased toward a negative amplitude. Figure 4c shows the Wigner function of the DOPO state. The oscillatory behavior and negative values of the Wigner function are another manifestation of the quantum coherence between $|\alpha\rangle$ and $|-\alpha\rangle$ states.[40] Such a highly non-classical feature is quickly destroyed by increasing a cavity loss. However, it should be emphasized that the superposition of $|X\rangle$ eigenstates with positive and negative eigenvalues in a squeezed vacuum state, represented by Eq. (4), is rather robust against a cavity loss as mentioned above so that the resulting quantum parallel search survive in a low-Q cavity.

Amplitude and phase error correction capability given by phase-sensitive amplification

Fluctuations are induced in the DOPO fields by the external phase and amplitude noise injection. Both the amplitude and phase of the DOPO field are continuous variables so that the standard error detection and error correction techniques cannot be applied. Fortunately, the phase-sensitive amplification/deamplification mechanism of DOPO can stabilize the phase to either 0 or π as

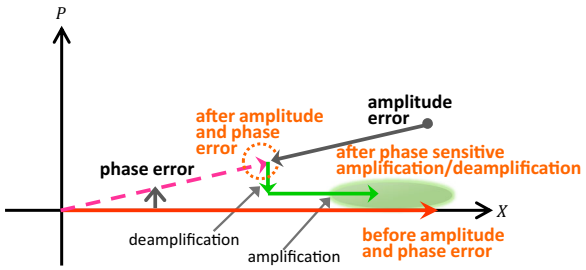

Fig. 5 Amplitude and phase error correction in DOPO. The amplitude error is corrected by saturated amplification along the X-axis and the phase error is corrected by phase-sensitive deamplification along the P-axis

long as the phase error is small compared to $\pm\pi/2$.[15] Moreover, the amplitude error can be also suppressed by the gain saturation mechanism of the DOPO. If the signal amplitude is increased to above the steady-state value, the pump amplitude is more strongly depleted which, in turn, recovers the steady-state amplitude through a reduced parametric gain. The opposite is true when the signal amplitude is decreased to below the steady-state value. The deterministic amplitude decay due to optical linear loss can be compensated for by the average gain. In this way, the output amplitude is stabilized through the balance between the linear cavity loss and the saturated gain in DOPO, while the phase is stabilized by the phase-sensitive deamplification of DOPO. This is schematically shown in Fig. 5.

The DOPO provides a unique opportunity as a robust analog memory against external noise injection. We can stably store the analog information at quantum limited accuracy with the above-mentioned mechanisms.

QUANTUM SYNAPSES

We wish to implement a target Hamiltonian as a cost function (effective loss) of the network consisting of quantum neurons. Our target Hamiltonian is the following Ising model:[41]

$$H=-\sum_{i<j}J_{ij}\sigma_i\sigma_j-\sum_i h_i\sigma_i, \tag{7}$$

where the Ising spin σ_i takes either +1 (up-spin) or −1 (down-spin). As described above, we wish to represent $\sigma_i=1$ or $\sigma_i=-1$ by the

positive or negative in-phase amplitude, which corresponds to 0-phase or π-phase oscillation, respectively (Fig. 2d). The Ising coupling J_{ij} and the local field h_i take continuous (real) values, the magnitude of which are determined by mapping a given combinatorial optimization problem on the Ising model.[13,21] The three-dimensional Ising model and the two-dimensional Ising model with local fields belong to the NP-hard class in complexity theory.[18] Therefore, you can imagine that many hard problems in the real world can be solved through the Ising model.[13,21] In order to implement the cost function described in Eq. (7) as the effective loss of the DOPO network, pairs of DOPOs must be coupled with the coupling constant J_{ij} and also the constant optical field h_i must be injected into each DOPO.[14]

Optical delay line coupling scheme

In the configuration shown in Fig. 6a, N independent DOPOs are simultaneously realized as N optical pulses circulating in a single fiber ring cavity with an internal PSA which is driven externally by pump pulse trains. Figure 6a shows one experimental scheme to implement the Ising coupling J_{ij} in the DOPO network, where a part of each DOPO pulse circulating in a fiber ring resonator is picked-off at every round trip by the output coupler, amplified by an external PSA, split into multiple optical delay lines including intensity and phase modulators and then injected back to the target DOPO pulse at appropriate timing.[15,22,23]

The external PSA is provided by another DOPA and boosts the in-phase amplitude $\hat{X}$ of a picked-off pulse before it is attenuated by the beam splitter and modulator loss. Using this technique, we can implement a relatively large Ising coupling constant ($J_{ij} \lesssim 1$) even if the system size is fairly large ($N \gtrsim 10^4$). Using such $N-1$ optical delay lines, any (jth) pulse can be connected to any other (ith) pulse with a coupling constant J_{ij}. Please note that asymmetric coupling $J_{ij} \neq J_{ji}$ can be easily implemented in this scheme and that $\sim N^2$ mutual coupling constants can be realized by ($N-1$) optical delay lines. Such an all-optical coupling scheme has been experimentally demonstrated for $N = 4$ and 16 pulses using free-space optics with a 4.8 m ring cavity and 250 MHz and 1 GHz repetition frequencies, respectively[15,22] and for $N = 10^4$ pulses using a planar lightwave circuit and fiber optics with 1 km fiber ring cavity and 2 GHz repetition frequency.[23] In the latter system, for example, the optical fiber length is 1 km so that a round trip time is 5 μs while the pulse interval is 0.5 ns, which leads to $N = 10^4$ DOPO pulses inside a single ring cavity.

This optical delay line coupling machine is capable of producing the quantum noise correlation among the internal

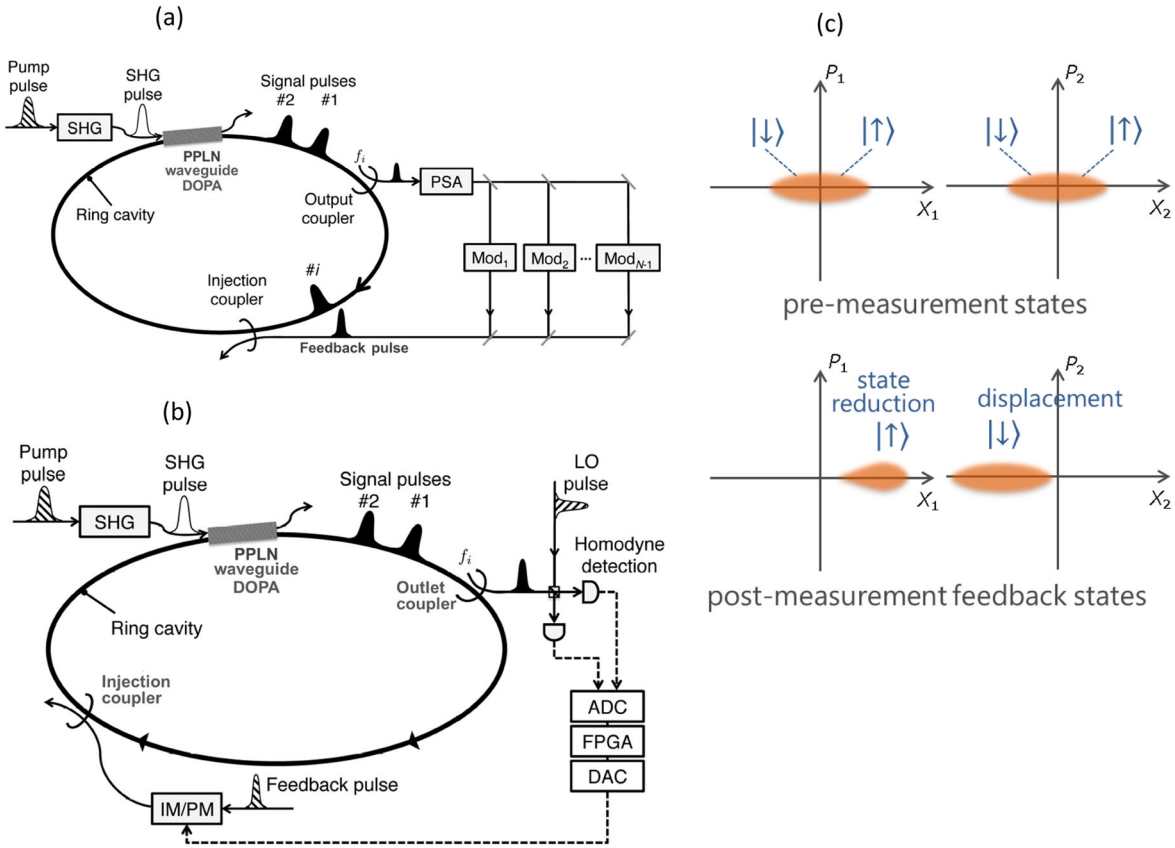

Fig. 6 Two types of coherent Ising machines (CIM). **a** A coherent Ising machine (CIM) based on the time-division multiplexed DOPO pulses with mutual coupling implemented by optical delay lines. A part of each pulse is picked off from the main cavity by the output coupler followed by an optical phase-sensitive amplifier (PSA) that amplifies the in-phase amplitude $\hat{X}$ of the extracted DOPO pulse. The feedback pulses, which are produced by combining the outputs from $N-1$ intensity and phase modulators, are injected back to the target DOPO pulse by the injection coupler.[15,22,23] **b** A CIM with a measurement-feedback circuit. A small portion of each DOPO signal pulse is out-coupled through the output coupler, and its in-phase amplitude is measured by optical balanced homodyne detectors, where LO pulse is directly obtained from the pulsed pump laser. Two detector outputs are converted to digital signals and input into an electronic digital circuit, where a feedback signal for the ith DOPO signal pulse is computed. The feedback pulse also taken from the pump laser is modulated in its intensity and phase to achieve the target amplitude proportional to $\sum_j J_{ij}\tilde{X}_j$ and coupled into the ith signal pulse by an injection coupler. Flows of optical fields and electrical signals are shown as solid and dashed lines, respectively.[24,25] **c** A partially reduced DOPO1 wave packet by the approximate measurement and a displaced DOPO2 wave packet by the feedback injection implements an anti-ferromagnetic order even after a single measurement-feedback

DOPO pluses.[16,17] In the section Quantum features of CIM, we will discuss this point in detail using the two measures for evaluating quantum noise correlation: quantum entanglement and discord.

Measurement-feedback coupling scheme

An alternative coupling scheme to implement the Ising coupling J_{ij} is shown in Fig. 6b.[24,25] Instead of directly connecting the DOPO pulses with optical delay lines, we can measure approximately the in-phase amplitude of the internal DOPO pulse by the optical balanced homodyne detectors. If the inferred in-phase amplitude of the jth DOPO pulse is represented by $\tilde{X}_j$, the feedback pulse to the ith DOPO pulse should have an in-phase amplitude proportional to $\sum_j J_{ij}\tilde{X}_j$. The complicated task of the synchronous computation of the vector-vector multiplication between J_{ij} and $\tilde{X}_j$, which must be completed in the pulse interval of 100 ps – 1 ns, is achieved by a single measurement-feedback circuit consisting of an analog-to-digital converter, a field programmable gate array (FPGA), a digital-to-analog converter and optical amplitude/phase modulators.[24,25] The feedback pulse used as an input to the optical modulator and the local oscillator pulse used for optical homodyne detection are both provided by a part of the pump laser output, as shown in Fig. 6b.

Such a measurement-feedback coupling scheme (Fig. 6b) is equivalent to an optical delay line coupling scheme (Fig. 6a) except for the following advantage/disadvantage. The advantage of the measurement-feedback scheme is that all-to-all coupling of the order of $\sim N^2$ connections can be implemented by a single measurement-feedback circuit, so that the daunting task of constructing $N-1$ optical delay lines and stabilizing their delay lengths (or optical phase) with an error much less than the optical wavelength can be avoided. In addition, many-body Ising-type interactions, such as $H = -\sum K_{ijk}\sigma_i\sigma_j\sigma_k$, can be readily implemented. The disadvantage of the measurement-feedback scheme is that the FPGA circuit must complete ~O(N) vector–vector multiplication and addition within each pulse interval. This imposes a limitation on the machine size N and the pulse repetition frequency. On the other hand, the optical delay line coupling scheme enjoys its inherent high-speed operation with a pulse repetition frequency limited only by optical device performance.

There is a subtle but important difference in the operational principles from quantum mechanical viewpoints between the optical delay line coupling and measurement-feedback coupling schemes. The measurement-feedback scheme does not produce any entanglement among DOPO pulses. The total density operator stays in the product state of individual DOPO pulse density operators, because the coupling is provided by the local operation and classical communications (LOCC). However, the measurement-feedback scheme is capable of implementing the non-unitary state reduction to pump out a spurious entropy from the network and prepare each DOPO pulse close to the Heisenberg limit (minimum uncertainty wave packet) induced by quantum measurements. The wave packet is actually non-Gaussian at above threshold, which contributes to enhanced quantum tunneling in the quantum parallel search process.

Figure 6c shows schematically this scenario. Suppose the two DOPOs are connected by the anti-ferromagnetic Ising coupling constant $J_{21} < 0$. At the pump rate near threshold and before measurement feedback, the two DOPOs are independently in squeezed vacuum states. If the DOPO1 is measured and a positive amplitude is obtained ($\tilde{X}_1 > 0$), the center position of the DOPO1 wave packet is shifted to the positive side and the variance is reduced, but a long tail exists toward $X_1 = 0$ and $X_1 < 0$ (non-Gaussian). Simultaneously, the center position of the DOPO2 wave packet is shifted to the negative side by injecting a feedback pulse with an amplitude of $J_{21}\tilde{X}_1$. In this way, the anti-ferromagnetic order is implemented in the two quantum states even with a single measurement-feedback step.

MAPPING OF AN ISING HAMILTONIAN TO DOPO NETWORK LOSS

In this section, the in-phase amplitude $x_j(t)$ of the jth DOPO is considered as a classical variable. The simple classical picture presented in this section still captures an essential part of the mapping principle.

Pitchfork bifurcation

We consider in this section simple equations of motion that capture the essential computational principle for the sake of clarity. We consider the dynamics of N analog variables x_j, with $x_j \in]-\infty, +\infty[$ and $j \in \{1,\cdots, N\}$, given as follows:

$$\dot{x}_j(t) = -\frac{\partial V}{\partial x_j}, \tag{8}$$

with

$$V = \sum_l \left[V_b(x_j)\right] + \epsilon V_l(x), \tag{9}$$

where $V_b(x_j)$ is the archetype bistable potential in x_j with $V_b(x_j) = -\frac{1}{2}\alpha x_j^2 + \frac{1}{4}x_j^4$ and $V_l(x)$ the analog version of the Ising Hamiltonian with $V_l(x) = -\sum_j \omega_{jl} x_j x_l$. Here $t = (\gamma_s/2)\tau$ is the unitless time normalized by the signal field lifetime $2/\gamma_s$, γ_s is the signal photon decay rate, $p = \alpha + 1 = F_p/F_{th}$ is the unitless pump rate normalized by the solitary DOPO threshold pump photon flux $F_{th} = \gamma_s\sqrt{\gamma_p}/4\kappa$, γ_p is the pump photon decay rate, κ is the parametric coupling constant, $x_j = X_j/A_s$ is the normalized in-phase amplitude, and $A_s = \sqrt{\gamma_s\gamma_p/2\kappa^2}$ is the saturation amplitude which is equal to the actual DOPO oscillation amplitude at the normalized pump rate $p = 2$. $\alpha = -1 + p$ is the bifurcation parameter given by the normalized decay rate (−1) and linear gain (p) for the signal field. Moreover, the parameter ϵ is small enough such that $0 < \epsilon \ll 1$. Note that the Ising coupling constant ω_{jl} can be arbitrarily chosen under this constraint on the parameter ε. Equation (8) describes the dynamics of a gradient descent system. Note that the equations of motion for a solitary DOPO are given as $\dot{x}_j = \alpha x_j - x_j^3$ when $\varepsilon = 0$, which represents the normal form of the supercritical pitchfork bifurcation.

In the uncoupled case, i.e., $\varepsilon = 0$, the potential V_b is monostable with $x_j = 0$ and bistable with $x_j = \pm\sqrt{\alpha}$ when $\alpha < 0$ and $\alpha > 0$, respectively. In the coupled case when $0 < \epsilon \ll 1$ and α is large enough, the state of each variable x_j can be considered as approximately binary and can be mapped to the Ising spin. In particular, the Ising spin configuration $\{\sigma_j\}_j$ is associated to the state $\{x_j\}_j$ of the system described in Eq. (8) by considering that σ_j is equal to the sign of x_j, i.e., $\sigma_j = x_j/|x_j|$.

When α increases over a critical threshold value, there is a breaking of symmetry and the states x_j become either positive or negative. The emergence of the first non-zero stable steady state can be studied by considering linear stability analysis at the origin, i.e., the Jacobian matrix J associated with Eq. (8) and computed at the state $x_j = 0$, $\forall_{ij}$. This Jacobian matrix is given as follows:

$$J = \alpha I + \epsilon\Omega, \tag{10}$$

where Ω is the matrix with components ω_{jl} and I the identity matrix.[14,42] The destabilization of the zero state corresponds to the emergence of the first non-zero stable steady state after the symmetry breaking and occurs when the maximal eigenvalue of the Jacobian matrix J, noted λ_1, is equal to zero. The eigenvalue λ_1 is given as $\lambda_1 = \alpha + \epsilon\mu_1$, where $\mu_1(>0)$ is the maximal eigenvalue of the matrix Ω, which we consider unique for the sake of simplicity. The first non-zero steady state becomes stable when λ_1

= 0, i.e., when the bifurcation parameter a is equal to a_c given as follows:[14,42]

$$a_c = -\epsilon\mu_1. \tag{11}$$

As the bifurcation parameter a increases further above the threshold a_c, other states become stable and the system develops a complex attractor landscape. Note that Eq. (11) suggests that $a_c < 0$ in general.

Condition for successful mapping

The linear stability analysis of the zero state allows determining the bifurcation parameter a_c for which the zero state becomes unstable. Moreover, the eigenvector associated with the eigenvalue λ_1 indicates the locations of the first non-zero stable steady states just after the bifurcation has occurred but, in general case, these locations do not correspond to the ground-state configurations of the Ising Hamiltonian. However, it can be noticed that, under the particular condition that all analog amplitudes $|x_j|$ are equal, the first non-zero stable steady states are mapped rigorously to the ground-state configurations of the Ising Hamiltonian.[42] Indeed, the steady-state condition of Eq. (8) is given as follows when $x_j = \sigma_j x$, $\forall_{ij}$, with $x > 0$:

$$\begin{aligned} \dot{x}_j = 0 &\Rightarrow a x_j - x_j^3 + \epsilon \sum_{l\neq j} \omega_{jl} x_l = 0 \\ &\Rightarrow a - x^2 + \epsilon \sum_{l\neq j} \omega_{jl}\sigma_l\sigma_j = 0. \end{aligned} \tag{12}$$

Thus, the homogeneous squared amplitude x^2 can be written as follows:

$$x^2 = a - \frac{2\epsilon}{N} H, \tag{13}$$

with the Ising Hamiltonian H without a dc field. When $a - \frac{2\epsilon}{N}H < 0$, the only solution of Eq. (12) is $x = 0$. As a increases, the first non-zero steady state appears at the minimal value of the function $\frac{2\epsilon}{N}H$ as shown in Eq. (13). As shown in Fig. 1a, b, the ground-state configurations of the Ising Hamiltonian appear as a single oscillation mode. In a nutshell, the first non-zero stable steady states of Eq. (12) encode for the ground-state configurations of the Ising Hamiltonian if the steady-state amplitudes $|x_j|, \forall j$, are all equal.

Effect of the amplitude heterogeneity

In more general case, the amplitudes $|x_j|$ are not equal and the mapping between the first stable steady-state configurations and the minima of the Ising Hamiltonian does not hold.[14] In practice, approximate solutions to various combinatorial optimization problems can still be obtained under this approximate mapping.[14,42] In order to quantify the distance between the case of exact mapping and more general case, i.e., when amplitudes are homogeneous and heterogeneous, respectively, we quantify the amplitude heterogeneity using the quantity δ_x defined as follows:

$$\delta_x = \frac{\sqrt{\langle \delta_j^2 \rangle}}{\langle x_j^2 \rangle}, \tag{14}$$

with $\delta_j = x_j^2 - \langle x^2 \rangle$ and $\langle x^2 \rangle = \frac{1}{N}\sum_j x_j^2$. The quantity δ_x corresponds to the coefficient of variation of the squared amplitudes. For $a > 0$ and $\epsilon \ll 1$, the amplitudes x_j can be expressed as $x_j = x_j^{(0)} + \epsilon x_j^{(1)} + \mathrm{O}(\epsilon^2)$ with $x_j^{(0)} = \sigma_j\sqrt{a}$ and $x_j^{(1)} = (1/2\sqrt{a})\sum_l \omega_{jl}\sigma_l$ using the perturbation theory in order to find an approximate solution to the steady-state equation (12). Thus, the squared amplitudes can be expressed as follows in the general case:

$$\begin{aligned} x_j^2 &= \left[x_j^{(0)} + \epsilon x_j^{(1)} + O(\epsilon^2)\right]^2 \\ &= a + \epsilon \sum_l \omega_{jl}\sigma_j\sigma_i + O(\epsilon^2). \end{aligned} \tag{15}$$

Using Eq. (15), we can express x_j^2 as follows:

$$x_j^2 = \left\langle x_j^2 \right\rangle + \delta_j + O(\epsilon^2), \tag{16}$$

with $\delta_j = \epsilon\left(\sigma_j h_j - \left\langle \sigma_j h_j \right\rangle\right)$, $h_j = \sum_{l\neq j}\omega_{jl}\sigma_l$, and $\langle x^2 \rangle = a - \frac{2\epsilon}{N}H$. We can remark that the homogeneous case, for which there exists exact mapping between the first non-zero stable steady states and the ground-state configurations of the Ising Hamiltonian, is the limit of the general case when $\delta_j \to 0$. Thus, the approximate mapping of the Ising Hamiltonian results from the variations of the squared amplitudes δ_j.

Experimentally, $\delta_j \to 0$ can be achived by additional measurement-feedback control. That is, CIM measures x_j^2 at each round trip and computes $\delta_j = x_j^2 - a + \frac{2\epsilon}{N}H$. This error signal is negatively fed back to the jth pump pulse or directly to the jth DOPO pulse to suppress δ_j.

Figure 7a shows the modulated threshold gain a_c due to the mutual coupling ω_{ij} and the success probability P_s of finding a ground state vs. the normalized pump rate p for an $N = 8$ Ising problem, in which each vertex has three edges and there is no Zeeman term (MAX-CUT-3 problem).[14] In this problem, each vertex has three edges with the constant weight $\omega_{ij} = -0.1$ as shown in Fig. 7b. At high pump rates $p \geq 1.3$, the possible value of a_c is minimum for the ground state and increases monotonically with the order of excited states. That is, the mapping is successful. However, the threshold gain a_c of the first excited state decreases to below that of the ground state at $p < 1$. This unexpected result stems from the heterogeneity of the DOPO amplitudes. Figure 7b shows the steady-state amplitude (by circle size) and phase (by color) of each DOPO for the first excited state at $p \lesssim 1$. Five of the total eight DOPOs mutually couple without destructive interference so that they can store large amplitudes. On the other hand, the remaining three DOPOs mutually couple with destructive interference. By keeping amplitudes in these three DOPOs small, the whole network realizes an effective field decay rate even smaller than that of the ground state. Consequently, the first excited state oscillates at a smaller pump power than the ground state, which leads to the malfunction of the DOPO network.

At above the threshold ($p \geq 1$), the first excited state is metastable so that the oscillation at the first excited state switches to the oscillation at the ground state by quantum tunneling with a finite lifetime.[43] This is why the success probability P_s increases from $p = 1$ to $p = 1.3$. However, if the pump rate is too high ($p > 1.3$), the potential barrier becomes too high to tunnel from a local minimum to a global minimum, which stabilizes the false oscillation mode at the first excited state. This is the reason why P_s decreases at $p > 1.3$ in Fig. 7a.

We remark that a technique to ensure equal amplitudes of DOPOs is actually realized by using an error detection and correction method so that the ground state of the Ising Hamiltonian is selected by a single oscillation mode.

EXPERIMENTAL CIMS

Optical delay line coupling CIM

The first experimental CIM implements $N = 4$ Ising spins with a free-space multiple-pulse DOPO and achieves all-to-all connections with N–1(=3) optical delay lines.[15] The Ising coupling constants are uniform and anti-ferromagnetic type ($J_{ij} < 0$). When the pump power is increased gradually, the machine finds one of

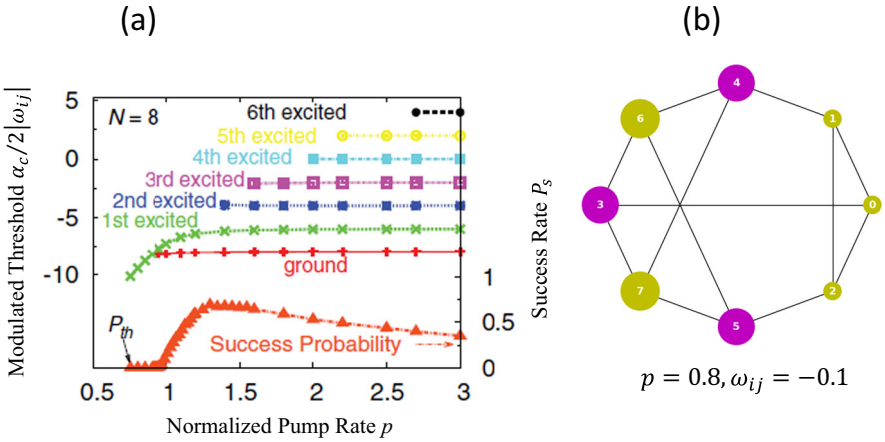

Fig. 7 The DOPO network with amplitude heterogeneity. **a** The modulated threshold gain α_c of several modes and the success probability P_s vs. the normalized pump rate p when $\omega_{ij} = -0.1$ for the worst-case instance with $N = 8$ vertices. **b** The inhomogeneous oscillation amplitudes for the first excited state at a pump rate $p = 0.8$. The size and color of the circle represent the amplitude and phase of each DOPO[14]

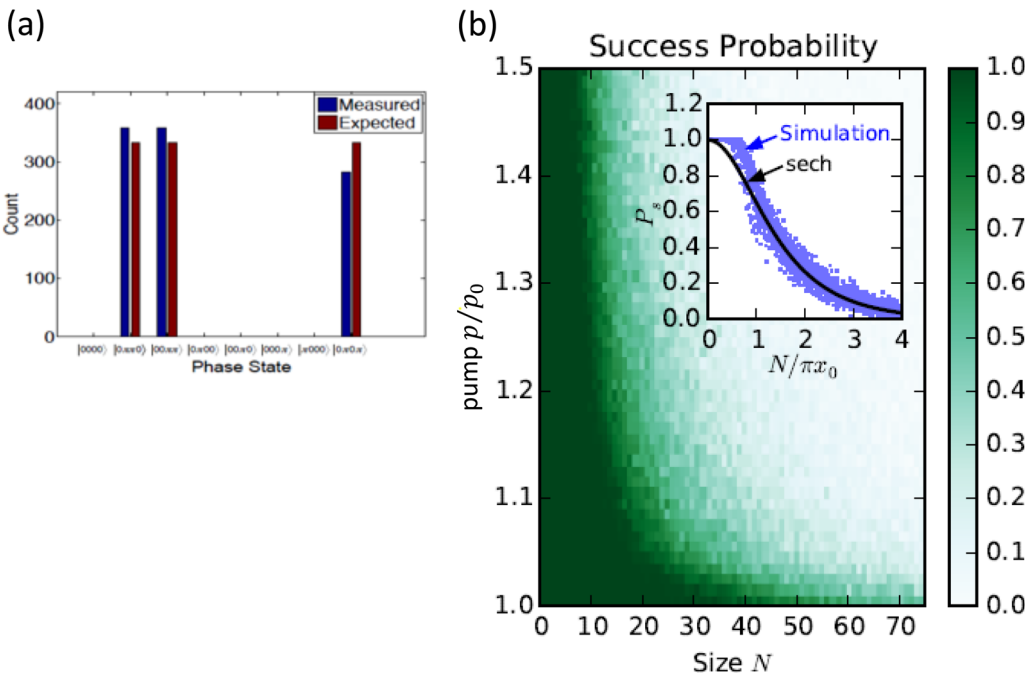

Fig. 8 Performance of optical delay line coupling CIM. **a** Final state histogram for 1000 runs in $N = 4$ CIM for the MAX-CUT-3problems.[15] **b** Success probability as a function of problem size N and pump rate p for the one-dimensional ring model[23]

three degenerate ground states of this complete graph with equal probabilities as shown in Fig. 8a. On the other hand, if the Ising coupling is switched-off, the machine randomly picks 1 of the 8 states by spontaneous symmetry breaking.

The second experimental CIM implements $N = 16$ Ising spins with a free-space multiple-pulse DOPO and connect them sparsely to realize one-dimensional ring or Möbius-Ladder graph configuration using two or three optical delay lines, respectively.[22] The Ising coupling constants are uniform. Again, the machine finds one of the degenerate ground states with almost equal probabilities. These two CIMs have a 100% success probability to find ground states and achieve the random sampling capability for the degenerate ground states.

The third experimental CIM implements $N = 10{,}000$ Ising spins with a fiber DOPO and realizes one-dimensional ring.[23] The success probability as a function of system size N normalized by the correlation length x_0 and pump rate p normalized by the threshold pump rate p_0 agrees fairly well with the theoretical model[48] as shown in Fig. 8b.

Measurement-feedback CIM as exact solvers

A setup of the measurement-feedback based CIM, where $N = 100$ DOPO pulses are fully connected by a single measurement-feedback circuit, is installed at Stanford University.[24] A fraction (~10%) of each pulse intensity is picked off and its in-phase amplitude $\hat{X}_j$ is measured by balanced optical homodyne detectors. Then, a feedback signal for the ith DOPO pulse, $\sum J_{ij}\hat{X}_j$, is computed by an FPGA. This signal is imposed on the feedback optical pulse for the ith DOPO pulse by intensity and phase modulators. Finally, the injection coupler combines the feedback optical pulse and the ith DOPO pulse. Similarly, a feedback optical pulse for the jth DOPO pulse, whose amplitude is proportional to $\sum_i J_{ji}\hat{X}_i$ is prepared and injected to the jth DOPO

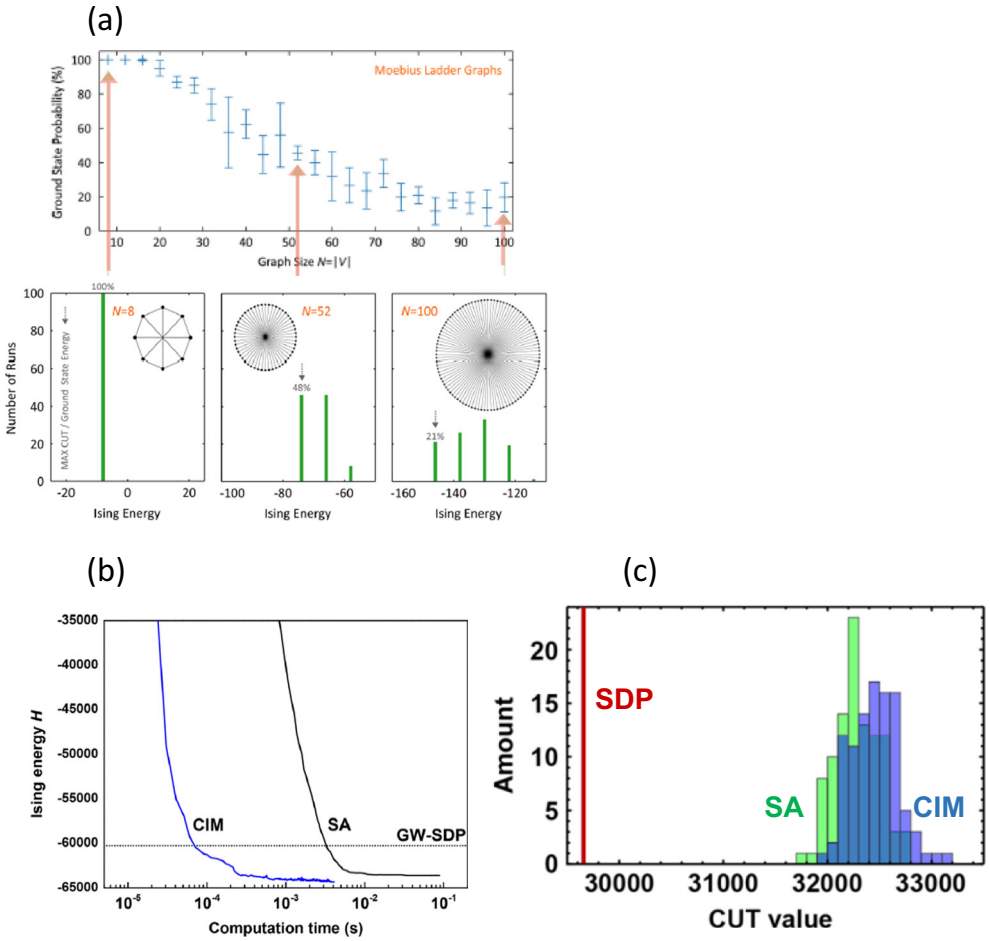

Fig. 9 Performance of measurement-feedback coupling CIM. **a** Results with the measurement-feedback CIM for various-size Möbius-Ladder graphs. (upper) Observed probability of obtaining a ground state of the Möbius-Ladder graph in a single run, as a function of the size N of the graph. Multiple 100-run batches were performed for each graph size to obtain the standard deviations, which are shown as error bars. (Lower) Histograms of obtained solutions in 100 runs for the graphs shown in the insets.[24] **b** Time evolutions of Ising energy obtained with CIM (blue curve) and SA (black curve) when solving the complete graph K_{2000}. The dotted line corresponds to the target Ising energy of −60,278 obtained with GW-SDP.[25] (**c**) Cut values obtained for the MAX-CUT problem on 2000-node graphs with the CIM, SA, and GW-SDP

pulse. After one round trip, the full Ising Hamiltonian is implemented via N injected feedback pulses. Note that asymmetric Ising coupling $J_{ij} \neq J_{ji}$ can be readily implemented in this scheme. On the other hand, a Zeeman term h_i can be implemented by injecting a constant optical pulse which is produced by appropriate amplitude and phase modulation of the master oscillator output.

Figure 9a shows the observed success probability of finding a ground state of the Möbius-Ladder graph as a function of the graph size.[24] Multiple 100-run batches were performed for each graph size and the standard deviations are plotted as error bars. Figure 9a also show the histograms of obtaining the low-energy excited states as well as the ground states. In the case of $N = 100$, the CIM can find the ground state out of $\sim 10^{30}$ solution candidates with a probability of $21 \pm 9\%$. Note that the amplitude heterogeneity is not actively suppressed in this experiment. It is theoretically predicted that the success probability is ~100% for the problem size up to $N = 100$ when the DOPO amplitudes are kept uniform.

Measurement-feedback CIM as approximate solvers

A similar setup of the larger-scale CIM, where $N = 2000$ DOPO pulses are fully connected by a single measurement-feedback circuit, is implemented at NTT.[25] Major changes from the Stanford CIM[24] are the fiber length increased from 300 m to 1 km and the pulse repetition frequency increased from 100 MHz to 1 GHz.

Figure 9b shows the observed Ising energy vs. computation time for the MAX-CUT problems in $N = 2000$ complete graphs with all-to-all connections.[25] It is impossible to find the ground state for this size of the complete graph with a current digital computer. We used the 87.8%, even in worst case, approximate solution obtained by the semi-definite programing (SDP)[20] as a benchmark target. The experimental CIM reaches this target with 70 μs,[25] while the simulated annealing (SA) implemented in the state of art Central Processing Unit (CPU) reaches the same target with 2.1 ms. Table 1 compares the computational time of experimental CIM with those of four different types of classical algorithms implemented on either CPU or supercomputer (PEZY-Shobu) at Riken.[44] The results summarized in Fig. 9b and Table 1 well constitute the evidence for a novel speedup of CIM. Figure 9c shows the histograms of the final energies (or cut values) for the SDP which achieves the best theoretical accuracy at the present time, SA which is the most popular heuristic at the present time, and CIM. The CIM also shows the better performance than SA and SDP in accuracy, too. The computation times for SDP, SA, and CIM are 100 s, 50 ms, and 5 ms, respectively, in this accuracy measurement.

Table 1. Four types of classical neural network models implemented as algorithms in CPU or supercomputer (PEZY-Shobu) and the computational time to reach the 87.8% accuracy level provided by Goemans–Williamson semi-definite programing[44]

Coherent Ising machine	Experimental CIM 0.07 ms	Virtual CIM 100 s
Classical neural network	Deterministic	Stochastic
Discrete variable	Hopfield Network (HN) 0.924 (ms) by CPU	Simulated Annealing (SA) 2.10 (ms) by CPU
Continuous variable	Hopfield-Tank Neural Network (HTNN) 7.04 (ms) by SC	Langevin Dynamics (LD) 100 (ms) by SC

CPU: Intel Xeon, E3–1225 v3, 3.2 GHz. Supercomputer: PEZY-SC, 733 MHz, 1024 cores. The computational time by experimental CIM[23] and virtual CIM using the CSDE and replicator dynamics are also shown, where 16 PEZY-SC are used in parallel to implement the virtual CIM

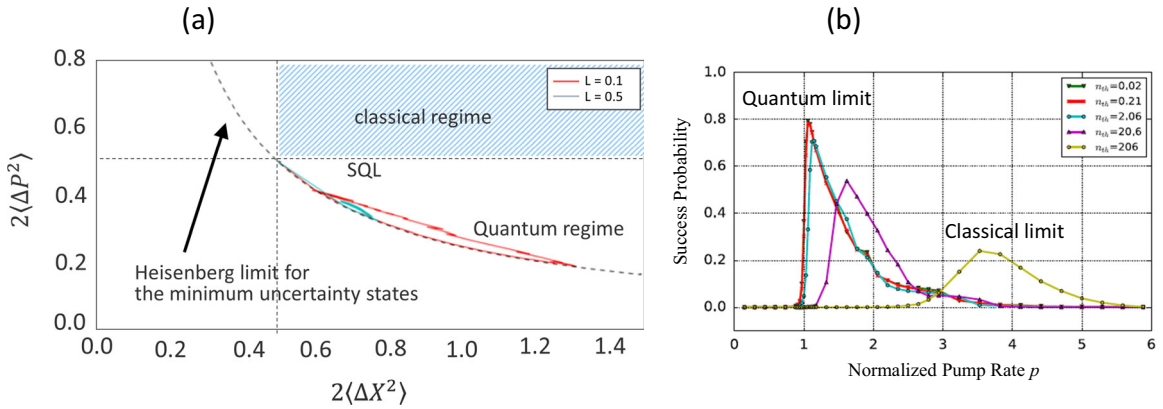

Fig. 10 CIM operating at the quantum and classical limits. **a** The trajectory of $2\langle\Delta\hat{X}^2\rangle$ and $2\langle\Delta\hat{P}^2\rangle$ of the high-Q and low-Q cavity CIM consisting of two anti-ferromagnetically coupled spins.[40] **b** The success probability of finding a solution for an $N = 16$ one-dimensional Ising spin model for various values of the temperature parameter $n_{th} = k_B T/\hbar\omega$[17]

QUANTUM FEATURES OF CIM

Optical neural network at the quantum limit

One of the unique features of optical neural networks is the continuous crossover of their operational modes from quantum limit to classical limit. In this section, we will present various numerical results showing the difference between the two regimes and shedding a light on the quantum-to-classical crossover.

Figure 10a shows the trajectory of the variances $2\langle\Delta\hat{X}^2\rangle$ and $2\langle\Delta\hat{P}^2\rangle$ for the DOPO field in the measurement-feedback CIM consisting of two anti-ferromagnetically coupled spins.[40] A minimum uncertainty state at the Heisenberg limit satisfies $\langle\Delta\hat{X}^2\rangle\langle\Delta\hat{P}^2\rangle = 1/16$, which is shown by the dashed line in Fig. 10a. The CIM with a high-Q or low-Q cavity, in which a round trip loss is either 10% (−0.5 dB) or 50% (−3dB), continuously excites the quantum states which are close to the Heisenberg limit, as shown by red or blue line in Fig. 10a. On the other hand, optical neural networks at thermal noise limit ($k_B T \gg \hbar\omega$) should operate in the classical regime defined by $2\langle\Delta\hat{X}^2\rangle \geq 1/2$ and $2\langle\Delta\hat{P}^2\rangle \geq 1/2$, which is shown by the shaded area in Fig. 10a.

The success probability P_s of finding the ground state of an $N = 16$ one-dimensional Ising spin model, in which only nearest neighbor anti-ferromagnetic interaction exists, is numerically evaluated for various values of the temperature parameter $n_{th} = k_B T/\hbar\omega$ and the result is shown in Fig. 10b.[17] Here the optical delay line coupling CIM is assumed. In this numerical simulation, the pump rate is abruptly increased from $p = 0$ to $p = p_0$ at $t = 0$. In the case of $\hbar\omega \gg k_B T$ (quantum noise limit), the squeezed vacuum state allows the quantum parallel search during a transient time before the steady-state amplitude is formed. The transient time decreases with the final pump rate p_0, so that the maximum success probability is achieved at a final pump rate just above the threshold value, $p_0 \simeq p_{th} = 1$, which allows the DOPO network to have a sufficient time to search for the solution by creating the quantum noise correlation. If the final pump rate p_0 is far above the threshold, the coherent field with random 0-phase or π-phase is formed quickly in each DOPO through spontaneous symmetry breaking, before the quantum search establishes the sufficient quantum noise correlation and identifies a correct solution. The quantum tunneling is not strong enough to overcome the potential barrier separating 0-phase and π-phase when the oscillation field is strong. In this way, the DOPO network is trapped in one of the excited states (local minima). This reasoning explains the monotonic decrease in P_s for $n_{th} \ll 1$ as a function of $p(\geq 1)$ in Fig. 10b.

In the case of $\hbar\omega \ll k_B T$ (thermal noise limit), the squeezed thermal state, which is formed before reaching the threshold, does not allow the quantum parallel search during a transient time before the steady-state amplitude is formed. This is because the quantum coherence between the states $|X\rangle$ and $|-X\rangle$ is ruined by the destructive interference among different photon number eigenstates in a thermal state. The maximum success probability is achieved at a final pump rate $p_0 \gg p_{th} = 1$, where the coherent mean-field is produced and searches for a solution. This mode of operation requires that the mean-field amplitude is larger than the thermal noise amplitude. This is an operation regime of classical neural network (CNN).

Three-step quantum computation at criticality

The three-step quantum computation of the optical delay line coupling CIM is illustrated in Fig. 11a. The success rates to find either one of the two degenerate ground states in $N = 16$ one-

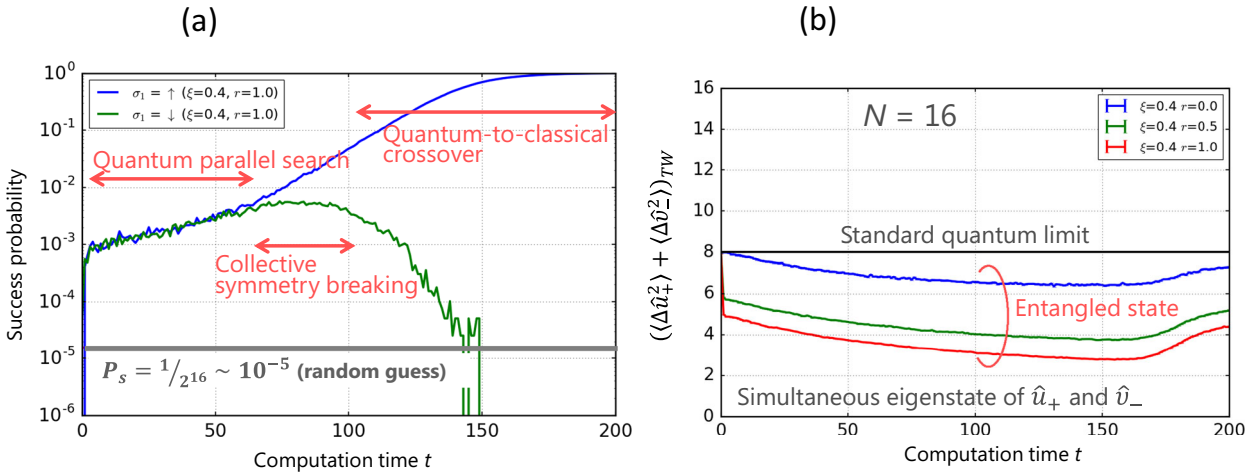

Fig. 11 Three-step quantum computation and entanglement in CIM. **a** The success rates to find the two degenerate ground states of the $N = 16$ 1D Ising spins with anti-ferromagnetic coupling vs. normalized time t/t_c. **b** The total variance of the EPR-like operator vs. normalized time t/t_c, where $\xi = \omega_{ij}$ is the mutual coupling constant and r is the squeezing parameter[17]

dimensioned Ising spins with anti-ferromagnetic coupling are plotted as a function of normalized computation time t/t_c, where t_c is a round trip time.[17] The pump rate is linearly increased from below threshold to above threshold. The success rate for a random guess is $P_s = (1/2)^{16} \simeq 10^{-5}$. After a few round trips, the success rates are increased by two orders of magnitude due to the formation of quantum noise correlation and this trend continues to $t/t_c = 60$, where the average photon number per DOPO pulse reaches $n = 1$ and the collective symmetry breaking is kicked-in. One ground state is selected, while the other is not. The probability to find the selected ground state increases exponentially while that to find the unselected ground state decreases exponentially. This exponential increase in the success rate is made possible by the bosonic final state stimulation and the associated cross-gain saturation. The quantum parallel search at below threshold, the collective symmetry breaking at threshold, and the quantum-to-classical crossover at above threshold are the three major steps in CIM.

Quantum entanglement

In the above example of an $N = 16$ one-dimensional ring consisting of anti-ferromagnetically coupled Ising spins, the ground state should have a negative correlation between the in-phase amplitudes X_i and $X_{i\pm1}$ of neighboring DOPOs. This observation motivates us to define an EPR-like operator,

$$\hat{u}_+ = \hat{X}_1 + \hat{X}_2 + \cdots + \hat{X}_{16}, \tag{17}$$

$$\hat{v}_- = \hat{P}_1 - \hat{P}_2 - \cdots - \hat{P}_{16}, \tag{18}$$

to demonstrate the entanglement that exists in the DOPO network. Since $\hat{u}_+$ and $\hat{v}_-$ commute, the simultaneous eigenstate for $\hat{u}_+$ and $\hat{v}_-$ should exist and such a state satisfies $\langle \Delta\hat{u}_+^2 \rangle = \langle \Delta\hat{v}_-^2 \rangle = 0$. On the other hand, if all DOPOs are independent (separable), it is shown that $\langle \Delta\hat{u}_+^2 \rangle + \langle \Delta\hat{v}_-^2 \rangle \geq 8$.[17] This means that if $\langle \Delta\hat{u}_+^2 \rangle + \langle \Delta\hat{v}_-^2 \rangle < 8$, such a system has the entanglement due to mutual coupling. Figure 11b demonstrates that the optical delay line coupling CIM indeed establishes the quantum entanglement in the system over a wide range of the computation time (or pump rate).[17]

Quantum coherence

Figure 12a–d show the contour maps of the density matrix elements $\langle X|\hat{\rho}|X'\rangle$ of the measurement-feedback CIM consisting of two DOPOs with anti-ferromagnetic coupling as a function of

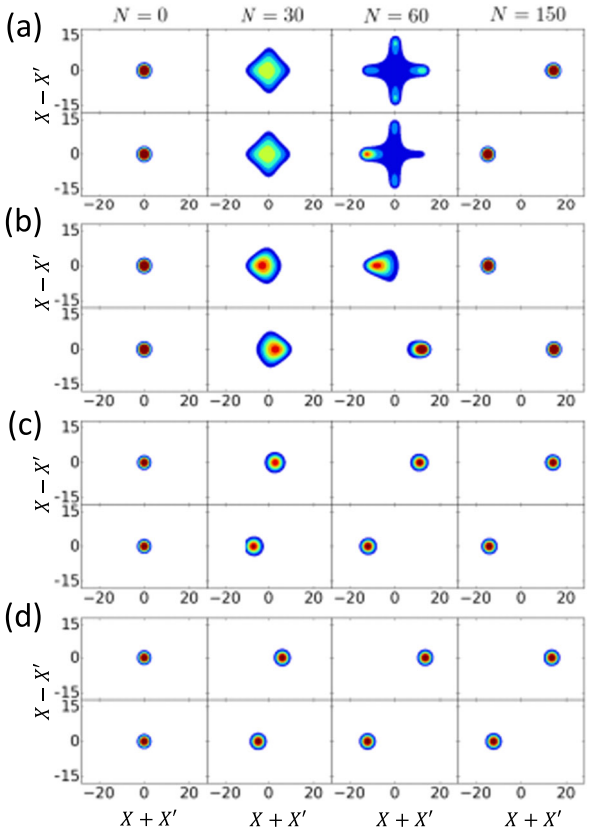

Fig. 12 Evolution of the density operator in CIM. The contour maps of typical conditional density matrices $\langle X|\hat{\rho}|X'\rangle$ of two oscillators governed by the measurement results x_m with (**a**) $T = 0.999$, (**b**) $T = 0.99$, (**c**) $T = 0.9$, and (**d**) $T = 0.5$ at round trips of $N = 0$, 30, 60, 150 in front of the PSA, where T is the transmission coefficient per round trip

normalized computation time $N = t/t_c$.[40] In a high-Q cavity with a round trip loss of 0.1% (Fig. 12a), the DOPO state near threshold ($t/t_c = 60$) is indeed in a Schrödinger cat-like state. However, in a low-Q cavity with a round trip loss of 50% (Fig. 12d), the DOPO state evolves from a vacuum state at $t/t_c = 0$, a squeezed vacuum state at threshold to a coherent state at above threshold. The quantum

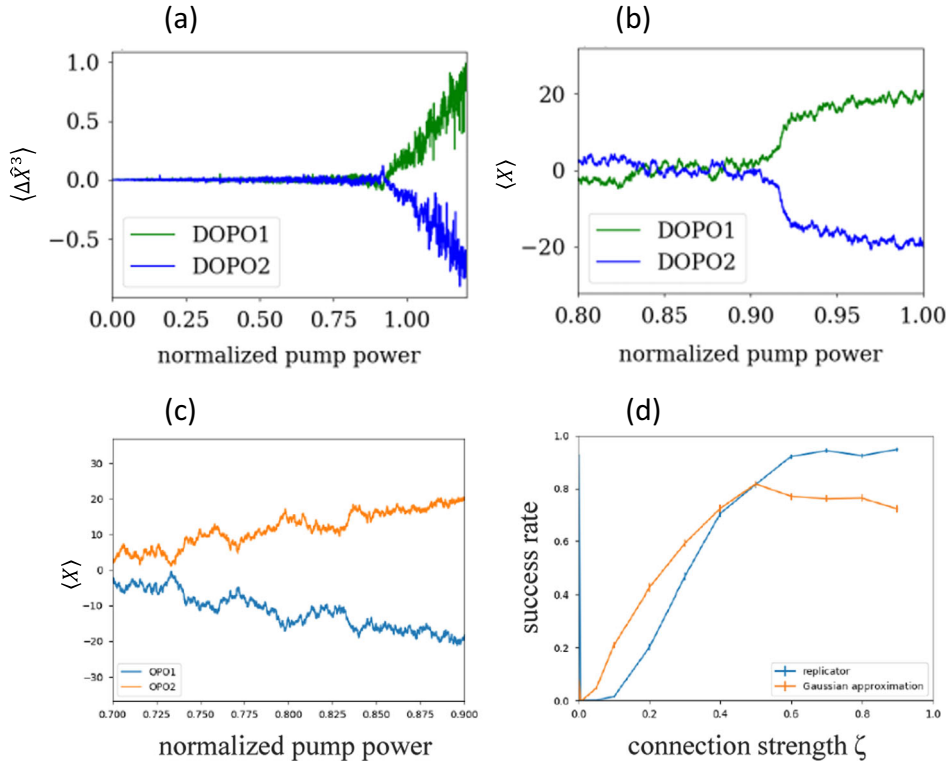

Fig. 13 Non-Gaussian states in CIM. **a** Time evolution of $\langle \Delta\hat{X}^3 \rangle$ of the in-phase amplitudes of the two DOPO fields. **b** Time evolution of $\langle \hat{X} \rangle$ under the replicator dynamics generated by the ensemble over 10,000 runs. **c** Time evolution of $\langle \hat{X} \rangle$ under the Gaussian approximation. **d** The success rates for fining a ground state of $N = 16$ one-dimensional Ising spin problem by the exact and Gaussian approximation[45]

states of the DOPO seem to stay always in Gaussian states in Fig. 12d, which is actually not the case as we will discuss in the next section. In any case, the quantum coherence $\langle X|\hat{\rho}|-X\rangle$, represented by the vertical slice, always exists in the CIM, no matter how large the cavity loss is.

Non-Gaussian wavepackets

The archetype bistable potential $V_b(x_j) = -\frac{1}{2}\alpha x_i^2 + \frac{1}{4}x_i^4$ at above threshold has a steep potential for large $|x_j|$ values but a shallow potential barrier for small $|x_j|$ values when the DOPO is pumped not far above the threshold, i.e. $0 < \alpha \lesssim 1$. As a result of this asymmetric potential profile, the DOPO wave packet has a rapidly decaying tail for large $|x_j|$ values and slowly decaying tail for small $|x_j|$ values.

Figure 13a shows the time evolution of skewness $\langle \Delta\hat{X}^3 \rangle$ of the two DOPO fields.[45] At below threshold, $\langle \Delta\hat{X}^3 \rangle$ is close to zero, which is expected for a Gaussian state. At threshold, however, $\langle \Delta\hat{X}^3 \rangle$ of the two DOPO fields depart to opposite directions, which indicates that the internal DOPO state has a slowly decaying tail toward a central potential and a rapidly decaying tail toward outer potential. Such a non-Gaussian state allows more frequent quantum tunneling before the final decision is made as demonstrated in the simulation result shown in Fig. 13b.[45] This switching behavior is a characteristic quantum parallel search of CIM, which is in sharp contrast to that of a hypothetical machine supporting Gaussian states (Fig. 13c).[45] In this case, the two DOPOs never switch their polarities due to the suppressed quantum tunneling.

Gottesman–Knill theorem

Not all quantum dynamics are difficult to simulate by classical digital computers. Some of representative and important quantum processes, including entangled state generation and purification, can be efficiently simulated by classical methods, so that such a quantum system alone is unlikely to achieve a computation power exceeding the current states of art in digital computing technology.

Gottesman and Knill were the first to point out this subtle distinction between classical and quantum information processing.[46] The statement of the Gottesman–Knill theorem can be summarized as follows:

If a quantum process starts with

1. computational basis states, such as ground states $|0\rangle_1|0\rangle_2\cdots|0\rangle_N$,
2. employs a limited set (Clifford group) of unitary gates such as Hadamard gates, phase gates and controlled-NOT gates,and ends with
3. projective measurements along the computational basis states $\{|0\rangle|1\rangle\}$,

such a quantum process can be efficiently simulated by classical digital computers. A reader who is familiar with the famous Shor's factoring algorithm knows that it requires implementation of a fractional phase which is not included in the above Clifford group constraint, so that the Shor algorithm is outside of the above limitation.

A continuous variable (harmonic oscillator) version of the above theorem was developed by Bartlett et al.[47] The statement of this theorem runs as follows:

If a quantum process with harmonic oscillators starts with

1. Gaussian states, such as coherent states $|\alpha\rangle_1|\alpha\rangle_2\cdots|\alpha\rangle_N$,
2. employs a limited set of unitary gates such as squeezing gates and displacement gates,and ends with
3. projective measurements of one quadrature amplitude (by homodyne detection) or two quadrature amplitudes (by heterodyne detection),

such a quantum process can be efficiently simulated by classical digital computers.

If we consider the above theorem against the CIM, we can identify the gain saturation (or two photon absorption) and single photon loss as the two essential quantum dynamics which make the CIM difficult to simulate efficiently by classical methods. The gain saturation term, $-x_j^3$, and the dissipation term, αx_j, in Eq. (12) produce a non-Gaussian state and provide the non-classical nature to CIM. Indeed, as shown in Fig. 13d, the success rate to find a ground state of $N = 16$ one-dimensional Ising spin problem based on the exact theory is higher than that based on the Gaussian approximation.[45]

SUMMARY

Optical neural networks can operate at the quantum limit ($k_BT/\hbar\omega \ll 1$) at room temperatures and realize a quantum parallel search based on squeezed vacuum states. Such CIM are composed of the two constituent devices: quantum neurons and quantum synapses.

Quantum neurons are provided by degenerate optical parametric amplifiers/oscillators, which realize the quantum parallel search at below the oscillation threshold, the decision making at the threshold based on collective symmetry breaking, and the quantum-to-classical amplification of computational results at above the threshold using bosonic final state stimulation. Quantum synapses are provided by either direct coupling with optical delay lines or indirect coupling with a measurement-feedback circuit. They utilize distinct computational resources: quantum noise correlation (entanglement) in the optical delay line coupling CIM and measurement-induced wave packet reduction to non-Gaussian states in the measurement-feedback CIM.

CIM can solve various combinatorial optimization problems by mapping them on NP-hard Ising problems. Optical neural networks at the quantum limit (CIM) can outperform those at classical regime by exploiting superposition states for quantum parallel search of solutions. The gain saturation and single photon loss are the two indispensable elements to make CIM hard to simulate by classical methods. These two dissipation processes and associated fluctuations from external reservoirs are crucial resources to accelerate the search process.

CIM is based on single-mode squeezing and oscillation at degenerate frequency. This preferred mode of operation is made possible by employing a pump rate not far above the oscillation threshold and a narrow bandpass filter in a fiber ring cavity. Nonlinear dynamics of DOPO in various parameter ranges, including crossover from degenerate to non-degenerate modes, are presented in refs. 31, 48, 49.

Another type of optical neural networks, coherent XY machines, have been recently demonstrated using coupled lasers,[50,51] non-degenerate OPOs,[52] and polaritons.[53] The Kuramoto model and the continuous optimization problems, in general, can be potentially implemented on these machines.

Data and code availability

All relevant data and codes are available from the authors upon request.

ACKNOWLEDGEMENTS

The authors wish to thank R. L. Byer, H. Mabuchi, R. D. Drummond, P. McMahon, A. Marandi, C. Langrock, T. Inagaki, T. Honjo, K. Inaba, R. Hamerly, Y. Haribara, D. Maruo, K. Takata, Z. Wang, A. Yamamura, and T. Shoji for their useful discussions. This research was supported by ImPACT Program of Council for Science, Technology and Innovation (Cabinet Office, Government of Japan).

AUTHOR CONTRIBUTIONS

Y.Y. wrote the manuscript, K.A., K.K., T.L., S.K., M.F., K.I., and H.T. critically read and revised the manuscript.

ADDITIONAL INFORMATION

Competing interests: The authors declare that they have no competing financial interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Change history: A correction to this article has been published and is linked from the HTML version of this article.

REFERENCES

1. Hopfield, J. J. & Tank, D. W. Computing with neural circuits: a model. *Science* **233**, 625–633 (1986).
2. Tank, D. & Hopfield, J. Simple 'neural' optimization networks: an A/D converter, signal decision circuit, and a linear programming circuit. *IEEE Trans. Circ. Syst.* **33**, 533–541 (1986).
3. Ercsey-Ravasz, M. & Toroczkai, Z. Optimization hardness as transient chaos in an analog approach to constraint satisfaction. *Nat. Phys.* **7**, 966–970 (2011).
4. Qian, L., Winfree, E. & Bruck, J. Neural network computation with DNA strand displacement cascades. *Nature* **475**, 368–372 (2011).
5. Farhi, E. et al. A quantum adiabatic evolution algorithm applied to random instances of an NP-complete problem. *Science* **292**, 472–475 (2001).
6. Kadowaki, T. & Nishimori, H. Quantum annealing in the transverse Ising model. *Phys. Rev. E* **58**, 5355–5363 (1998).
7. Brooke, J., Bitko, D., Rosenbaum, T. F. & Aeppli, G. Quantum annealing of a disordered magnet. *Science* **284**, 779–781 (1999).
8. Johnson, M. W. et al. Quantum annealing with manufactured spins. *Nature* **473**, 194–198 (2011).
9. Rønnow, T. F. et al. Defining and detecting quantum speedup. *Science* **345**, 420–424 (2014).
10. Boixo, S. et al. Evidence for quantum annealing with more than one hundred qubits. *Nat. Phys.* **10**, 218–224 (2014).
11. Barends, R. et al. Digitized adiabatic quantum computing with a superconducting circuit. *Nature* **534**, 222–226 (2016).
12. Bunyk, P. I. et al. Architectural considerations in the design of a superconducting quantum annealing processor. *IEEE Trans. Appl. Supercond.* **24**, 1700110 (2014).
13. Rieffel, E. G. et al. A case study in programming a quantum annealer for hard operational planning problems. *Quant. Inform. Process.* **14**, 1–36 (2015).
14. Wang, Z., Marandi, A., Wen, K., Byer, R. L. & Yamamoto, Y. Coherent Ising machine based on degenerate optical parametric oscillators. *Phys. Rev. A* **88**, 063853 (2013).
15. Marandi, A., Wang, Z., Takata, K., Byer, R. L. & Yamamoto, Y. Network of time-multiplexed optical parametric oscillators as a coherent Ising machine. *Nat. Photonics* **8**, 937–942 (2014).
16. Takata, K., Marandi, A. & Yamamoto, Y. Quantum correlation in degenerate optical parametric oscillators with mutual injections. *Phys. Rev. A* **92**, 043821 (2015).
17. Maruo, D., Utsunomiya, S. & Yamamoto, Y. Truncated Wigner theory of coherent Ising machines based on degenerate optical parametric oscillator network. *Phys. Scr.* **91**, 083010 (2016).
18. Barahona, F. On the computational complexity of Ising spin glass models. *J. Phys. Math. Gen.* **15**, 3241–3253 (1982).
19. Istrail, S. *Proc. Thirty-second Annual ACM Symposium on Theory of Computing, STOC*, 87–96 (2000).
20. Haribara, Y., Utsunomiya, S. & Yamamoto, Y. Computational principle and performance evaluation of coherent Ising machine based on degenerate optical parametric oscillator network. *Entropy* **18**, 151 (2016).
21. Lucas, A. Ising formulations of many NP problems. *Front. Phys.* **2**, 5 (2014).
22. Takata, K. et al. A 16-bit coherent Ising machine for one-dimensional ring and cubic graph problems. *Sci. Rep.* **6**, 34089 (2016).
23. Inagaki, T. et al. Large-scale Ising spin network based on degenerate optical parametric oscillators. *Nat. Photonics* **10**, 415–419 (2016).
24. McMahon, P. L. et al. A fully programmable 100-spin coherent Ising machine with all-to-all connections. *Science* **354**, 614 (2016).
25. Inagaki, T. et al. A coherent Ising machine for 2000-node optimization problems. *Science* **354**, 603 (2016).
26. Nigg, S. E., Lörch, N. & Tiwari, R. P. Robust quantum optimizer with full connectivity. *Sci. Adv.* **3**, e1602273 (2017).

27. Puri, S., Andersen, C. K., Grimsmo, A. L. & Blais, A. Quantum annealing with all-to-all connected nonlinear oscillators. *Nat. Commun.* **8**, 15785 (2017).
28. Walls, D. F. & Milburn, G. J. *Quantum Optics* (Springer, Berlin-Heidelberg, 2008).
29. Byer, R. L., Oshman, M. K., Young, J. F. & Harris, S. E. Visible CW parametric oscillator. *Appl. Phys. Lett.* **13**, 109 (1968).
30. Serkland, D. K., Fejer, M. M., Byer, R. L. & Yamamoto, Y. Squeezing in a quasi-phase-matched LiNbO_3 waveguide. *Opt. Lett.* **20**, 1649 (1995).
31. Hamerly, R. et al. Reduced models and design principles for half-harmonic generation in synchronously pumped optical parametric oscillators. *Phys. Rev. A* **94**, 063809 (2016).
32. Nambu, Y. Nobel Lecture: Spontaneous symmetry breaking in particle physics: a case of cross fertilization. *Rev. Mod. Phys.* **81**, 1015 (2009).
33. Goto, H. Bifurcation-based adiabatic quantum computation with a nonlinear oscillator network. *Sci. Rep.* **6**, 21686 (2016).
34. Goto, H. Universal quantum computation with a nonlinear oscillator network. *Phys. Rev. A* **93**, 050301 (R) (2016).
35. Goto, H., Lin, Z. & Nakamura, Y. Dissipative quantum bifurcation machine: quantum heating of coupled nonlinear oscillators. arXiv:1707.00986 [quant-ph] (2017).
36. Mandel, L. & Wolf, E. *Optical Coherence and Quantum Optics* (Cambridge University Press, 1995).
37. Wolinsky, M. & Carmichael, H. J. Quantum noise in the parametric oscillator: from squeezed states to coherent-state superpositions. *Phys. Rev. Lett.* **60**, 1836 (1988).
38. Krippner, L., Munro, W. J. & Reid, M. D. Transient macroscopic quantum superposition states in degenerate parametric oscillation: calculations in the large-quantum-noise limit using the positive P representation. *Phys. Rev. A* **50**, 4330 (1994).
39. Drummond, P. D., Dechoum, K. & Chaturvedi, S. Critical quantum fluctuations in the degenerate parametric oscillator. *Phys. Rev. A* **65**, 033806 (2002).
40. Yamamura, A., Aihara, K. & Yamamoto, Y. Quantum model for coherent Ising machines: Discrete-time measurement feedback formulation. *Phys. Rev. A* **96**, 053834 (2017).
41. Ising, E. Beitrag zur theorie des ferromagnetismus. *Z. Phys.* **31**, 253–258 (1925).
42. Leleu, T., Yamamoto, Y., Utsunomiya, S. & Aihara, K. Combinatorial optimization using dynamical phase transitions in driven-dissipative systems. *Phys. Rev. E* **95**, 022118 (2017).
43. Kinsler, P. & Drummond, P. D. Quantum dynamics of the parametric oscillator. *Phys. Rev. A* **43**, 6194 (1991).
44. Haribara, Y., Ishikawa, H., Utsunomiya, S., Aihara, K. & Yamamoto, Y. Performance evaluation of coherent Ising machines against classical neural networks. *Quant. Sci. Tech.* **2**, 044002 (2017).
45. Shoji, T., Aihara, K. & Yamamoto, Y. Quantum model for coherent Ising machine: Stochastic differential equations with replicator dynamics. *Phys. Rev. A* **96**, 053833 (2017).
46. Gottesman, D. The Heisenberg representation of quantum computers. arXiv: quant-ph/9807006 (1998).
47. Bartlett, S. D., Sanders, B. C., Braunstein, S. L. & Nemoto, K. Efficient classical simulation of continuous variable quantum information processes. *Phys. Rev. Lett.* **88**, 097904 (2002).
48. Hamerly, R. et al. Topological defect formation in 1D and 2D spin chains realized by network of optical parametric oscillators. *Int. J. Mod. Phys. B* **30**, 1630014 (2016).
49. Roslund, J., Medeiros de Araújo, R., Jiang, S., Fabre, C. & Treps, N. Wavelength-multiplexed quantum networks with ultrafast frequency combs. *Nature Photonics* **8**, 109 (2014).
50. Nixon, M., Ronen, E., Friesem, A. A. & Davidson, N. Observing geometric frustration with thousands of coupled lasers. *Phys. Rev. Lett.* **110**, 184102 (2013).
51. Tamate, S., Yamamoto, Y., Marandi, A. McMahon, P. & Utsunomiya, S. Simulating the classical XY model with a laser network. arXiv:1608.00358 [physics.optics] (2016).
52. Takeda, Y. et al. Boltzmann sampling for an XY model using a non-degenerate optical parametric oscillator network. *Quant. Sci. Tech.* **3**, 014004 (2018).
53. Berloff, N. G., Kalinin, K., Silva, M., Langbein, W. & Lagoudakis P. G. Realizing the XY Hamiltonian in polariton simulators. arXiv:1607.06065 [cond-mat.mes-hall] (2016).

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2017