

I. Introduction to Quantum Neural Networks

In this Chapter, we will introduce the basic concepts and operational principles of a novel computing machine, optical neural networks at quantum limit, and describe their unique characteristics. We start with the discussion how to construct such physical devices as the quantum version of classical neurons and synapses.

1.1 Quantum neurons

Nonlinear devices that have a characteristic input-output relation as shown in Fig. 1 represent the typical activity of neurons in classical neural networks [1]. When the 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, the output signal level is clamped at a constant value b_s . This nonlinear input-output relation is essential for classical neural networks to find a stable operating point which manifests a solution for a given mathematical problem [1]. Each neuron state is expressed by a continuous variable x_i and obeys a continuous time evolution governed by

$$\frac{d}{dt}x_i = -x_i + f(x_i) - \frac{\partial V}{\partial x_i}, \quad (1)$$

where the first term of R.H.S. of E.(1) expresses the linear loss and the second term describes the nonlinear gain function shown in Fig. 1, which can be interpreted as a self-feedback excitation to the neuron i . The third term represents the mutual coupling among neurons and the interaction potential V implements a given mathematical problem. We emphasize that, in the case of E.(1), a simultaneous mutual coupling between neurons without introducing undesired instability or periodic oscillation is made possible by the gradient descent character of the third term.

1.1.1 Degenerate optical parametric amplifiers/oscillators

Any optical amplifier has a similar input-output relation as that shown in Fig. 1 due to its inherent gain saturation effect. We focus here on the degenerate optical parametric amplifier

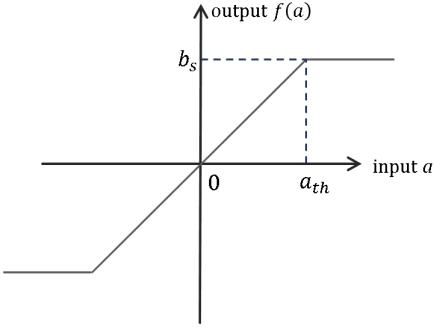


Figure 1: The characteristic input-output relation of a neuron in classical neural networks.

(DOPA) shown in Fig. 2. This particular device is composed of a second-order nonlinear crystal placed between two mirrors. 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 [2] is expressed by

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

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 the 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 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 along the X -axis and P -axis, respectively, in (X, P) phase space. As a result of such quantum interference, if a DOPA is driven by an external vacuum state (zero-point fluctuation) as shown in Fig. 2, the zero-point fluctuation is amplified and deamplified along X -axis and P -axis, respectively, as shown in Fig. 3(a). The resulting state is called a squeezed vacuum state, which is a minimum uncertainty wavepacket and thus 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. This is just an illustrative example of the Liouville theorem of quantum mechanics, which dictates that the uncertainty area of a quantum state is invariant against a unitary time evolution.

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. 4, where a person (corresponding to the pump in DOPA) completes a full cycle (up-down-up)

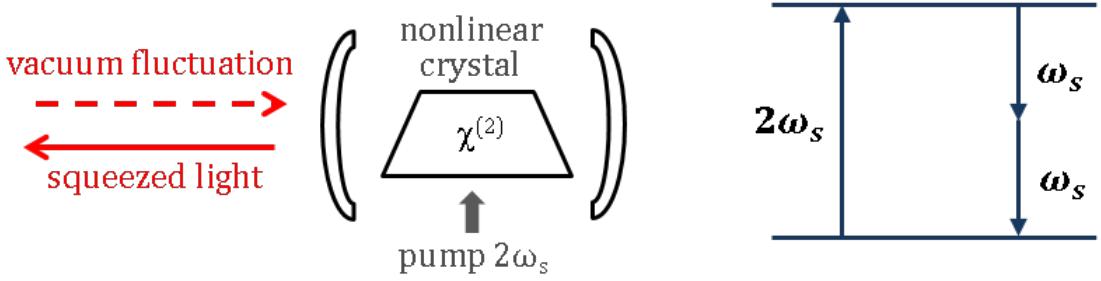


Figure 2: A degenerate optical parametric amplifier/oscillator.

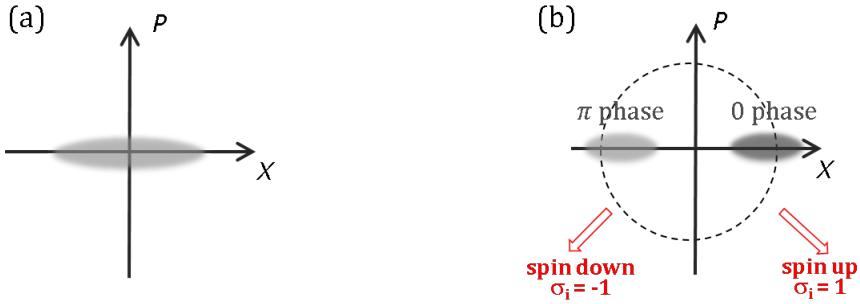


Figure 3: Quantum Noise distributions of (a) DOPA at below the threshold and (b) DOPO at above the threshold.

while a swing (corresponding to the signal) makes only a half cycle (left-to-right). Note that the phase of the person and the phase of the swing are mutually locked in order to realize an amplitude amplification process. This corresponds to the situation in which 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 it was time to go home. You stood up at the center of the swing period and crouched at both ends. Then, the swing (signal) amplitude was attenuated. This corresponds to the situation that the quadrature amplitude \hat{P} is deamplified.

An optical parametric oscillator has a long history of development at Stanford University [3]. In the quantum neural network (QNN), a periodically polled lithium-niobate (PPLN) waveguide device is used as such a phase sensitive amplifier (PSA) for signal pulses. The first experimental demonstration of deamplification (squeezing) and amplification (anti-squeezing) for input optical pulses in vacuum states using this particular device was reported in 1995 [4]. Figure 5 shows the suppressed quadrature-phase amplitude noise and enhanced in-phase amplitude noise by the PPLN waveguide DOPA. It is a general property of the squeezing experiments that the degree of squeezing (vacuum noise deamplification) is lim-

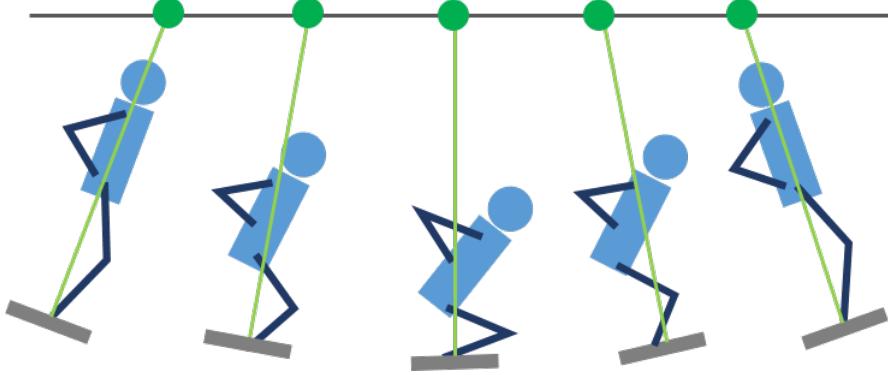


Figure 4: A phase sensitive amplification/deamplification process of a swing, in which a person makes a full cycle (up-down-up) but the swing makes only half cycle (left-to-right).

ited by experimental system imperfections, in particular, by a linear optical loss but that the degree of anti-squeezing (vacuum noise amplification) is not ruined so much by the optical linear loss. Note that the quantum parallel search in the QNN is based on the anti-squeezed quantum noise rather than squeezed quantum noise, so that the operation of QNN is inherently robust against the optical loss.

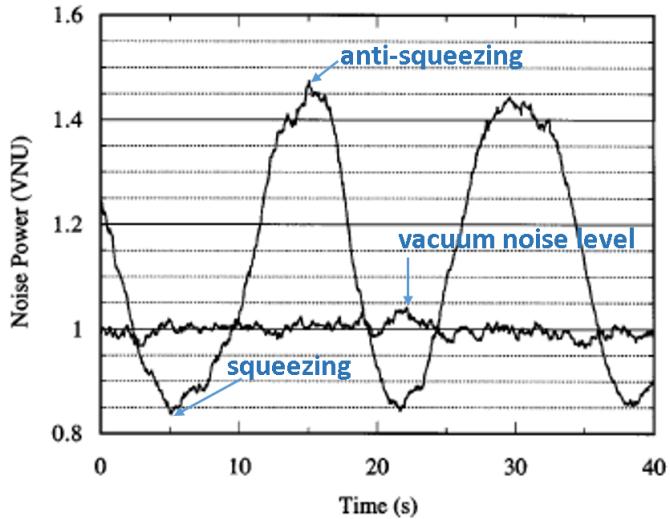


Figure 5: Vacuum noise amplification/deamplification characteristics of a PPLN waveguide PSA for optical signal pulses [4].

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. 1. 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 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. For quantum neurons, a related but slightly different nonlinear effect is employed to amplify a solution state to a macroscopic level, which we will explain next.

The two mirrors in a DOPO shown in Fig. 2 confine an optical field as a resonator. If the parametric amplifier gain exceeds the field decay rate from the resonator, the system can sustain a steady state field by creating a finite average field along the X -direction. The center position of a quantum wavepacket in X - P phase space is no more $\langle \hat{X} \rangle = \langle \hat{P} \rangle = 0$ but there emerges a non-zero $\langle \hat{X} \rangle$ value, as shown in Fig. 3(b). This is the oscillation threshold point of DOPO. The probability for creating a positive $\langle \hat{X} \rangle$ value and negative $\langle \hat{X} \rangle$ 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 [5]. In the language of nonlinear dynamics theory, the DOPO threshold represents a supercritical pitchfork bifurcation. The device pumped at above the oscillation threshold is called a degenerate optical parametric oscillator (DOPO). The two stable states with positive and negative $\langle \hat{X} \rangle$ values correspond to the firing and non-firing states of a classical neural network or up-spin and down-spin states of an Ising problem. When the input signal level to the DOPA is varied while the pump rate is set at below the oscillation threshold, the output signal level is first linearly amplified and then clamped due to the gain saturation. Finally, a DOPA with an increasing input signal level at a fixed pump rate below the threshold eventually exceeds the oscillation threshold. In this way, the nonlinear input/output relation shown in Fig. 1 is realized in the transition from the DOPA to DOPO.

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.

1.1.2 Linear superposition states in DOPA/DOPO

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. 6) can be mathematically constructed as a superposition of photon number eigenstates with even eigenvalues:

$$|\psi_{sv}\rangle = c_0 |0\rangle + c_2 |2\rangle + c_4 |4\rangle + \dots . \quad (3)$$

A simple physical picture behind E.(3) is explained as follows: the photon number of an intense pump field exhibits large quantum noise, 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 $|0\rangle, |2\rangle, |4\rangle, \dots$ photon number eigenstates. Because of the lack of which-path information, the proper expression for the quantum state of the output signal field must be a superposition of those states. Except for an irrelevant phase factor, the probability amplitudes $c_0, c_2, c_4 \dots$ have an identical phase, for instance, positive real numbers [6]. In the (X, P) phase space, the different photon number eigenstates constructively interfere with each other along X -axis, while they destructively interfere with each other along P -axis, as shown in Fig. 6(a).

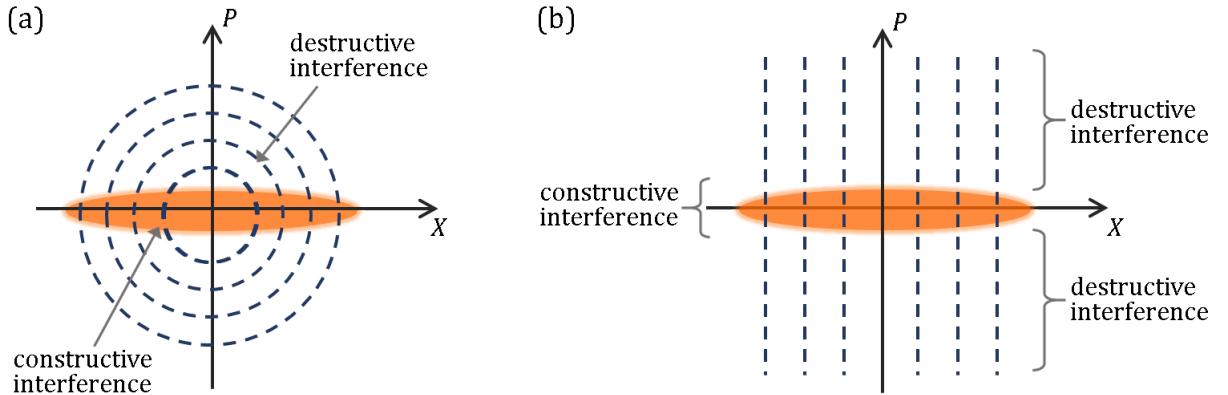


Figure 6: Two representations for a squeezed vacuum state. (a) Superposition of photon number eigenstates with even-number eigenvalues. (b) Superposition of in-phase amplitude eigenstates.

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} :

$$|\psi_{sv}\rangle = \int_{-\infty}^{\infty} C(X) |X\rangle dX. \quad (4)$$

It is not difficult to show that the different in-phase amplitude eigenstates constructively interfere with each other in the small P -regime centered at $\langle \hat{P} \rangle = 0$, while they destructively interfere with each other in the large P -regime, as shown in Fig. 6(b). 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

$$\hat{\rho}_{\text{mix}} = \int_{-\infty}^{\infty} P(X) |X\rangle \langle X| dX, \quad (5)$$

instead of E.(4), covers the same range of the eigenvalues X but require an infinite amount of energy to realize, since the localization of the state along P -axis near $\langle \hat{P} \rangle = 0$ is now impossible due to the lack of phase coherence between different $|X\rangle$ eigenstates. The infinite uncertainty along P -axis means an infinite energy. A squeezed vacuum state, represented by E.(4) and produced by the DOPA, allows a quantum parallel search with a finite energy, while a classical mixture state, represented by E.(5), cannot be created with a finite energy.

A particularly unique feature of the DOPA/DOPO is that the aforementioned superposition survives not only below the threshold but also at the threshold. At well above the oscillation threshold, the DOPO produces either 0-phase or π -phase coherent field as shown in Fig. 3(b). However, a DOPO at just above the oscillation threshold maintains the superposition of 0-phase and π -phase, in spite of finite photon leakage from the resonator [7–9]. This is 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 X -axis. The anti-squeezed in-phase amplitude noise realizes a so-called “quantum erasure” for which-path information.

One of the evidence for the above statement is shown in Fig. 7(a). The density matrix elements $\langle X | \hat{\rho} | X' \rangle$ for the two out-of-phase coupled DOPOs, when the pump rate indexed by N is increased from below to above the threshold, are plotted. In Fig. 7(a) at $N = 60$, the probability distribution $\langle X | \hat{\rho} | X \rangle$, which is given as the slice along x -axis, and the quantum coherence $\langle X | \hat{\rho} | -X \rangle$, which is given as the slice along y -axis, of the density operator feature a macroscopically separated coherent state $|\alpha\rangle$ and $|-\alpha\rangle$ that maintains quantum coherence [10]. Figure 7(b) compares the density matrix elements $\langle X | \hat{\rho} | X' \rangle$ for a superposition state, $|\psi\rangle = \frac{1}{\sqrt{2}} (|\alpha\rangle + |-\alpha\rangle)$, and a mixed state, $\hat{\rho} = \frac{1}{2} (|\alpha\rangle \langle \alpha| + |-\alpha\rangle \langle -\alpha|)$, respectively. We can conclude from Fig. 7(a) that the two DOPO are in Schrödinger’s cat states and yet their centers of gravity are negatively correlated due to out-of-phase coupling. Figure 7(c) shows the Wigner function of the DOPO state at the threshold. The oscillatory behavior and negative values of the Wigner function is another manifestation of the quantum coherence

between $|\alpha\rangle$ and $|-\alpha\rangle$ states [10]. Such a highly non-classical feature is destroyed by an increasing cavity loss, but the superposition of $|X\rangle$ eigenstates, represented by E.(3), survives against a large cavity loss. We will discuss this important point in detail in Chapter V.

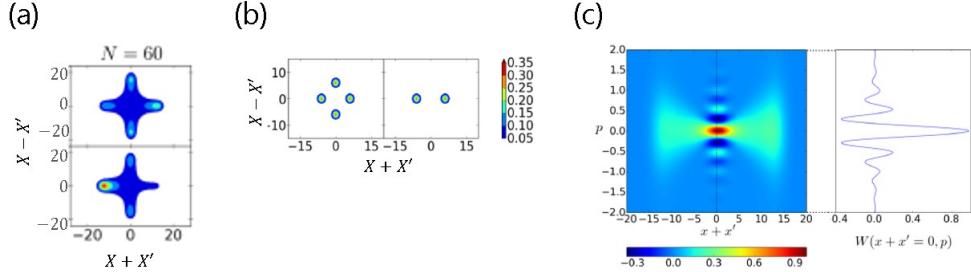


Figure 7: (a) The probability distribution $\langle X | \hat{\rho} | X \rangle$ and the quantum coherence $\langle X | \hat{\rho} | -X \rangle$ of the two DOPO states in a high-Q regime at the threshold pump rate. The round trip loss is 0.1%. (b) The density matrix elements $\langle X | \hat{\rho} | X' \rangle$ for a superposition state, $|\psi\rangle = 1/\sqrt{2}(|\alpha\rangle + |-\alpha\rangle)$ and a mixed state, $\hat{\rho} = 1/2(|\alpha\rangle\langle\alpha| + |\alpha\rangle\langle-\alpha|)$. (c) The corresponding Wigner functions to Fig. 7(a) [10].

1.1.3 Amplitude and phase error correction by phase sensitive amplification

The DOPO fields are fluctuated by the external phase and amplitude noise injection. Both amplitude and phase of the DOPO field are continuous variables so that the error detection and error correction are not straightforward, just like classical analog devices. Fortunately, the phase sensitive amplification mechanism of a DOPO can stabilize the phase to either 0 or π as far as the phase error is small compared to $\pm\pi/2$. Moreover, the amplitude error can be 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. 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. 8.

The DOPO provides a unique opportunity as a stable analog memory. We can stably store the analog information (both amplitude and phase) at quantum limited accuracy with the above mentioned mechanisms.

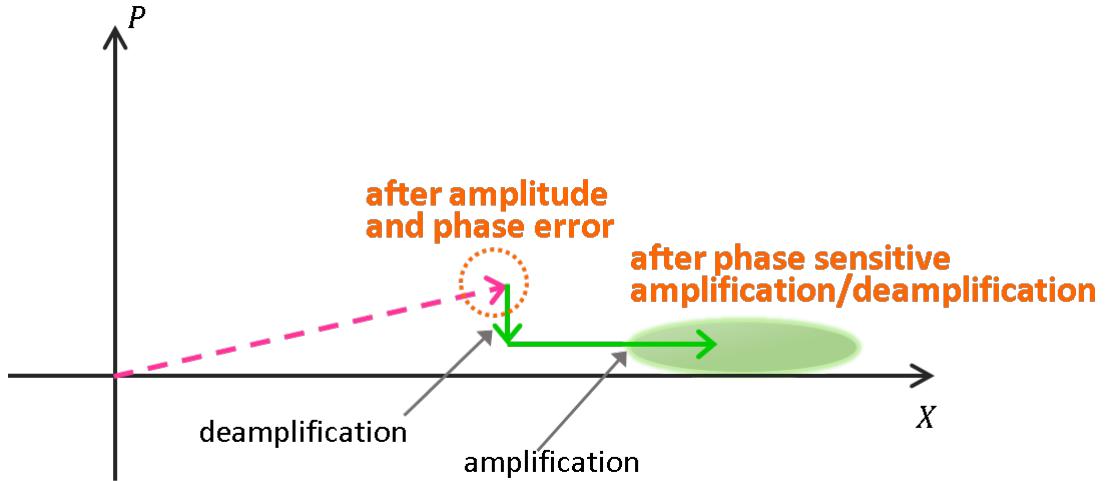


Figure 8: Amplitude and phase error correction by DOPO, in which the amplitude error is corrected by saturated amplification along X -axis and the phase error is corrected by phase sensitive deamplification along Y -axis.

1.2 Quantum synapses

We wish to implement a target Hamiltonian as a cost function using the network of quantum neurons. One target Hamiltonian is that of the Ising model [11]:

$$\mathcal{H} = - \sum_{i < j} J_{ij} \sigma_i \sigma_j - \sum_i h_i \sigma_i, \quad (6)$$

where the Ising spin σ_i takes either $+1$ (up-spin) or -1 (down-spin). As described above, we wish to represent $\sigma_i = 1$ and $\sigma_i = -1$ by the positive or negative in-phase amplitude, which corresponds to 0 -phase or π -phase oscillation, respectively (Fig. 3(b)). The Ising coupling J_{ij} and the Zeeman term h_i take (real) continuous values, which are determined by mapping a given real-world problem on the Ising model. Chapters VIII - XII describe various real-world problems that are mapped on the Ising Hamiltonian. The three-dimensional Ising model and the two-dimensional Ising model with local fields belong to the NP-hard class in complexity theory [12], and many hard problems can be solved through the Ising model. In order to implement the cost function described in E.(6) 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 [13].

1.2.1 Optical delay line coupling scheme

Figure 9 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 phase sensitive amplifier (PSA), split into multiple optical delay lines including intensity and phase modulators [14–16] and then injected back to the target DOPO pulse at appropriate timing. In the configuration shown in Fig. 9, N independent DOPOs are simultaneously realized as N optical pulses circulating in a fiber ring cavity with an internal PSA (PPLN waveguide optical parametric amplifier) which is driven externally by pump pulse trains.

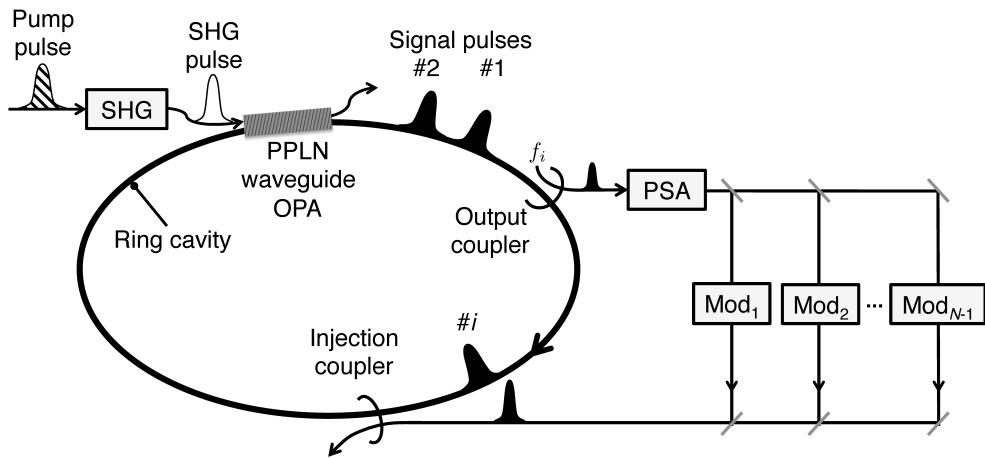


Figure 9: A quantum neural network based on the time-division multiplexed DOPO pulses with mutual coupling implemented by optical delay lines. A part of each pulse is picked off from the main cavity by the output coupler followed by an optical phase sensitive amplifier (PSA) which amplifies the in-phase 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 [14–17].

The external phase sensitive amplifier 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 such $N - 1$ optical delay lines, any (i th) pulse can be connected to any other (j th) pulse with a coupling constant J_{ij} . Such an all-optical coupling scheme has been experimentally demonstrated for $N = 4$ and 16 pulses using free-space optics [14, 15] and for $N = 10^4$ pulses using a planar lightwave circuit (PLC) [16]. In the last system, for example, the optical fiber length is 2 km so that a round trip time is $10 \mu\text{s}$ while the pulse repetition frequency is 1 GHz so that the pulse interval is 1 nsec, which leads to $N = 10^4$ DOPO pulses inside a single ring cavity [16].

This optical delay line coupling machine is capable of producing the quantum noise correlation among the internal DOPO pulses [18, 19]. In Chapter III, we will develop the quantum theory of the optical delay line coupling QNN and discuss this point in detail using the two measures for evaluating quantum noise correlation: quantum entanglement and discord.

1.2.2 Measurement feedback coupling scheme

An alternative experimental scheme to implement the Ising coupling J_{ij} is shown in Fig. 10 [20, 21]. 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 j -th DOPO pulse is represented by \tilde{X}_j , the feedback pulse to the i -th 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, which must be completed in the pulse interval time of 100 ps - 1 ns, is achieved by a single measurement-feedback circuit consisting of an analog-to-digital converter (ADC), field programmable gate array (FPGA), digital-to-analog converter (DAC) and optical amplitude/phase modulators [20, 21]. The feedback pulse and local oscillator pulse used for optical homodyne detection are both provided by a part of the pump pulse, as shown in Fig. 10.

Such a measurement feedback coupling scheme (Fig. 10) is equivalent to an optical delay line coupling scheme (Fig. 9) except for the following advantage/disadvantage. The advantage of the measurement-feedback scheme is that all-to-all coupling of the order of N^2 connection can be implemented by a single measurement feedback circuit, so that the daunting task of constructing an $N - 1$ 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 $\mathcal{H} = -\sum K_{ijk} \sigma_i \sigma_j \sigma_k$, can be readily implemented. The disadvantage of the measurement feedback scheme is that the FPGA circuit must complete $\sim O(N)$ 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 difference in the operational principles from quantum mechanical view-

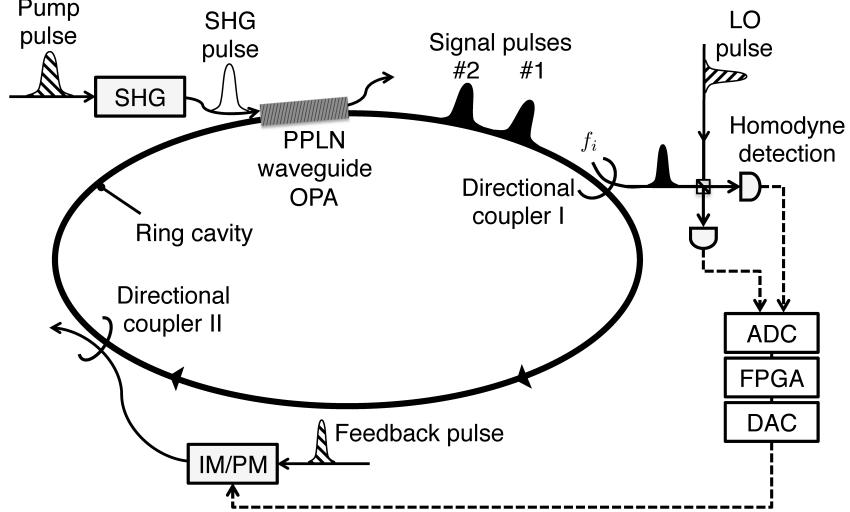


Figure 10: A quantum neural network with a measurement-feedback circuit. A small portion of each DOPO signal pulse is out-coupled through the directional coupler I, and its in-phase amplitude is measured by optical balanced homodyne detectors, where LO pulse is directly obtained from the pulsed pump laser. Two detector outputs are converted to digital signals and input into an electronic digital circuit, where a feedback signal for the i -th DOPO signal pulse is prepared. The feedback pulse also taken from the pump laser is modulated in its intensity and phase to achieve the target amplitude proportional to $\sum_j J_{ij} \hat{X}_j$ and coupled into i -th signal pulse by directional coupler II. Flows of optical fields and electrical signals are shown as solid and dashed lines, respectively [17, 20, 21].

points between the optical delay line coupling and measurement feedback coupling schemes. The measurement feedback scheme does not produce any entanglement. The total density operator stays in the product state of individual DOPO pulse density operators. However, the measurement feedback scheme is capable of implementing the non-unitary state reduction to realize non-Gaussian states induced by quantum measurements, which contributes to enhance quantum tunneling in the search process. In Chapter V, we will discuss this point in detail using the quantum theory of the measurements feedback QNN. Before we discuss these advanced subjects, let us describe the basic and common principle of the quantum neural networks in the next two sections.

1.3 Mapping of an Ising model to DOPO network: coherent Ising machines

In Chapters III and V, we will develop the full quantum theory of the quantum neural networks (QNN) with optical delay line coupling scheme and measurement feedback coupling scheme, respectively.

1.3.1 Pitchfork bifurcation

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

$$\dot{x}_j(t) = -\frac{\partial V}{\partial x_j}, \quad (7)$$

with

$$V = \sum_l [V_b(x_l)] + \epsilon V_l(x), \quad (8)$$

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 $V_l(x) = -\sum_{j,l} \omega_{jl} x_j x_l$. Here $t = \frac{\gamma_s}{2}\tau$ is the unitless time normalized by the signal field lifetime $2/\gamma_s$, γ_s is the signal photon decay rate, $p = F_p/F_{th}$ 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, $A_s = \sqrt{\gamma_s \gamma_p / 2\kappa^2}$ is the saturation amplitude which is equal to the DOPO oscillation amplitude at the normalized pump rate $p = 2$. $\alpha = -1 + p$ is the bifurcation parameter given by the normalized decay rate and linear gain for the signal field. Moreover, the parameter ϵ is small enough such that $0 < \epsilon \ll 1$. Equation (7) describes the dynamics of a gradient system. Under the approximation that quadrature-phase components are always zero, the equations of motion for a solitary DOPO are given as $\dot{x}_j = \alpha x_j - x_j^3$ when $\epsilon = 0$, which represents the normal form of the supercritical pitchfork bifurcation.

In the uncoupled case, i.e., $\epsilon = 0$, (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 E. (7) 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, there is a breaking of symmetry and the states x_j become either positive or negative. The emergence of the first nonzero stable steady state

can be studied by considering linear stability analysis at the origin, i.e., the Jacobian matrix J associated with E. (7) and computed at the state $x_j = 0, \forall j$. This Jacobian matrix is given as follows:

$$J = \alpha I + \epsilon \Omega, \quad (9)$$

where Ω is the matrix with components Ω_{jl} and I the identity matrix. The destabilization of the zero state corresponds to the emergence of the first nonzero 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 μ_1 is the maximal eigenvalue of the matrix Ω , which we consider unique for the sake of simplicity. The first nonzero steady state becomes stable when $\lambda_1 = 0$, i.e., when the bifurcation parameter α is equal to α_c given as follows [13]:

$$\alpha_c = -\epsilon \mu_1. \quad (10)$$

As the bifurcation parameter α increases further above the threshold α_c , other states become stable and the system develops a complex attractor landscape. Note that E.(10) suggests that $\alpha_c < 0$ in general.

1.3.2 Conditional mapping of the Ising Hamiltonian

The linear stability analysis of the zero state allows determining the bifurcation parameter α_c for which the zero state becomes unstable. Moreover, the eigenvector associated with the eigenvalue λ_1 indicates the locations of the first nonzero stable steady states just after the bifurcation has occurred and, in the 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 nonzero stable steady states map rigorously the ground-state configurations of the Ising Hamiltonian. Indeed, the steady-state condition of E.(7) is given as follows when $x_j = \sigma_j x, \forall j$, with $x > 0$:

$$\begin{aligned} \dot{x}_j = 0 &\Rightarrow \alpha x_j - x_j^3 + \epsilon \sum_{l \neq j} \omega_{jl} x_l = 0 \\ &\Rightarrow \alpha - x^2 + \epsilon \sum_{l \neq j} \omega_{jl} \sigma_l \sigma_j = 0. \end{aligned} \quad (11)$$

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

$$x^2 = \alpha - \frac{2\epsilon}{N}\mathcal{H}, \quad (12)$$

with the Ising Hamiltonian \mathcal{H} without a dc field. When $\alpha = \frac{2\epsilon}{N}\mathcal{H}$, the only solution of E.(12) is $\mathbf{x} = \mathbf{0}$. As α increases, the first nonzero steady state is reached for the minimal value of the function $\frac{2\epsilon}{N}\mathcal{H}$ as shown in E.(12) and Fig. 11, i.e., in the regions of the state space that encode for the ground-state configurations of the Ising Hamiltonian. In a nutshell, the first nonzero stable steady states of E.(11) encode for the ground-state configurations of the Ising Hamiltonian if the steady-state amplitudes $|x_j|, \forall j$, are all equal.

1.3.3 Effect of the amplitude heterogeneity

In the 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. In practice, approximate solutions to various combinatorial optimization problems can still be obtained under this approximate mapping [13, 22]. In order to quantify the distance between the case of exact mapping and the 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:

$$\delta_x = \frac{\sqrt{\langle \delta_j^2 \rangle}}{\langle x_j^2 \rangle}, \quad (13)$$

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 $\alpha > 0$ and $\epsilon \ll 1$, the amplitudes x_j can be expressed as $x_j = x_j^{(0)} + \epsilon x_j^{(1)} + \mathcal{O}(\epsilon^2)$ with $x_j^{(0)} = \sigma_j \sqrt{\alpha}$ and $x_j^{(1)} = \frac{1}{2\sqrt{\alpha}} \sum_l \omega_{jl} \sigma_l$ using the perturbation theory in order to find an approximate solution to the steady-state E.(11). Thus, the squared amplitudes can be expressed as follows in the general case:

$$\begin{aligned} x_j^2 &= \left[x_j^{(0)} + \epsilon x_j^{(1)} + \mathcal{O}(\epsilon^2) \right]^2 \\ &= \alpha + \epsilon \sum_l \omega_{jl} \sigma_j \sigma_l + \mathcal{O}(\epsilon^2). \end{aligned} \quad (14)$$

Using E.(14), we can express x_j^2 as follows:

$$x_j^2 = \langle x^2 \rangle + \delta_j + \mathcal{O}(\epsilon^2), \quad (15)$$

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

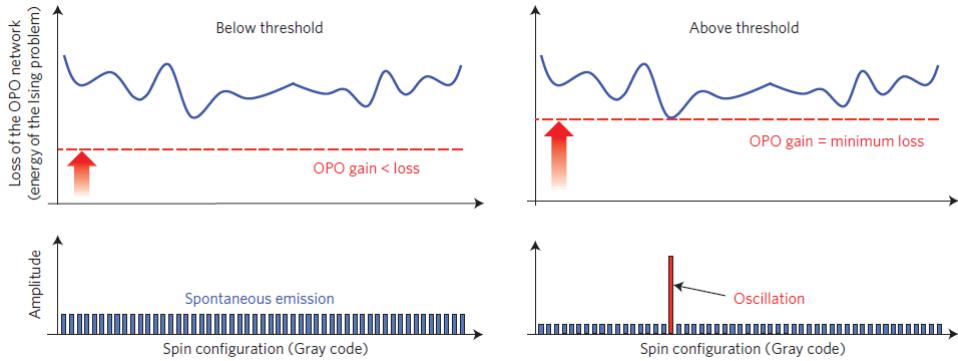


Figure 11: Principle of operation of the QNN for Ising problems. Gradual pumping of the DOPO network makes it oscillate at the ground state if all DOPOs have identical absolute amplitudes. The parametric gain is gradually increased from below threshold to reach the minimum loss rate of the ground state. Because the lowest loss of the network corresponds to the spin configuration of the ground state of the implemented Ising problem, QNN is expected to find the ground state as a single oscillation mode [14].

Figure 12(a) shows the modulated threshold gain α_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) [13]. In this problem, each vertex has three edges with constant weight $\omega_{ij} = -0.1$ as shown in Fig. 12(b). At high pump rates $p \geq 1.3$, the possible value of α_c is minimum for the ground state and increase monotonically with the order of excited states. That is, the mapping is successful. However, the threshold gain α_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 12(b) shows the steady state amplitude (by circle size) and phase (by color) of each DOPO for the first excited state at $p < 1$. Five of the total eight DOPOs mutually couple without destructive interference so that they 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 earlier 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 [23]. 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 wrong minimum to a correct 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. 12(a).

In Chapter VII, we will describe a technique for how to ensure equal amplitudes of DOPOs at pump rates lower than the threshold so that the ground state of the Ising Hamiltonian is selected by a single oscillation mode.

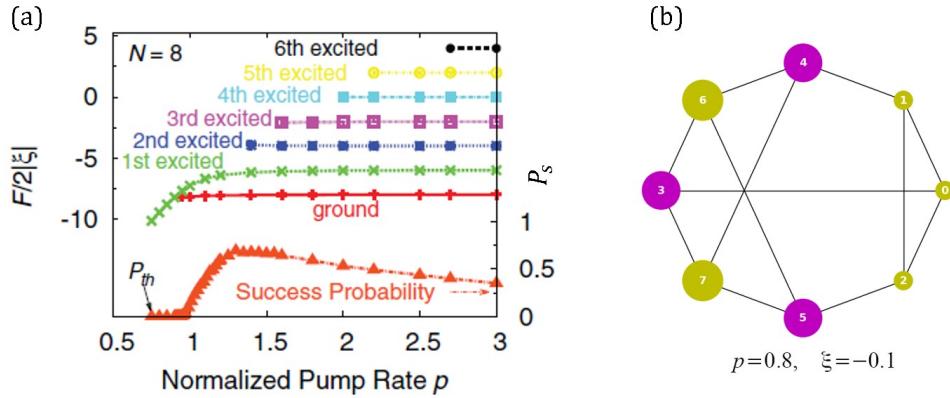


Figure 12: (a) The dependence of the success probability P_s and the modulated threshold gain α_c of the whole network vs. the normalized pump rate p when $\omega_{ij} = -0.1$ for the worst-case instance with $N = 8$ vertices. (b) The inhomogeneous oscillation amplitudes for the first excited state at a pump rate $p = 0.8$. The size and color of the circle represent the amplitude and phase of each DOPO [13].

1.4 Optical neural network operating at the quantum limit and classical limit

One of the unique features of optical neural networks is the continuous crossover of their operational modes from the quantum limit to the classical limit. In this section, we will present various numerical results showing the difference between the two regimes, i.e. quantum neural network (QNN) vs. classical neural network (CNN).

Figure 13 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 based coherent Ising machine (CIM) consisting of anti-ferromagnetically coupled two spins [10]. 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. 13. The CIM with a high-Q or low-Q cavity, in which a round trip loss is either 10% (-0.5 dB) or 50% (-3 dB), continuously excites the quantum states which are close to the Heisenberg limit, as shown by red or green line in Fig. 13. A pump rate is linearly increased from below to above the DOPO threshold. The initial state is a vacuum state $|0\rangle$ and the final state is a coherent state $|\alpha\rangle$, so that both of them satisfy $\langle\Delta\hat{X}^2\rangle = \langle\Delta\hat{P}^2\rangle = 1/4$. This result suggests that the approximate measurement of \hat{X} per every round trip pumps out an extra-entropy from the DOPO system and the whole computation is performed with the quantum states close to the pure states. On the other hand, optical neural networks at the thermal noise limit ($k_B T \gg \hbar\omega$) should operate in the classical regime defined by $\langle\Delta\hat{X}^2\rangle \geq 1/4$ and $\langle\Delta\hat{P}^2\rangle \geq 1/4$, which is shown by the shaded area in Fig. 13.

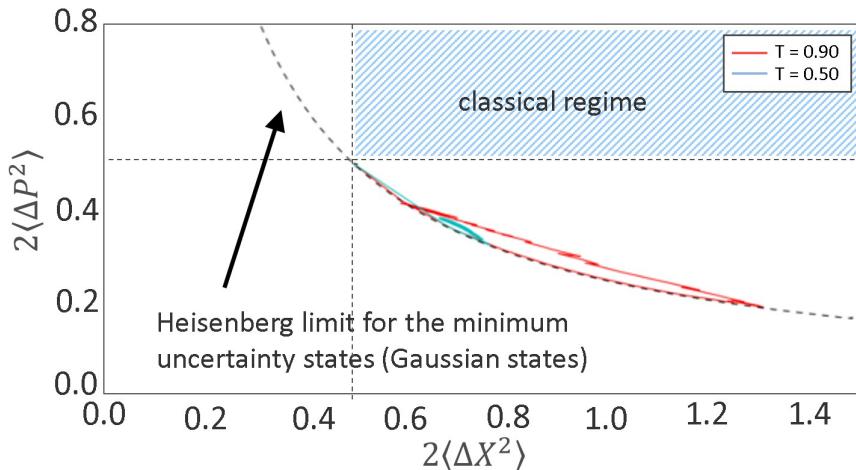


Figure 13: The trajectory of $2\langle\Delta\hat{X}^2\rangle$ and $2\langle\Delta\hat{P}^2\rangle$ of the high-Q or low-Q cavity CIM consisting of two anti-ferromagnetically coupled spins [10].

The success probability P_s of finding the ground state of $N = 16$ one-dimensional Ising spin model, in which only nearest neighbor anti-ferromagnetic interaction exists, is numerically evaluated for various temperature parameters $n_{th} = k_B T / \hbar\omega$ and is shown in Fig. 14 [19]. The optical delay line coupling QNN 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 of each DOPO realizes the quantum par-

allel search and establishes the quantum correlation (entanglement) among DOPOs during a transient time before the steady state amplitude is formed [18, 19] so that the maximum success probability is achieved at a final pump rate just above the threshold ($p_0 \simeq p_{th} = 1$), which allows the network to have a sufficient time to search the solution with quantum noise. 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, before the quantum parallel search establishes the quantum noise correlation and identifies a correct solution before each DOPO oscillates. The quantum noise induced tunneling is also not strong enough to overcome the potential barrier separating 0-phase and π -phase when the oscillation field is strong. In this way, the network is trapped in one of the excited states (local minima). In the case of $\hbar\omega \ll k_B T$ (thermal noise limit), the squeezed thermal state, which is the statistical mixture of the amplitude eigenstates $|X\rangle$, does not allow the quantum parallel search during a transient time before the steady state amplitude is formed so that the maximum success probability is achieved at a final pump rate well above the threshold ($p_0 \gg p_{th} = 1$), where the coherent mean-field searches a solution against thermal noise as classical neural networks.

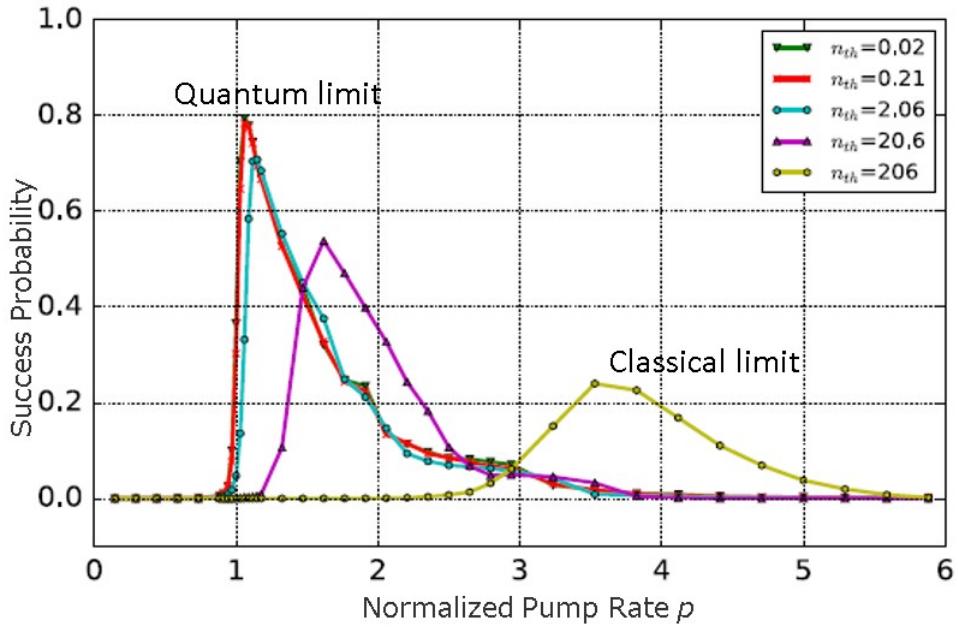


Figure 14: The success probability of finding a solution for $N = 16$ one-dimensional Ising spin model for various temperature parameters $k_B T/\hbar\omega$ [19].

In classical neural networks, the neuron state (in-phase amplitude X) has a pre-determined value before a measurement is actually made. Thus, the action of measurement

does not change this value of X no matter what value is reported by the detector. On the other hand, the neuron state in quantum neural networks does not have a pre-determined X value. The measurement result \tilde{X} for the observable \hat{X} is determined by the joint action of the measured neuron (DOPO field) and the measuring devices (homodyne detectors). In this case, the post-measurement neuron state is changed by the specific measurement result (partial reduction of a wavepacket) and this effect can be computed by the so-called projection operator.

Figure 15 shows the success rate P_s to find the ground state vs. problem size N for a generalized Ising model by a measurement-feedback based CIM. The numerical simulation employs the c -number stochastic differential equations (CSDE) and the replicator equations, which are derived from the measurement-feedback master equation [26]. The success rate P_s decreases exponentially with the problem size N and $P_s \sim 0.1$ at $N = 2000$. If the partial reduction of the DOPO wavepacket induced by the approximate measurement of \hat{X} is neglected, the above quantum equations are reduced to the classical Langevin equations. Figure 15 also shows the success rate P_s vs. the problem size N by the classical Langevin equations. The success rate P_s decreases more rapidly in this case and $P_s < 0.01$ at $N = 2000$, which highlights the importance of the measurement-induced wavepacket reduction in the CIM.

1.5 Gottesman-Knill theorem and non-Gaussian wavepackets

Not all quantum dynamics are difficult to simulate by classical digital computers. Some of representative quantum processes, including entangled state generation and purification, can be efficiently simulated by classical methods, so that such a quantum process alone does not likely achieve an exceeding computation power compared to 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 [24]. The statement of the Gottesman-Knill theorem can be summarized as follows:

If a particular quantum process starts with

1. computational basis states, such as all ground states $|0\rangle_1|0\rangle_2\cdots|0\rangle_N$,

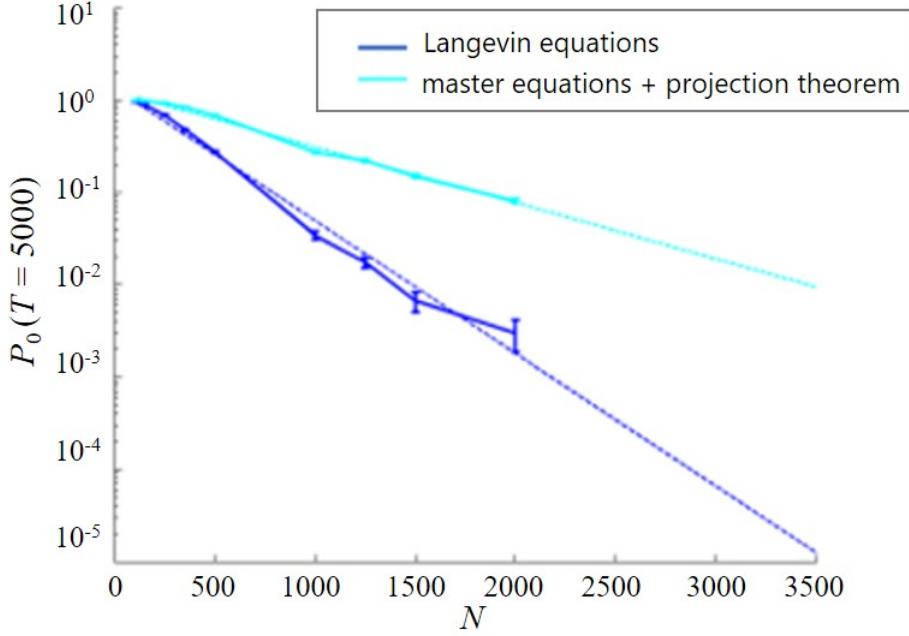


Figure 15: The success rate P_s to find the ground state vs. the problem size N . Quantum neural network (QNN) is described by the c -number stochastic differential equations at the replicator equations, while classical neural network (CNN) is described by the classical Langevin equations.

2. employs a limited set of unitary evolutions such as Hadamard gate, phase gate and controlled-NOT gate (Clifford group),

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 to implement 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. [25]. The statement of this theorem runs as follows:

If a quantum process with harmonic oscillators starts with

1. Gaussian states, such as all coherent states $|\alpha\rangle_1 |\alpha\rangle_2 \cdots |\alpha\rangle_N$,
2. employs a limited set of unitary evolutions such as squeezing gate and displacement gate,

and ends with

3. projective measurements of one quadrature amplitude (by homodyne detection) or simultaneous measurement of 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 measurement-feedback CIM (QNN), we can identify the gain saturation (or two photon loss) as the essential quantum dynamics which make the QNN difficult to simulate efficiently by classical methods. You can find the gain saturation term, $-(x_j^2 + p_j^2)x_j$, in E.(11) provide the non-classical nature to QNN.

Figure 16(a) shows the third-order dispersion $\langle \Delta \hat{X}^3 \rangle$ of the in-phase amplitude vs. the normalized pump rate p for the $N = 2$ anti-ferromagnetically coupled spins [26]. $\langle \Delta \hat{X}^3 \rangle$ is identically equal to zero for a Gaussian wavepacket, which is clearly observed at below the DOPO threshold in Fig. 16(a). However at above the threshold, $\langle \Delta \hat{X}^3 \rangle$ of the two DOPOs bifurcate to either positive or negative value. This is because the DOPO wavepacket at above the threshold has a rapidly decaying tail at larger amplitudes but a slowly decaying tail toward smaller amplitudes due to the highly asymmetric potential profile, as shown in Fig. 16(b). As a result of the asymmetric (non-Gaussian) wavepackets, the quantum tunneling easily occurs between 0-phase and π -phase states at the threshold region. Figure 17 confirms this effect. In the QNN shown in Fig. 17(a), the switching between the two degenerate ground states occurs frequently. However, in the Gaussian wavepacket approximation shown in Fig. 17(b), such a switching behavior is highly suppressed due to the rapid decay of the tail toward smaller amplitudes.

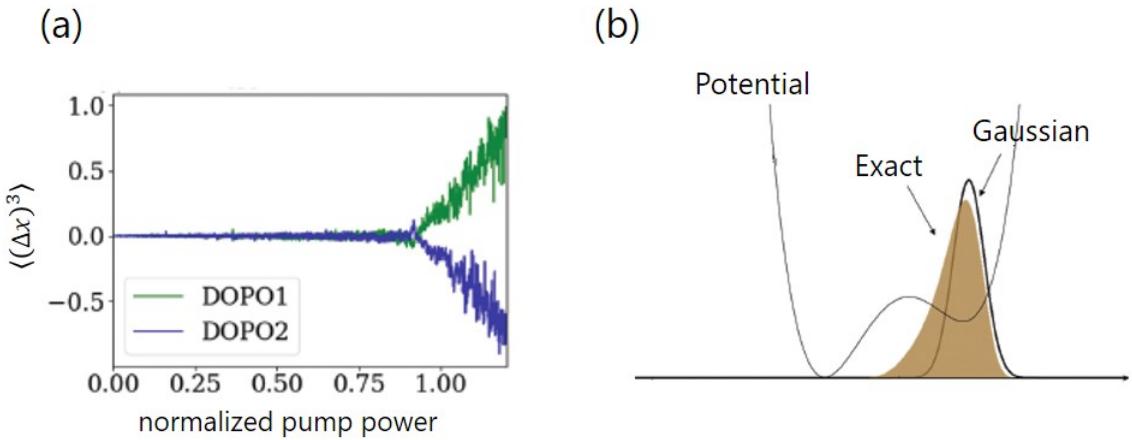


Figure 16: (a) The third-order dispersion $\langle \Delta \hat{X}^3 \rangle$ of the in-phase amplitude in the measurement-feedback CIM (QNN). (b) The non-Gaussian wavepacket at just above the threshold due to the turn-on of gain saturation [26].

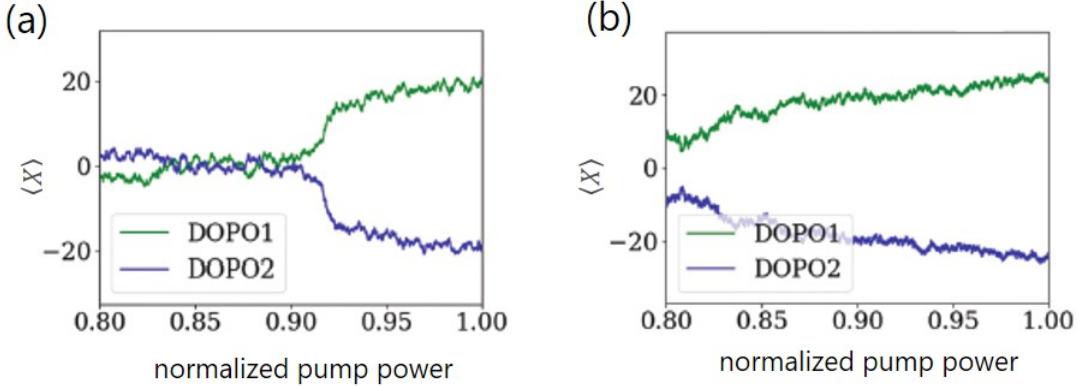


Figure 17: The average in-phase amplitudes $\langle \hat{X}_1 \rangle$ and $\langle \hat{X}_2 \rangle$ of the $N = 2$ anti-ferromagnetically coupled DOPOs. (a) Exact solution with non-Gaussian wavepackets. (b) Approximate solution with Gaussian wavepackets [26].

Quantum tunneling induced by the non-Gaussian wavepackets near the threshold has a substantial impact on the success rate of QNN. Figure 18 shows the success rate P_s to find the ground state vs. the anti-ferromagnetic coupling strength for $N = 16$ one dimensional Ising spin ring. When the coupling strength is weak and the feedback signal is overwhelmed by the background vacuum fluctuation, the success rate P_s is small. When the coupling strength is large and the feedback signal dominates over the vacuum fluctuation, the success rate P_s increases, where P_s with the non-Gaussian wavepackets exceeds P_s with the Gaussian wavepackets. This result suggests the important role of the quantum tunneling induced by non-Gaussian wavepackets.

1.6 Summary

Some of the important conclusions in Chapter I are summarized below.

1. Optical neural networks at room temperatures operate at $k_B T / \hbar \omega \ll 1$ and can realize a quantum parallel search. Such quantum neural networks (QNN) are composed of the two constituent devices: quantum neurons and quantum synapses.
2. 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 and the quantum-to-classical amplification of computational results at above the threshold.

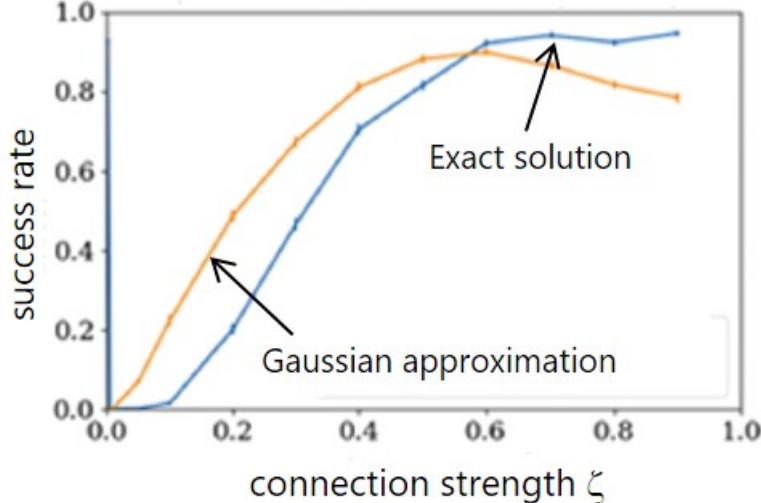


Figure 18: The success rates P_s by the exact theory (non-Gaussian wavepackets) and the Gaussian approximation vs. Ising coupling strength for the $N = 16$ one-dimensional Ising model [26].

3. Quantum synapses are provided by either direct coupling with optical delay lines or indirect coupling with measurement-feedback circuits. They utilize distinct computational resources: quantum noise correlation (entanglement) in the optical delay line coupling QNN and measurement-induced wavepacket reduction to non-Gaussian states in the measurement feedback QNN. We will develop the quantum theory of the two types of QNN in Chapters III and V.
4. QNN can solve various combinatorial optimization problems by mapping them either on NP-hard Ising problems or NP-complete k -SAT problems. The former system, called a coherent Ising machine, is constructed as a symmetric continuous-time homogeneous neural network and the latter system, called a coherent SAT-machine, is realized as an asymmetric continuous-time recurrent neural network. The details of these two machines will be presented in Chapters VI and VII (Chapter VII is in preparation).
5. Optical neural networks at quantum limit (QNN) can outperform those at classical limit (CNN) by exploiting superposition states for quantum search of solutions and quantum suppression of classical chaos.
6. The gain saturation and single photon loss are the two indispensable elements to make QNN hard to simulate by classical methods. These two dissipation processes

and associated fluctuations from external reservoirs are crucial resources to accelerate the search process.

- [1] J. J. Hopeld and D. W. Tank, *Science* 233, 625 (1986).
- [2] D. F. Walls and G. J. Milburn, *Quantum Optics* (Springer Berlin-Heidelberg, 1995).
- [3] R. L. Byer et al., *Appl. Phys. Lett.* 13, 109 (1968).
- [4] D. Serkland et al., *Opt. Lett.* 20, 1649 (1995).
- [5] Y. Nambu, *Rev. Mod. Phys.* 81, 1015 (2009).
- [6] L. Mandel and E. Wolf, *Optical Coherence and Quantum Optics* (Cambridge Univ. Press, 1995).
- [7] M. Wolinsky and H. J. Carmichael, *Phys. Rev. Lett.* 60, 1836 (1988).
- [8] L. Krippler et al., *Phys. Rev. A* 50, 4330 (1994).
- [9] P. D. Drummond et al., *Phys. Rev A* 65, 033806 (2002).
- [10] A. Yamamura et al., *Phys. Rev. A* 96, 053834 (2017).
- [11] E. Ising, *Zeitschrift fur Physik A* 31, 253 (1925).
- [12] F. Barahona, *J. Phys. Math. Gen.* 15, 3241 (1982).
- [13] Z. Wang et al., *Phys Rev. A* 88, 063853 (2013).
- [14] A. Marandi et al., *Nature Photonics* 8, 937 (2014).
- [15] K. Takata et al., *Sci. Rep.* 6, 34089 (2016).
- [16] T. Inagaki et al., *Nature Photonics* 10, 415 (2016).
- [17] Y. Haribara et al., *Entropy* 18, 151 (2016).
- [18] K. Takata et al., *Phys. Rev. A* 92, 043821 (2015).
- [19] D. Maruo et al., *Phys. Scr.* 91, 083010 (2016).
- [20] T. Inagaki et al., *Science* 354, 603 (2016).
- [21] P. McMahon et al., *Science* 354, 614 (2016).
- [22] T. Leleu et al., *Phys. Rev. E* 95, 022118 (2017).
- [23] P. Kinsler and P. D. Drummond *Phys Rev. A* 43, 6193 (1991).
- [24] D. Gottesman, arXiv:9807006 (1998).
- [25] S. D. Bartlett et al., arXiv:0109047v2 (2002).

[26] T. Shoji et al., Phys. Rev. A 96, 053833 (2017).

Written by Y. Yamamoto and T. Leleu

version 2