


16-channel photonic solver for optimization problems on a silicon chip



Jiayi Ouyang^{1,†}, Shengping Liu^{2,†}, Ziyue Yang¹, Wei Wang², Xue Feng^{1,*} , Yongzhuo Li¹ & Yidong Huang^{1,*}

¹Department of Electronic Engineering, Tsinghua University, Beijing 100084, China ²Chongqing United Microelectronics Center, Chongqing 100290, China

[†]These authors have equal contributions to this work.

*E-mails: x-feng@tsinghua.edu.cn (Xue Feng), yidonghuang@tsinghua.edu.cn (Yidong Huang)

Cite as: Ouyang, J. et al. 16-channel photonic solver for optimization problems on a silicon chip. *Chip* 4, 100117 (2025).

Received: 12 August 2024

Accepted: 1 November 2024

Published online: 8 November 2024

A programmable photonic solver for quadratic unconstrained binary optimization (QUBO) problems is demonstrated with a hybrid optoelectronic scheme, which consists of a photonic chip and an electronic driving board. The photonic chip is employed to perform the optical vector-matrix multiplication (OVMM) to calculate the cost function of the QUBO problem, while the electronic processor runs the heuristic algorithm to search for the optimal solution. Due to the parallel and low-latency propagation of lightwaves, the calculation of the cost function can be accelerated. The photonic chip was fabricated on the silicon on insulator (SOI) substrate and integrated 16 high-speed electro-optic modulators, 88 thermo-optic phase shifters, and 16 balanced photodetectors. The computing speed of the photonic chip is 1.66 TFLOP/s. As a proof of principle, two randomly generated 16-dimensional QUBO problems are solved with high successful probabilities. These results present the potential of fast-solving optimization problems with integrated photonic systems.

Keywords: Optical computing, Optimization problem, Integrated photonics

INTRODUCTION

The rapid advances in artificial intelligence (AI) desire huge computation resources. Although the training of large-scale neural networks has been significantly improved with electronic hardware, *e.g.* the graphics processing unit, the energy consumption is still heavy. Recently, various optical computing systems have been proposed to demonstrate the optical neural networks (ONNs) due to the superiority of the high-speed and low-loss parallel propagation of lightwaves. These platforms are based on spatial diffractive layers^{1–3}, on-chip optical matrix multiplications^{4–6}, on-chip diffractive optics^{7–9}, etc. Such optical

computing systems exhibit high computing speed and energy efficiency in specific computation tasks, including vowel classification, image classification, or content generation. Besides the identification and classification, the nondeterministic polynomial hard (NP-hard) optimization problem is also important in the AI domain. For the NP-hard optimization problem, it is required to find the input state that minimizes the cost function. Such problems are ubiquitous but important in physics¹⁰, finance¹¹, biology^{12,13}, etc., including the subset sum problem^{14,15}, the Ising problem¹⁶, and the traveling salesman problem (TSP)¹⁷. The solving of NP-hard optimization problems can also be accelerated with optical computing systems¹⁶, including those based on optical resonators^{18–24}, polaritons²⁵, laser networks²⁶, and optical matrix multiplications^{27–35}. Different from ONNs, the optical parameters have to be precisely configured in the optical solver to conduct a specific transmission function corresponding to the given optimization problem. In addition, the ability to map any arbitrarily given problems requires that the transfer function be highly programmable and reconfigurable. However, optical solvers based on long fiber resonators^{20–23} or complicated spatial-optical systems^{26–29,33–35} usually face challenges of large space occupation, high energy consumption, and ambient and vibrational destabilization. Besides, it is difficult to integrate multiple such systems to form a large-scale optical solver. One possible technical approach is to implement optical solvers with photonic chips. For instance, the subset sum problem can be solved with the silicon photonic circuits consisting of split junctions^{14,15}, however, the employed circuit is only for specific problems and is not reconfigurable. In the photonic recurrent Ising sampler (PRIS)^{30,31}, a programmable photonic circuit that can perform arbitrary OVMM is first employed to solve an arbitrary Ising problem, which shows the feasibility of implementing a high-speed optical solver with a photonic chip. However, only 4-spin problems with the solution space of 2^4 are solved, which are insufficient for practical applications. In addition, an integrated photonic solver based on tunable delay lines was proposed theoretically to solve large-scale TSPs, but only numerical simulations were performed¹⁷. It can be seen that, how to extend the dimensionality of the on-chip photonic solver still needs to be addressed, although the previous photonic solvers have shown the potential to achieve high computing speed and energy efficiency.

In this work, we proposed and demonstrated a photonic solver for QUBO problems, which is based on a hybrid optoelectronic scheme including a photonic chip for the OVMM and an electronic driving circuit. Two 16-dimensional QUBO problems, whose complexity is equivalent to 16-spin Ising problems³⁶, are successfully solved in the experiment. With the eigendecomposition of the weight matrix of the

QUBO problem, the calculation of the cost function can be accelerated with the on-chip OVMM. To conduct the required OVMM, the photonic chip integrated 16 optical amplitude modulators, a Mach–Zehnder interferometer (MZI) array including 64 multimode interference (MMI) couplers and 88 thermo-optic phase shifters, and 16 balanced photo-detectors (BPDs). The electronic driving circuit, which mainly comprises a field-programmable gate array (FPGA), was employed to control the photonic circuits, process the optical signals, and conduct the necessary heuristic algorithm. During the operation of the photonic solver, each single iteration takes 265.1 ns, and the computing speed of the photonic chip is 1.66 TFLOP/s. These results indicate that the process of solving the optimization problems can be accelerated with photonic chips.

ARCHITECTURE OF THE PHOTONIC SOLVER

The proposed photonic QUBO solver employs a hybrid optoelectronic architecture shown in Fig. 1, which consists of a 1550-nm laser (ID Photonics Model CBDX1-1-SC-FA), a photonic chip, and an electronic driving board. The photonic chip was fabricated on the silicon on insulator (SOI) substrate with the 130 nm process (CSiP130C) of silicon photonics in Chongqing United Microelectronics Center (CUMEC). The process available in CSiP130C consists of 3 steps of silicon etching, 4 levels of n-type and p-type doping, Ge epitaxy, TiN metal heater, doped Si heater, Tungsten plug contact, 2 levels of copper interconnections, Al bond pad, and deep etching. The photonic chip consists of coupling gratings, beamsplitters, amplitude modulators, an MZI array, mixers, and BPDs. Here, each electro-optic amplitude modulator consists of a 1×2 and a 2×2 multimode interference (MMI) couplers, and a phase shifter on each arm combined with traveling-wave electrodes and p-n junctions to achieve high modulation speed. Each mixer is a 2×2 MMI coupler. On the chip, the beam was first split into the signal and reference beams. The signal beam was further split into $N = 16$ beams as the input of the MZI array. The driving board is composed of an FPGA (Xilinx Zynq UltraScale + RFSoc XCZU29DR), digital-to-analog converters (DACs, Analog Devices AD5767) for the MZI array, operational amplifiers

(OAs, Analog Devices ADA4927-2), and trans-impedance amplifiers (TIAs, Texas Instruments LMH32401). The FPGA controls the voltages of the MZI array and the amplitude modulators. The optical signals from the MZI array are detected by the BPDs, and are finally received by the FPGA via the TIAs and the internal analog-to-digital converters (ADCs). Furthermore, the FPGA executed the heuristic algorithm. To obtain stable OVMM, a cooling module was employed on the photonic chip to maintain the chip temperature at 37 °C during the operation.

Our proposed photonic solver is employed to solve the NP-hard QUBO problem, where the N -dimensional binary vector $s \in \{0, 1\}^N$ needs to be found to minimize the cost function $C(s)^{37}$:

$$C(s) = -\frac{1}{2}s^T K s. \tag{1}$$

In Eq. (1), K is a real symmetric weight matrix, and the superscript T denotes the transpose. The weight matrix K can be decomposed by $K = Q^T D Q$, where D is the diagonal eigenvalue matrix and Q denotes the orthogonal eigenvector matrix. If all the eigenvalues of K are non-negative, the cost function C can be written as follows:

$$C = -\frac{1}{2}s^T Q^T D Q s = -\frac{1}{2}(A s)^T (A s), \tag{2}$$

where $A = \sqrt{D} Q$. In the OVMM, the input vector is set to s and the transformation matrix is set to A . Denoting the output vector of the OVMM by $E_{out} = A s$, Eq. (2) becomes

$$C = -\frac{1}{2} \sum_{i=1}^N |E_{out}^{(i)}|^2. \tag{3}$$

Eq. (3) indicates that the calculation of the cost function is accelerated by one step of OVMM. When solving a QUBO problem, the OVMM of the photonic chip is configured to A . In each iteration, the cost function is calculated and the spins are changed according to the adopted heuristic algorithm. In the experiment, the photonic solver executes a simulated-

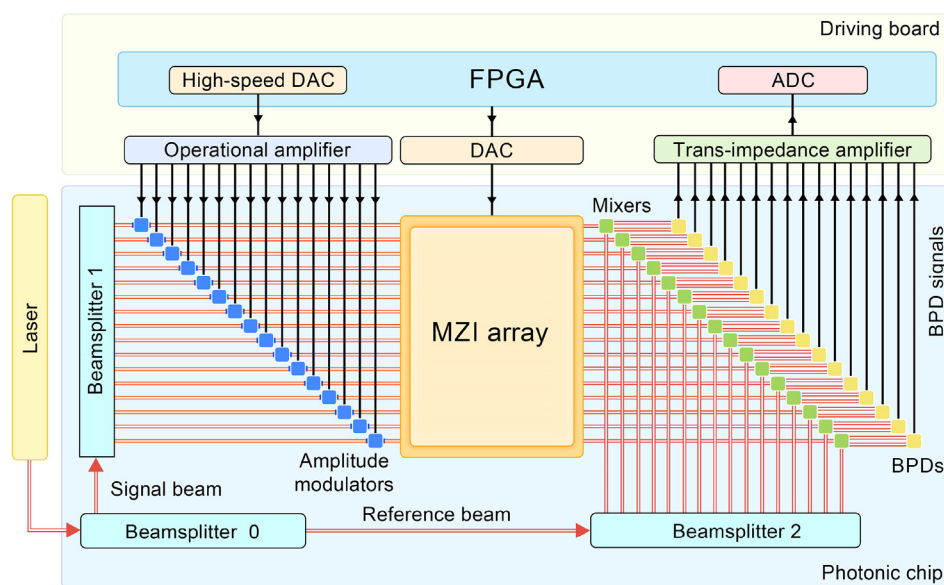


Fig. 1 | The architecture of the photonic solver.

annealing-like algorithm³⁴ to search for the near-optimal solutions of QUBO problems.

The OVMM on the photonic chip is based on the pseudo-real-value architecture³⁸, which has lower loss and higher scalability than the singular-value-decomposition architecture³⁸. Moreover, the balanced detection can reduce the signal noise. As a trade-off, the pseudo-real-value architecture has lower reconfigurability and can only perform real matrix transformations³⁸. This architecture with 4 channels is employed to demonstrate vision tasks in the previous work³⁸, while we only focus on solving QUBO problems in the current work. The size of the photonic chip is $7.5 \times 5 \text{ mm}^2$, as shown in the microscope image of Fig. 2a. Fig. 2b and c are the scanning electron microscope (SEM) images of a coupling grating from the cross-section view and the top view respectively. Fig. 2d and e depict the cross-section images of the waveguides. Fig. 2f shows the microscope image of the beamsplitter, and the SEM image of a 1×2 MMI coupler in the beamsplitter is shown in Fig. 2g. Fig. 2h shows the microscope image of a mixer. The 16-dimensional binary input vector s is encoded on the amplitudes of the signal beams with the electro-optic amplitude modulators. Then the signal beams are guided to the MZI array. The structure of the Fast-Fourier-transform-mesh (FFT-mesh) MZI array is shown in Fig. 3, which consists of 16 input ports, 4 layers of MZIs, and 16 output ports. In the MZI array, there are 88 thermo-optic phase shifters, which determine the transformation matrix. The cross-connections between the MZIs are implemented with crossing waveguides. To enhance the spurious-free dynamic range of the MZI array, the “equal length, equal loss” design was employed, which means that the light passes through the same optical path and the same number of crossing waveguides from any input port to any output port. Such a design provides robustness against temperature fluctuation. The detailed design of the MZI array is provided in the Supplemental Material.

Theoretically, the MZI array can perform a 16×16 unitary matrix transformation $U(V)$, which is the function of the bias voltages $V = [v_1, v_2, \dots, v_{88}]$ on the thermo-optic phase shifters. The output complex amplitude vector of the MZI array is $E_{\text{out}} = Us$. As shown in Fig. 1,

each output beam of the MZI array is input to a mixer along with 1 of 16 reference beams from beamsplitter 2. In mixer i ($i = 1, 2, \dots, 16$), the complex amplitude of the beam from the MZI array is denoted by $|E_{\text{out}}^{(i)}| \exp(i\phi_{\text{out}}^{(i)})$, and that of the reference beam is $|E_{\text{ref}}| \exp(i\phi_{\text{ref}}^{(i)})$ (the amplitudes of the reference beams are the same). Therefore, the intensities of the two output ports of mixer i are

$$I_{\pm}^{(i)} = |E_{\text{ref}}|^2 + |E_{\text{out}}^{(i)}|^2 \pm 2|E_{\text{ref}}||E_{\text{out}}^{(i)}| \cos(\phi_{\text{out}}^{(i)} - \phi_{\text{ref}}^{(i)}), \quad (4)$$

which would then be detected by BPD i , resulting in the differential current signal

$$I_{\text{BPD}}^{(i)} = I_{+}^{(i)} - I_{-}^{(i)} = 4|E_{\text{ref}}||E_{\text{out}}^{(i)}| \cos(\phi_{\text{out}}^{(i)} - \phi_{\text{ref}}^{(i)}). \quad (5)$$

Since $|E_{\text{out}}^{(i)}| \cos(\phi_{\text{out}}^{(i)}) = \text{Re}(E_{\text{out}}^{(i)})$ ($\text{Re}(\cdot)$ represents taking the real part), by setting $\phi_{\text{ref}}^{(i)}$ to 0, $I_{\text{BPD}}^{(i)} = 4|E_{\text{ref}}| \text{Re}(E_{\text{out}}^{(i)})$. The input vector s consists of real elements $s_i \in \{0, 1\}$, hence $\text{Re}(E_{\text{out}}) = \text{Re}(Us) = \text{Re}(U)s$, and we have

$$I_{\text{BPD}} = 4|E_{\text{ref}}| \text{Re}(E_{\text{out}}) = 4|E_{\text{ref}}| \text{Re}(U)s. \quad (6)$$

Eq. (6) indicates that, if the BPD signals I_{BPD} is regarded as the output vector, the OVMM is indeed conducting a real matrix transformation $A \propto \text{Re}(U)$ on the input vector s . To avoid confusion, in the following discussion, the output vector denotes the BPD signals, and the real matrix denotes the corresponding real transformation matrix A .

According to Eq. (2), any transformation matrix A corresponds to a QUBO problem with the weight matrix $K = A^T A$, which is positive semi-definite since $x^T K x = (Ax)^T Ax \geq 0, \forall x \in \mathbb{R}^N$. Thus, according to Eqs. (3) and (6), the experimental cost function C_{exp} can be calculated from the output signal

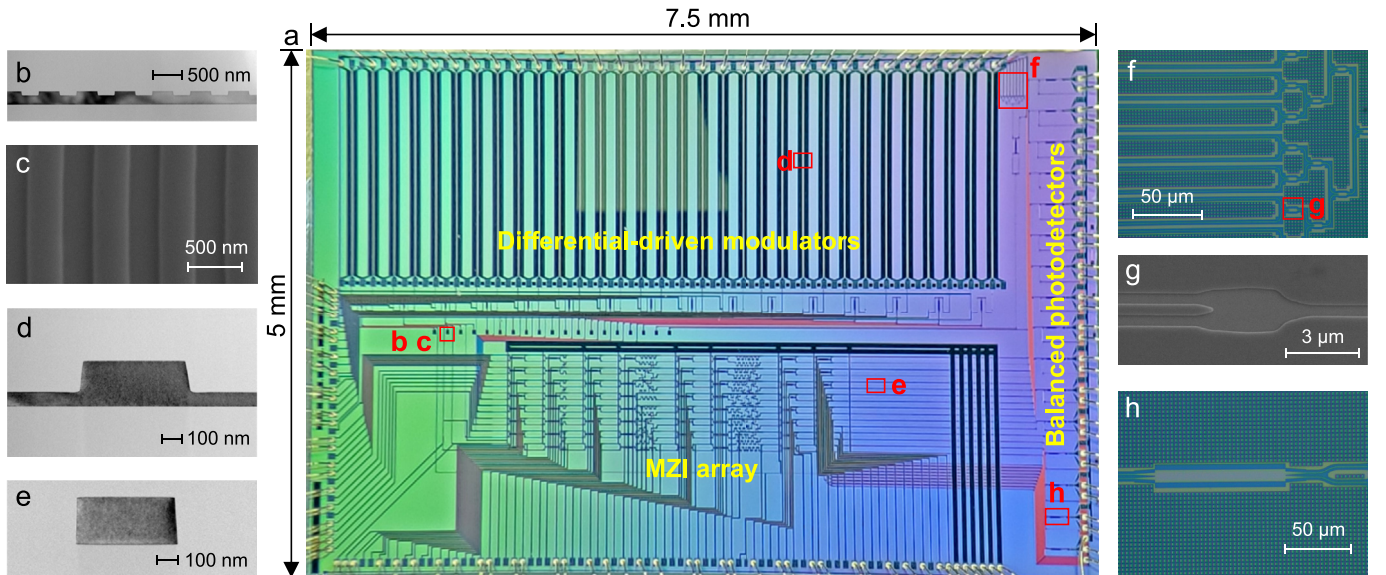


Fig. 2 | The image of the photonic chip. **a**, The microscope image of the photonic chip. **b**, **c**, SEM images of a coupling grating from the cross-section view and the top view respectively. **d**, **e**, The SEM images of the waveguides from the cross-section view. **f**, The microscope image of the beamsplitter. **g**, The SEM images of a 1×2 MMI coupler in **f**. **h**, The microscope image of a mixer. The positions of the components are denoted by the red boxes. Abbreviations: MMI, multimode interference; SEM, scanning electron microscope.

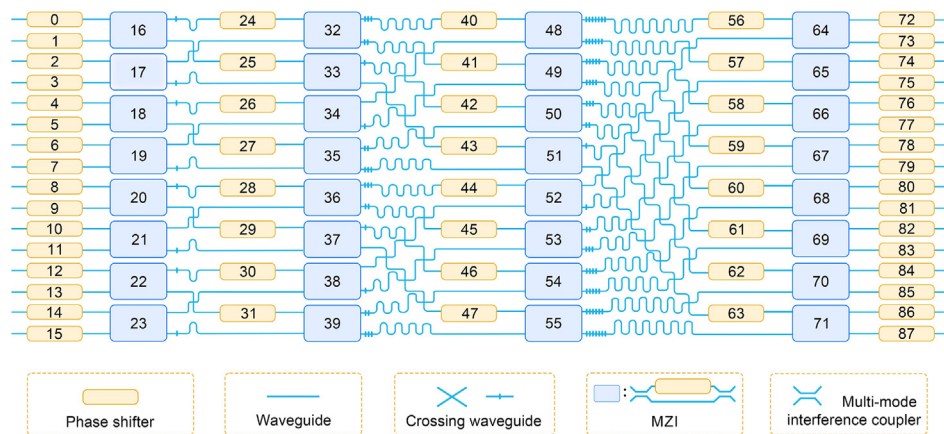


Fig. 3 | The structure of the FFT-mesh MZI array. Each yellow box denotes a thermo-optic phase shifter. Each blue box represents an MZI consisting of two multimode interference couplers and one phase shifter. Abbreviations: FFT, fast-Fourier-transform-mesh; MZI, Mach-Zehnder interferometer.

$$C_{\text{exp}} = -\frac{1}{2} \sum_{i=1}^N \left(I_{\text{BPD}}^{(i)} \right)^2. \quad (7)$$

EXPERIMENTAL DEMONSTRATION

In the experiment, we generated two configurations of the bias voltages V_1 and V_2 (the 32 voltages on the MZIs in Fig. 3 are randomly generated while the others are 0), and then employed them on the photonic chip respectively. Thus, two transformation matrices A_1 and A_2 can be obtained by measuring the output matrices when the input is the identity matrix. The transformation matrices A_1 and A_2 correspond to the QUBO problems with weight matrices K_1 and K_2 , respectively. The two problems are denoted by Q1 and Q2 for simplicity, and their weight graphs are shown in the insets of Fig. 4a and b, respectively, where the black dots denote the variables and the color bar denotes the weight.

Each problem is solved for 100 times, and the accepted states in all iterations are recorded. To evaluate the solving performance, the cost functions C of the accepted states are calculated with Eq. (1), and the evolution curves of C of Q1 and Q2 are plotted in Fig. 4a and b, respectively. It can be seen that C fluctuates above C_{\min} denoted by the black dashed line (C_{\min} is the lowest C obtained in numerical simulations) within a small range after about 400 iterations. To quantify the solving performance, a run is considered to be successful if the theoretical cost function of the final accepted state is lower than ηC_{\min} , where $\eta \in (0, 1]$ is the tolerance coefficient. Then the successful probability is defined as the proportion of the successful runs. The successful probabilities of Q1 and Q2 under $\eta = 0.960$ to $\eta = 0.995$ are calculated, and the results are shown in Fig. 4c and d, respectively. It can be seen from Fig. 4c and d that the successful probabilities are close to 1 when $\eta \leq 0.98$, while they tend to decrease quickly when $\eta > 0.98$. Such results indicate that our photonic solver is capable of solving QUBO problems. The corresponding simulation results are provided in Fig. 4c and d for comparison, respectively, which show that the successful probabilities in simulation are close to 1 even when $\eta = 0.995$. The decay in the experiment is mainly attributed to the fluctuation of the experimental cost function, which may result from the detection noise of the photodetector, the deviation of the OVMM, and the fluctuation of the laser power. They will be discussed in the next section.

RESULTS

Accuracy of the OVMM We first discuss the stability of the OVMM, which is directly associated with the accuracy of the experimental cost function. Due to the high sensitivity to the temperature of the photonic chip, even a small temperature variation would result in a completely different optical matrix transformation. Therefore, a cooling module is employed to stabilize the working temperature of the photonic chip at 37 °C. In the experiment, the temperature fluctuation is less than 0.005 °C. To quantify the stability of the OVMM, the fidelity F of each output vector is calculated by the following equation:

$$F = \frac{\left| \sum_i I_{\text{BPD}}^{(i)} I_T^{(i)} \right|}{\sqrt{\sum_i \left[I_{\text{BPD}}^{(i)} \right]^2 \sum_i \left[I_T^{(i)} \right]^2}}, F \in [0, 1], \quad (8)$$

where $I_T = \begin{bmatrix} I_T^{(1)} \\ \vdots \\ I_T^{(N)} \end{bmatrix} = A s$ is the theoretical output vector. The fidelity indicates the parallelism of the vectors I_{BPD} and I_T . The fidelity close to 1 can indicate the high stability of the OVMM. The average fidelity and the corresponding standard deviation of each run for Q1 and Q2 are shown in Fig. 4e. It can be seen that all the average fidelities are higher than 0.99 with small distribution, and the total average fidelities of Q1 and Q2 are 0.9953 ± 0.0017 and 0.9934 ± 0.0025 , respectively. Such results indicate that the OVMM of the photonic chip is quite stable during the solving process.

Besides the fidelity, the noise of the BPD and the fluctuation of the laser power can also degrade the accuracy of the OVMM. If the total noise level of the cost function is larger than the variation of the cost function, the photonic solver would give the incorrect solution. Therefore, the scale factor of the experimental cost function is used to evaluate the total noise level. The scale factor P in each iteration is defined as $P = C_{\text{exp}}/C_{\text{theo}} = |I_{\text{BPD}}|^2/|I_T|^2$, where $C_{\text{theo}} = -|I_T|^2/2$ is the theoretical cost function. In ideal conditions, P should be a constant. In the experiment, the fluctuation of P mainly results from the laser power fluctuation and the detection noise. When C_{theo} are equal, we have

$$\text{SNR} = \frac{C_{\text{exp}}}{\Delta C_{\text{exp}}} = \frac{P}{\Delta P}, \quad (9)$$

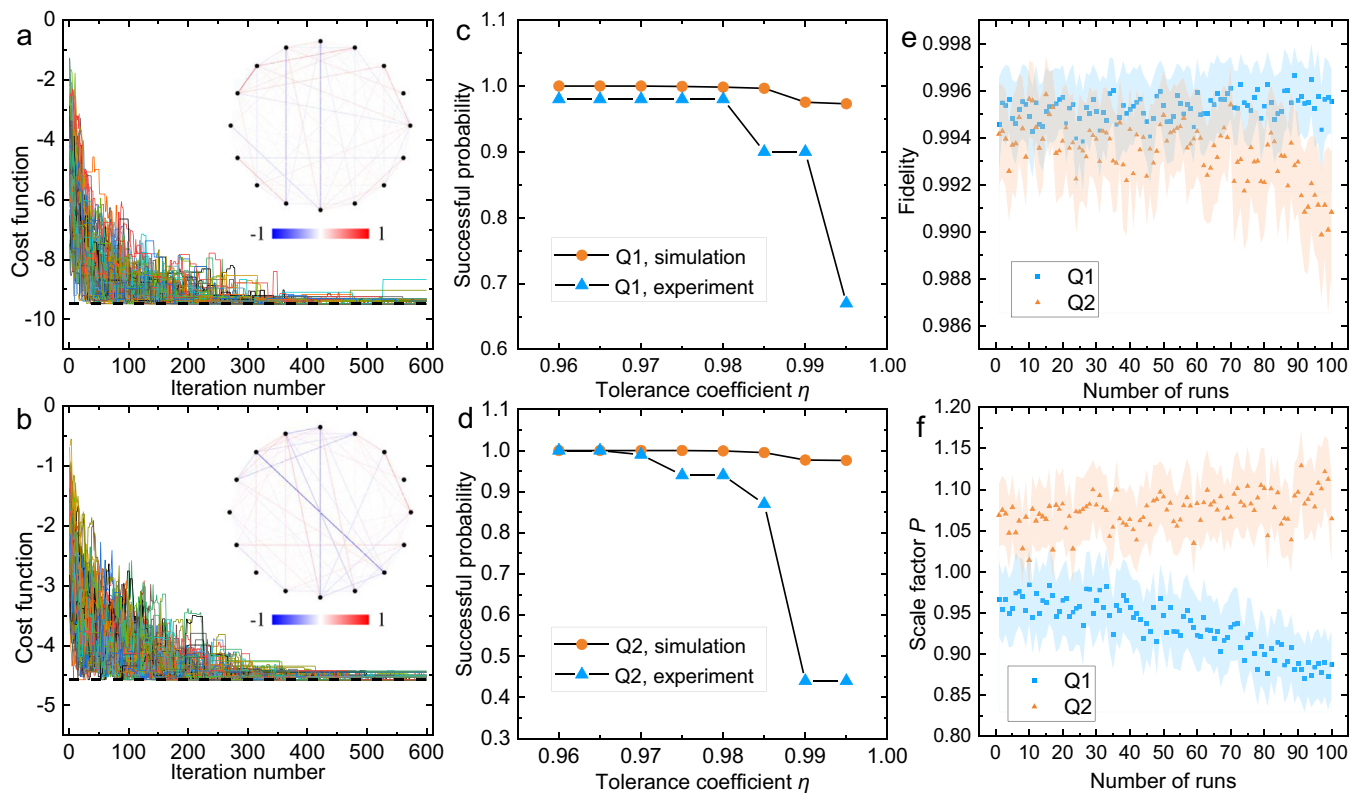


Fig. 4 | The experimental results. **a, b**, The normalized experimental Hamiltonian evolution curves of Q1 and Q2, respectively. The insets show the corresponding weight graphs, where the weight can be inferred from the color bar. **c, d**, The simulation and the experimental results of the final successful probabilities under different tolerance coefficients of Q1 and Q2, respectively. **e, f**, The average fidelity and scale factor in each run of Q1 and Q2, respectively. The blue and red regions denote the standard deviations, respectively.

where SNR denotes the signal-to-noise ratio of the experimental cost function, ΔC_{exp} is the noise level of C_{exp} , and ΔP is the standard deviation of the scale factor P . For convenience, the average scale factor and the corresponding standard deviation of each run are calculated in all iterations, and the results of Q1 and Q2 are shown in Fig. 4f. It can be seen that in each run the scale factor P only distributes in a small range. The average SNRs of 100 runs of Q1 and Q2 are 26.6 dB and 28.2 dB, respectively, which indicates that the noise level of the experimental cost function is small. The inverse of the SNR can also be regarded as the “resolution” of the cost function $R = 1/\text{SNR}$, hence the corresponding R of Q1 and Q2 are 4.67% and 3.88%, respectively. Such a parameter indicates that the photonic solver cannot distinguish two states if $R > |\Delta C_{\text{theo}}/C_{\text{min}}|$, where ΔC_{theo} is the difference of their theoretical cost functions and C_{min} is the minimum of the cost function.

To further investigate how much the noise level would affect the searching process near the ground state, the theoretical relative variation of the cost function, $C_r = |\Delta C_{\text{theo}}/C_{\text{min}}|$, between the sampled state and the previous state of all samplings are calculated in iterations 400–600. Obviously, if $R > C_r$ and $\Delta C_{\text{theo}} > 0$ in a single iteration, though the theoretical cost function increases, which means that the sampled state should not be accepted, the experimentally measured cost function might decrease due to the noise. Therefore, a wrong acceptance of the sampled state would occur. Among these samplings, it is found that the proportions of the samples that satisfy $R > C_r$ and $\Delta C_{\text{theo}} > 0$ are 29.9% and 26.8% for models 1 and 2, respectively. This indicates that our photonic

solver can distinguish two states in most iterations of the demonstrated problems, hence high successful probabilities can be obtained even when the tolerance coefficient is close to 1. The above analysis indicates that our OVMM is quite stable and accurate during the solving process.

Computing speed and power consumption As mentioned in the previous section, the calculation of the photonic solver includes the pretreatment and the iterative sampling of the state vector. In the pretreatment, the eigendecomposition of the weight matrix of the given problem is calculated to obtain the necessary transformation matrix of the OVMM. In the experimental demonstration, the transformation matrices of the OVMM are randomly generated; hence, we only discuss the time consumption in the iterative sampling process.

The main sequence diagram of each sampling iteration is shown in Fig. 5a. The clock rate of the FPGA is 245.76 MHz, which is corresponding to the clock cycle of $t_0 = 4.069$ ns. At $t = 0$, the FPGA sends the TX (transmit) signal containing the voltages on the amplitude modulators to the DAC. It is observed that the RX (receive) signal appears at $t = 40t_0$. Such latency is composed of five parts in sequence: the DAC latency, the modulator response time, the propagation time of lightwaves in the OVMM, the photodetector response time, and the ADC latency. According to the frequency responses of the modulator and the photodetector shown in Fig. 5b, the -3 dB bandwidth f_B of the amplitude modulator is 28.0 GHz and that of the photodetector at -2 V bias voltage is 41.3 GHz. Therefore, the response time of the modulator and the photodetector can be estimated

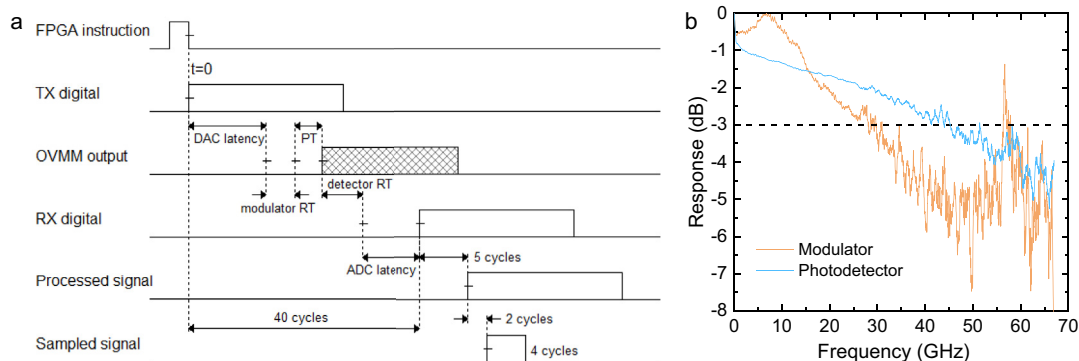


Fig. 5 | Analysis of time consumption. **a**, Main sequence diagram of each sampling iteration. **b**, The frequency responses of the modulators and the photodetectors. Abbreviations: PT, propagation time; RT, response time.

to be $\tau_m = 12.5$ ps and $\tau_p = 8.5$ ps, respectively, according to $\tau \approx 0.35/f_B$ ³⁹. The length from the amplitude modulator to the photodetector is 9.3 mm, and the group index of silicon under the working temperature of 37 °C is 4.28⁴⁰, hence the propagation time of light is 132.7 ps. Thus, the computing of the OVMM takes $\tau_{OVMM} \approx 153.7$ ps in each iteration, and the total latency of the DAC, ADC, and signal synchronization can be estimated as $\tau_{DAC/ADC} \approx 162.6$ ns, which indicates that τ_{OVMM} is much lower than $\tau_{DAC/ADC}$. After the signal processing operations, the processed RX signal appears at $t = 45t_0$, and the sampled signal rises at $t = 47t_0$ and falls at $t = 51t_0$. According to the experimental results, each iteration takes $\tau_{iter} \approx 265.1$ ns. Thus, the time consumption of the heuristic algorithm on the FPGA takes $\tau_{FPGA} = \tau_{iter} - 51t_0 = 57.6$ ns.

In each iteration, calculating the cost function requires the multiplication between an $N \times N$ real matrix and an $N \times 1$ vector, which includes N^2 floating-point operations (FLOPs). In the experimental demonstration, the computing speed of the OVMM on chip is 1.57 GFLOP/s, which is mainly limited by the latency of the digital-analog conversion. As mentioned above, the computing time of the OVMM is much lower than the total latency of the DAC, ADC, and signal synchronization, which is the actual bottleneck of the computation speed of the photonic solver. Adopting the low-latency DAC, ADC, and the FPGA with a higher operating frequency would reduce the iteration time, but the improvement would be very limited in the near future. An alternative approach is to utilize all-analog and hybrid optoelectronic architecture. Recently, the all-analog chip combining electronic and light computing (ACCEL) for vision tasks has been reported, in which the conversion between digital and analog signals are not involved, and the superiority of the photonic chip is fully exploited⁴¹. Likewise, if the latency of the digital-analog conversion can be overcome, the maximum computing speed of our photonic chip can be estimated as $N^2/\tau_{OVMM} \approx 1.66$ TFLOP/s, and the corresponding area efficiency of the photonic chip is 44.4 GFLOP/mm² with the chip area of 37.5 mm². Another approach is to operate the OVMM with the pipelining method, where the data input and output are uninterrupted. Consequently, the computing speed only depends on the bandwidth of the modulator and the detector, rather than the propagation time of lightwave and the latency from the digital-analog conversion, and the computing speed would achieve $N^2/\max(\tau_m, \tau_p) \approx 20.5$ TFLOP/s⁴². In the future work, more compact chip sizes and more channels of the OVMM based on advanced fabrications would benefit both the computing speed and the area efficiency of the photonic chip.

The power consumption of the presented photonic chip can be estimated as follows. When the applied voltages on all phase shifters are 7 V and 0 V in the MZI array, the power consumption of the driving board is $p_1 = 46.67$ W and $p_2 = 45.95$ W, respectively. Therefore, the power consumption of the photonic chip is estimated to be $p_1 - p_2 = 0.72$ W, which is proportional to the number of the thermo-optic phase shifters, $N \log_2 N + 3N/2$. The energy efficiency of the chip is 0.458 nJ/FLOP in the current experimental demonstration (corresponding to the computing speed of 1.57 GFLOP/s), if the latency of the digital-analog conversion can be overcome like that in the all-analog architecture, the energy efficiency could be as high as 434 fJ/FLOP, while only considering the maximum computing speed of the photonic chip (1.66 TFLOP/s). Therefore, it is believed that the photonic chip is capable of achieving higher computing speed and energy efficiency.

CONCLUSIONS

We have proposed and demonstrated a 16-channel photonic solver based on an integrated photonic chip for QUBO problems. The amplitude modulators, the MZI array for the optical matrix transformation, and the photodetectors were all integrated on the same photonic chip to achieve low optical latency. The calculation of the cost function can be accelerated with the photonic chip, while heuristic algorithms are employed on the FPGA to search for the optimal solution. Two randomly generated 16-dimensional QUBO problems have been successfully solved with the photonic solver, and the successful probabilities of both problems are larger than 0.94 with the tolerance coefficient of 0.98. Such results indicate the high stability and low noise level of the photonic solver. According to the experimental results, each iteration takes 265.1 ns, which is mainly ascribed to the DAC and ADC latency (162.6 ns). The maximum computing speed and the area efficiency of the photonic chip are 1.66 TFLOP/s and 44.4 GFLOP/mm², respectively. The energy efficiency is 0.458 nJ/FLOP in the demonstration. Our proposed photonic solver shows the possibility of the integrated photonic system to accelerate the solving process of computationally complex problems.

REFERENCES

1. Lin, X. et al. All-optical machine learning using diffractive deep neural networks. *Science* **361**, 1004–1008 (2018). <https://doi.org/10.1126/science.aat8084>.

2. Zhou, T. et al. Large-scale neuromorphic optoelectronic computing with a reconfigurable diffractive processing unit. *Nat. Photonics* **15**, 367–373 (2021). <https://doi.org/10.1038/s41566-021-00796-w>.
3. Gao, S. et al. Super-resolution diffractive neural network for all-optical direction of arrival estimation beyond diffraction limits. *Light: Sci. Appl.* **13**, 161 (2024). <https://doi.org/10.1038/s41377-024-01511-4>.
4. Shen, Y. et al. Deep learning with coherent nanophotonic circuits. *Nat. Photonics* **11**, 441–446 (2017). <https://doi.org/10.1038/nphoton.2017.93>.
5. Feldmann, J. et al. Parallel convolutional processing using an integrated photonic tensor core. *Nature* **589**, 52–58 (2021). <https://doi.org/10.1038/s41586-020-03070-1>.
6. Ashtiani, F., Geers, A. J. & Aflatouni, F. An on-chip photonic deep neural network for image classification. *Nature* **606**, 501–506 (2022). <https://doi.org/10.1038/s41586-022-04714-0>.
7. Yan, T. et al. All-optical graph representation learning using integrated diffractive photonic computing units. *Sci. Adv.* **8**, eabn7630 (2022). <https://doi.org/10.1126/sciadv.abn7630>.
8. Fu, T. et al. Photonic machine learning with on-chip diffractive optics. *Nat. Commun.* **14**, 70 (2023). <https://doi.org/10.1038/s41467-022-35772-7>.
9. Xu, Z. et al. Large-scale photonic chiplet Taichi empowers 160-TOPS/W artificial general intelligence. *Science* **384**, 202–209 (2024). <https://doi.org/10.1126/science.adl1203>.
10. Nguyen, P. H. & Boninsegni, M. Superfluid transition and specific heat of the 2D x-y model: Monte Carlo simulation. *Appl. Sci.* **11**, 4931 (2021). <https://doi.org/10.3390/app11114931>.
11. Gilli, M., Maringer, D. & Schumann, E. *Numerical Methods and Optimization in Finance*. (Elsevier Inc., 2019). <https://doi.org/10.1016/C2017-0-01621-X>.
12. Bryngelson, J. D. & Wolynes, P. G. Spin glasses and the statistical mechanics of protein folding. *Proc. Natl. Acad. Sci. U.S.A.* **84**, 7524–7528 (1987). <https://doi.org/10.1073/pnas.84.21.7524>.
13. Degasperis, A., Fey, D. & Kholodenko, B. N. Performance of objective functions and optimisation procedures for parameter estimation in system biology models. *npj Syst. Biol. Appl.* **3**, 1–9 (2017). <https://doi.org/10.1038/s41540-017-0023-2>.
14. Xu, X.-Y. et al. A scalable photonic computer solving the subset sum problem. *Sci. Adv.* **6**, eaay5853 (2020). <https://doi.org/10.1126/sciadv.aay5853>.
15. Jiang, X. et al. Programmable photonic solver for computationally complex problems. *ACS Photonics* **10**, 4340–4348 (2023). <https://doi.org/10.1021/acsp Photonics.3c01164>.
16. Mohseni, N., McMahon, P. L. & Byrnes, T. Ising machines as hardware solvers of combinatorial optimization problems. *Nat. Rev. Phys.* **4**, 363–379 (2022). <https://doi.org/10.1038/s42254-022-00440-8>.
17. Yajima, S. & Shoji, Y. Integrated photonic traveling salesman problem probabilistic solver with polynomial calculation time complexity. *Jpn. J. Appl. Phys.* **61**, 052003 (2022). <https://doi.org/10.35848/1347-4065/ac55e2>.
18. 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). <https://doi.org/10.1103/PhysRevA.88.063853>.
19. 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). <https://doi.org/10.1038/nphoton.2014.249>.
20. McMahon, P. L. et al. A fully programmable 100-spin coherent Ising machine with all-to-all connections. *Science* **354**, 614–617 (2016). <https://doi.org/10.1126/science.aah5178>.
21. Inagaki, T. et al. A coherent Ising machine for 2000-node optimization problems. *Science* **354**, 603–606 (2016). <https://doi.org/10.1126/science.aah4243>.
22. Honjo, T. et al. 100,000-spin coherent Ising machine. *Sci. Adv.* **7**, eabh0952 (2021). <https://doi.org/10.1126/sciadv.abh0952>.
23. Cen, Q. et al. Large-scale coherent Ising machine based on optoelectronic parametric oscillator. *Light: Sci. Appl.* **11**, 333 (2022). <https://doi.org/10.1038/s41377-022-01013-1>.
24. Takeda, Y. et al. Boltzmann sampling for an XY model using a non-degenerate optical parametric oscillator network. *Quantum Sci. Technol.* **3**, 014004 (2017). <https://doi.org/10.1088/2058-9565/aa923b>.
25. Berloff, N. G. et al. Realizing the classical XY Hamiltonian in polariton simulators. *Nat. Mater.* **16**, 1120–1126 (2017). <https://doi.org/10.1038/nmat4971>.
26. Gershenzon, I. et al. Exact mapping between a laser network loss rate and the classical XY Hamiltonian by laser loss control. *Nanophotonics* **9**, 4117–4126 (2020). <https://doi.org/10.1515/nanoph-2020-0137>.
27. Pierangeli, D., Marcucci, G. & Conti, C. Large-scale photonic Ising machine by spatial light modulation. *Phys. Rev. Lett.* **122**, 213902 (2019). <https://doi.org/10.1103/PhysRevLett.122.213902>.
28. Pierangeli, D., Marcucci, G. & Conti, C. Adiabatic evolution on a spatial-photonic Ising machine. *Optica* **7**, 1535–1543 (2020). <https://doi.org/10.1364/OPTICA.398000>.
29. Pierangeli, D., Marcucci, G., Brunner, D. & Conti, C. Noise-enhanced spatial-photonic Ising machine. *Nanophotonics* **9**, 4109–4116 (2020). <https://doi.org/10.1515/nanoph-2020-0119>.
30. Roques-Carnes, C. et al. Heuristic recurrent algorithms for photonic Ising machines. *Nat. Commun.* **11**, 1–8 (2020). <https://doi.org/10.1038/s41467-019-14096-z>.
31. Prabhu, M. et al. Accelerating recurrent Ising machines in photonic integrated circuits. *Optica* **7**, 551–558 (2020). <https://doi.org/10.1364/OPTICA.386613>.
32. Pierangeli, D., Rafayelyan, M., Conti, C. & Gigan, S. Scalable spin-glass optical simulator. *Phys. Rev. Appl.* **15**, 034087 (2021). <https://doi.org/10.1103/PhysRevApplied.15.034087>.
33. Yamashita, H. et al. Low-rank combinatorial optimization and statistical learning by spatial photonic Ising machine. *Phys. Rev. Lett.* **131**, 063801 (2023). <https://doi.org/10.1103/PhysRevLett.131.063801>.
34. Ouyang, J. et al. On-demand photonic Ising machine with simplified Hamiltonian calculation by phase encoding and intensity detection. *Commun. Phys.* **7**, 1–9 (2024). <https://doi.org/10.1038/s42005-024-01658-x>.
35. Ouyang, J. et al. Programmable and reconfigurable photonic simulator for classical XY models. *Phys. Rev. Appl.* **22**, L021001 (2024). <https://doi.org/10.1103/PhysRevApplied.22.L021001>.
36. Lucas, A. Ising formulations of many NP problems. *Front. Phys.* **2**, 5 (2014). <https://doi.org/10.3389/fphy.2014.00005>.
37. Kochenberger, G. et al. The unconstrained binary quadratic programming problem: a survey. *J. Combin. Optim.* **28**, 58–81 (2014). <https://doi.org/10.1007/s10878-014-9734-0>.
38. Tian, Y. et al. Scalable and compact photonic neural chip with low learning-capability-loss. *Nanophotonics* **11**, 329–344 (2022). <https://doi.org/10.1515/nanoph-2021-0521>.
39. Li, Z.-Y. et al. Silicon waveguide modulator based on carrier depletion in periodically interleaved PN junctions. *Opt. Express* **17**, 15947–15958 (2009). <https://doi.org/10.1364/OE.17.015947>.
40. Eric, D. et al. Group index and group velocity dispersion in silicon-on-insulator photonic wires. *Opt. Express* **14**, 3853–3863 (2006). <https://doi.org/10.1364/OE.14.003853>.
41. Chen, Y. et al. All-analog photoelectronic chip for high-speed vision tasks. *Nature* **623**, 48–57 (2023). <https://doi.org/10.1038/s41586-023-06558-8>.
42. Li, S. et al. All-optical image identification with programmable matrix transformation. *Opt. Express* **29**, 26474–26485 (2021). <https://doi.org/10.1364/OE.430281>.

MISCELLANEA

Supplementary materials Supplementary data to this article can be found online at <https://doi.org/10.1016/j.chip.2024.100117>.

Acknowledgments Funding from the National Key Research and Development Program of China (2023YFB2806703), the National Natural Science Foundation of China (Grant No. U22A6004, 92365210) is greatly acknowledged. This work was also supported by Beijing National Research Center For Information Science And

Technology, Frontier Science Center for Quantum Information, Beijing Academy of Quantum Information Sciences, Tsinghua University Initiative Scientific Research Program. The authors would like to thank Mr. Hejia Zhang for his helpful comments.

Author contributions Jiayi Ouyang: Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. Shengping Liu: Writing – original draft, Software, Resources, Project administration, Methodology, Investigation, Conceptualization. Ziyue Yang: Validation, Software,

Data curation. **Wei Wang:** Validation, Software, Investigation, Data curation. **Xue Feng:** Writing – review & editing, Supervision, Resources, Project administration, Funding acquisition, Conceptualization. **Yongzhuo Li:** Writing – review & editing, Supervision. **Yidong Huang:** Writing – review & editing, Supervision.

Declaration of Competing Interest The authors declare no competing interests.

© 2024 The Author(s). Published by Elsevier B.V. on behalf of Shanghai Jiao Tong University. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).