REVIEW **Open Access**

The challenges of modern computing and new opportunities for optics

Chong Li, Xiang Zhang, Jingwei Li, Tao Fang and Xiaowen Dong*

* Correspondence: xiaowen.dong@huawei.com
Huawei Technologies Co., Ltd, Shenzhen, China

Abstract

In recent years, the explosive development of artificial intelligence implementing by artificial neural networks (ANNs) creates inconceivable demands for computing hardware. However, conventional computing hardware based on electronic transistor and von Neumann architecture cannot satisfy such an inconceivable demand due to the unsustainability of Moore's Law and the failure of Dennard's scaling rules. Fortunately, analog optical computing offers an alternative way to release unprecedented computational capability to accelerate varies computing drained tasks. In this article, the challenges of the modern computing technologies and potential solutions are briefly explained in Chapter 1. In Chapter 2, the latest research progresses of analog optical computing are separated into three directions: vector/matrix manipulation, reservoir computing and photonic Ising machine. Each direction has been explicitly summarized and discussed. The last chapter explains the prospects and the new challenges of analog optical computing.

Keywords: Optical computing, Vector matrix multiplier, Artificial neural network, Reservoir computing, Photonic Ising machine, Hybrid optical-electrical system

Introduction

The extraordinary development of complementary-metal-oxide-semiconductor (CMOS) technology facilitates an unprecedented success of integrated circuits. As predicated by Gordon E. Moors in 1965, the transistor number on a computing chip is doubled in every 18–24 months. Moreover, Dennard's scaling rule explains the benefit of reducing a transistor's dimensions in further [1]. Nowadays, Moore's law has made central processor units (CPUs) 300 times faster than that in 1990. However, such an incredible development is unsustainable as predicted by the International Technology Roadmap of Semiconductors (ITRS) in 2016. After 5 nm technology node, the semiconductor industry is difficult to move forward. In addition, the proliferation of artificial intelligence (AI) applications create exponentially increasing amounts of data that can hardly processed by conventional computing systems and architectures. Such a desperate discrepancy boosts numerous investigations of novel approaches and alternative architectures for data processing.

Comparing to electrical devices, optical devices can process information instantaneously with negligible energy consumption and heat generation. Furthermore, optical

devices have much better parallelism than electrical devices in data processing by employing multiplex schemes, such as wavelength division multiplexing (WDM) and mode division multiplexing (MDM). With adopting the properties of light, the architecture and layout of many complex computing systems can be potentially simplified by introducing optical computing units.

In general, optical computing can be classified in two different categories: the digital optical computing and the analog optical computing. The digital optical computing based on Boole logics, using similar mechanism as the general-purpose computing based on transistor, has been developed for more than 30 years. However, it is difficult to beat the conventional digital computing in terms of the low integration density of optical device. In contrast, analog optical computing utilizes the physical characteristics of light, such as amplitude and phase, and the interactions between light and optical devices to achieve certain computing functions. It is a dedicated computing because of the unique mathematical depiction of computational process in one certain analog optical computing system. Compared to the conventional digital computing, the analog optical computing can realize better data processing acceleration in specific tasks, such as pattern recognition and numerical computation. Therefore, as one of the most promising computing technologies in post-Moore era, large amount of research work has been drawn into the investigation of analog optical computing systems.

In this paper, the challenges of modern computing and the potential opportunities of analog optical computing have been discussed separately. The first chapter briefly explains the main factors impeding the sustainability of Moore's law, the growing demands of information processing, and the latest researches in the semiconductor industry. In the second chapter, the progresses of analog optical computing over last decade have been reviewed in three sections. In the last chapter, a systematical analysis of the hybrid computing system has been given followed by a discussion of the new challenges and potential opportunities of analog optical computing.

Moore's law and the new challenges

The challenges of Moore's law

Originally, Moore's law and Dennard's scaling rules show the reduction of transistor's dimensions is a viable way to boost computational capability without increasing energy dissipation. While, the continuous development CMOS technologies induces the failure of Dennard's scaling rules, because the shrunk transistor cannot maintain a constant energy density. Utilizing a higher clock frequency in CPUs would be another plausible way to further enhance computational capability. However, the thermal effects from power dissipation will become a new bottleneck of CPUs' performance by employing high clock frequency. Today, the computational capabilities of CPUs, with the 5 GHz clock speed constrains, are alternatively improved by utilizing a parallel architecture.

Apart from the thermal effects from power dissipation, the limitations of manufacturing process also challenge the Moore's law. To extend the downscaling of transistor in CPUs, the new top-down patterning methods should be introduced into current manufacturing line. Extreme ultraviolet (EUV) lithography, at the 13.5 nm wavelength, is the core technology to extend the Moore's law because of the shorter wavelength allows the higher resolution [2]. For EUV interference lithography, the theoretical limit of

half-pitch is around 3.5 nm. Similarly, electron beam lithography (EBL) as another fabrication technology, is also able to create the extremely fine patterns of integrated circuits with high resolution. Though EBL provides ultra-high resolution closing to the atomic level and adapts to work with a variety of materials, the processing time is much longer and more expensive than optical lithography [3].

These scale down methodologies for silicon-based CMOS circuits are classified as 'More Moore' technologies which are used to maintain the Moore's law. However, following the size reduction of transistor's gate channel by employed better fabrication technologies, the quantum effects, such as quantum tunneling and quantum scattering, will bring other unpredictable problems. For example, in the latest sub-5 nm all-around gate (GAA) of the fin field-effect transistor (FinFET), the threshold voltage is increased as the effective fin width reduced by quantum effect [4]. Therefore, the enhancement of computational capability will not be able to sustain by shrinking the transistor size continuously.

The challenges of AI applications

On top of the challenges from physical limitations of Moore's Law mentioned in the "The challenges of Moore's law" section, the computational capability of conventional digital systems is challenged by the thriving AI applications as well. The most popular AI implementations are deep neural networks (DNNs) which contain two most important types: convolution neural networks (CNNs) and long short-term memory (LSTM). In CNNs, there are a series of convolution and sub-sampling layers followed by a fully connected layer and a normalizing layer. Convolution is the main computing task for inference and back-propagation is used solely for training all parameters in CNN [5]. LSTM consists of blocks of memory cell which are dictated by the input, forget and output gates. The output of the LSTM blocks are calculated via the cell values [6–9]. To promote high accuracy of output results, DNNs have been developing large number of parameters. The first DNN model LeNet [10] only contains 5 convolution layers with 60 K parameters. In 2012, AlexNet [11] became the best performance DNN model with 60 M parameters. Nowadays, the Megatron model [6] contains 3.9 G parameters and it needs several weeks to train with millions level USD costing.

All the processes of DNNs mentioned above contains many complex computing tasks and it consume large volume of computing resource. A metric researched by OpenAI shows that the prosperity of AI has increased the demand of computational capability more than 300,000 times from 2012 to 2018, while Moore's law would yield only a 7 times enhancement [7]. In short, AI applications have become more and more complex, precise and computing resources drained. There is a great thirst for higher computational capability systems to meet these challenges.

New attempts under the challenges

It is clear that extending the Moore's law is one critical factor to gain the computational capability. To promote the semiconductor technologies, there are two other technical paths 'More than Moore' and 'Beyond CMOS', apart from 'More Moore' [12]. 'More than Moore' encompasses the engineering of complex heterogeneous systems that can meet certain needs and advanced applications, with varies technologies (such

as system on chip, system in package, network on chip et al.). 'Beyond CMOS' explores the new materials to improve the performance of CMOS transistor, such as carbon nanotubes (CNT) [13]. The motivation of introducing CNT in computing system is that the CNT based transistors have low operation voltages and exceptional performance as they have shorter length of current-carrying channel than current design. Because CNT can be either metallic or semiconducting, the isolation of purely semiconducting nanotubes is essential for making high performance transistors. However, the purifying and controllably positioning for these 1 nm diameter molecular cylinders is still a formidable challenge today [14–17].

Besides extending the Moore's law, developing new systematic architectures can also gain the computational capability of conventional digital systems. In-memory computing architecture has been extensively explored in CMOS based static random access memory (SRAM) [18, 19]. However, CMOS memories have limitation in density which is slow in scaling trends. Researchers are motivated to explore in-memory computing architectures with the emerging non-volatile memory (NVM) technologies, such as phase change material (PCM) [20] and resistive random-access memory (RRAM) [21]. NVM devices are configured in a form of two-dimensional crossbar array which enables high performance computing as NVM devices allow non-volatile multiple states. NVM crossbars can do multiplication operation in parallel and result higher energy efficiency and speed than conventional digital accelerators by eliminating data transfer [18]. The high density NVM crossbars provide massively parallel multiplication operations and lead to the exploration of analog in-memory computing systems [19].

However, the approaches mentioned above still seem to be incompetent at meeting the challenges which are from the applications with extreme computational complexity, such as large scale optimization, large molecules simulation, large number decomposition, etc. These applications require large size of memory which the most powerful supercomputers can hardly meet. In addition, processing of these applications needs the runtimes on the order of tens of years or more. Therefore, it is essential to investigate the new computing paradigms which are different with the conventional computing systems based on Boole logics and von Neumann architecture. Currently, quantum computing, DNA computing, neuromorphic computing, optical computing, etc. called as physical computing paradigms are attracting more and more researcher attention. These physical computing paradigms, providing more complexity operators than Boole logics in device level, can be used to build exceptional accelerators. Compared to the low-temperature requirement in quantum computing, and the dynamic instabilities of DNA and neuromorphic computing, optical computing has loose environment requirement and solid systemic composing. Therefore, optical computing has been considered as one of the most promising ways to tackle intractable problems.

Analog optical computing: an alternative approach at post-Moore era

Optical computing is not a brand-new concept. Back to the middle of twentieth century, the optical correlator had already been invented [22], and it can be treated as an preliminary prototype of optical computing system. Other technologies underpinned by the principles of Fourier optics, such as 4F-system and vector matrix multiplier (VMM), were well developed and investigated during last century [22–25]. The great success of digital electrical computer promotes the investigations of digital optical

computer in which the optical logic gates have been concatenated [26–33]. The idea of replacing electrical transistor by optical transistor was considered as a competitive approach to build a digital optical computer due to the intrinsic merits of photon, such as high bandwidth, negligible heat generation and ultra-fast response. However, this tantalizing idea has not yet been systematically verified since the middle of twentieth century. D. B. Miller proposed some practical criteria for optical logic in 2010, and he pointed out that current technologies were incompetence to meet these criteria. These criteria include logic-level restoration, cascadability, fan-out, input–output isolation, absence of critical biasing and independent loss at logic level [34]. Until now, a digital optical computer is still a fascinate blueprint. Digital electrical computer still is a practical and reliable system due to its compatibility and flexibility. Alternatively, analog optical computing harnessing physical mechanisms opens up new possibilities for optical computing because it relieves the requirement of high integration density by implementing arithmetic operation rather than Boole logic operation. In this chapter, VMM, reservoir computing and photonic Ising machine are illustrated as three typical instances of analog optical computing. "Vector and matrix manipulation in optical domain" section explains the principle of VMM and its applications toward complex computing. "Optical reservoir computing" section and "Photonic Ising Machine" section summarize the principle and research progresses of reservoir computing and photonic Ising machine, respectively.

Vector and matrix manipulation in optical domain

Since optical computing has not yet been verified as a viable approach to realize universal computing via logical operations, people start to explore the potential opportunities in arithmetic computing, such as multiplication and addition. In this section, the relevant researches are briefly summarized and sequentially explained. Firstly, a principle explanation of multiplication is followed by a typical realization called fan-in/out VMM introduced by Goodman [24] in last century. Many creative schemes and new technologies are introduced as well. Then complex computing is introduced, such as Fourier transformation (FT) and convolution. A typical way of realizing FT and convolution are explicitly explained. At last, other optical computing schemes are mentioned as well.
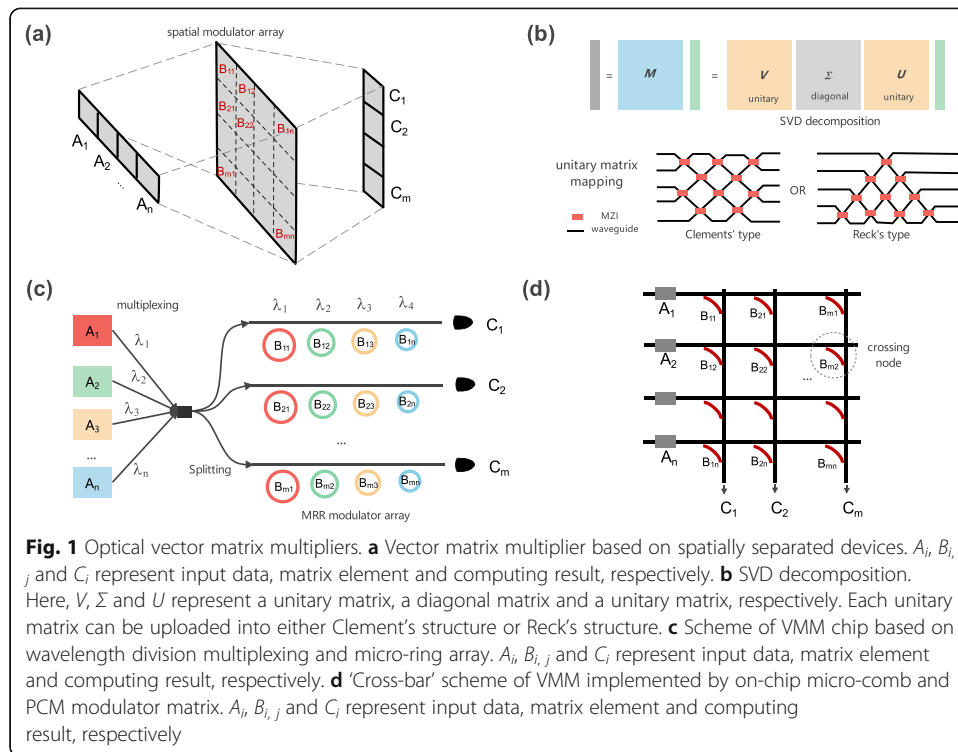
VMM-vector matrix multiplier

As mentioned above, the first fan-in/out VMM was proposed as early as 1978 [24]. This multiplier is designed to compute multiplication between a vector and a matrix as follows

$$C_j = \sum_i B_{ji} \cdot A_i, \tag{1}$$

where A and B are a vector and matrix, respectively. The jth-row of the matrix B times with the vector A in an element-wise way, and a scalar result C_j is obtained after summation. After traversing each row of matrix B, the final result of the VMM is obtained.

The traditional free-space fan-in/out VMM scheme shows in Fig. 1(a). The input vector A and matrix B are loaded into an array of light sources and a series planar spatial light modulators (SLM), respectively. One or several lenses are used to expand each light beam from a A_i source to illuminate all the pixels at i-th column of SLM. Then, a

Fig. 1 Optical vector matrix multipliers. **a** Vector matrix multiplier based on spatially separated devices. A_i, $B_{i, j}$ and C_i represent input data, matrix element and computing result, respectively. **b** SVD decomposition. Here, V, Σ and U represent a unitary matrix, a diagonal matrix and a unitary matrix, respectively. Each unitary matrix can be uploaded into either Clement's structure or Reck's structure. **c** Scheme of VMM chip based on wavelength division multiplexing and micro-ring array. A_i, $B_{i, j}$ and C_i represent input data, matrix element and computing result, respectively. **d** 'Cross-bar' scheme of VMM implemented by on-chip micro-comb and PCM modulator matrix. A_i, $B_{i, j}$ and C_i represent input data, matrix element and computing result, respectively

cylinder lens (other collimating lenses may be used to improve the precision) is used to focus all the beams in the horizontal direction, and a line array of spots can be detected at last. Theoretically, the intensity of spots is proportional to the computing result C. In this scheme, the lenses before SLM are used to broadcast the vector A and map it onto each row of SLM, and the SLM is respond for element-wise production. The lenses after SLM are used to do the summation. Assuming the vector has a length of N and the matrix size is $N * N$, this architecture can effectively achieve $\sim N^2 MAC$ in 'one flash' if all the data has been loaded (MAC, multiply–accumulate operation, each contains one multiplication and one adding). Although the light propagates very fast, the loading time of data and the detecting time of optical signal cannot be ignored. Thereby, the effective peak performance of such apparatus is $\sim F \cdot N^2 MAC$/s. The F is the working frequency of the system, which is mainly limited by the refreshing rate of the SLM. An impressive engineering practice is Enlight256 developed by Israeli company Lenslet at 2003. It supports the multiplication between a 256-length vector and a matrix with the size of 256*256 at 125 MHz refreshing rate. In other words, its computational capability can reach ~ 8 *TMAC*/s, and it is faster than the digital signal processor (DSP) at that time by 2–3 orders [35]. The key technology of Enlight256 is the high speed gallium arsenide (GaAs) based SLM which is different with the traditional ones with 10^{0-1} ms typical response time based on liquid crystal.

Moreover, benefiting from the quickly developed liquid-crystal-on-silicon (LCoS) technology and driving from the display industry, the resolution of SLM or DMD becomes fairly large (4 K resolution is commercially available). But the crosstalk error is the main obstacle to demonstrate the utmost performance of VMM employing high resolution SLM or DMD [36]. Though the crosstalk issue could be circumvented by enlarging the pixel size of SLM or DMD, the functional area of SLM or DMD restricts

the size of matrix. Meanwhile, the diffraction of light cannot be ignored even if using incoherent light source. This limitation is named as space–bandwidth product similar to the time-bandwidth product in the traditional communication system.
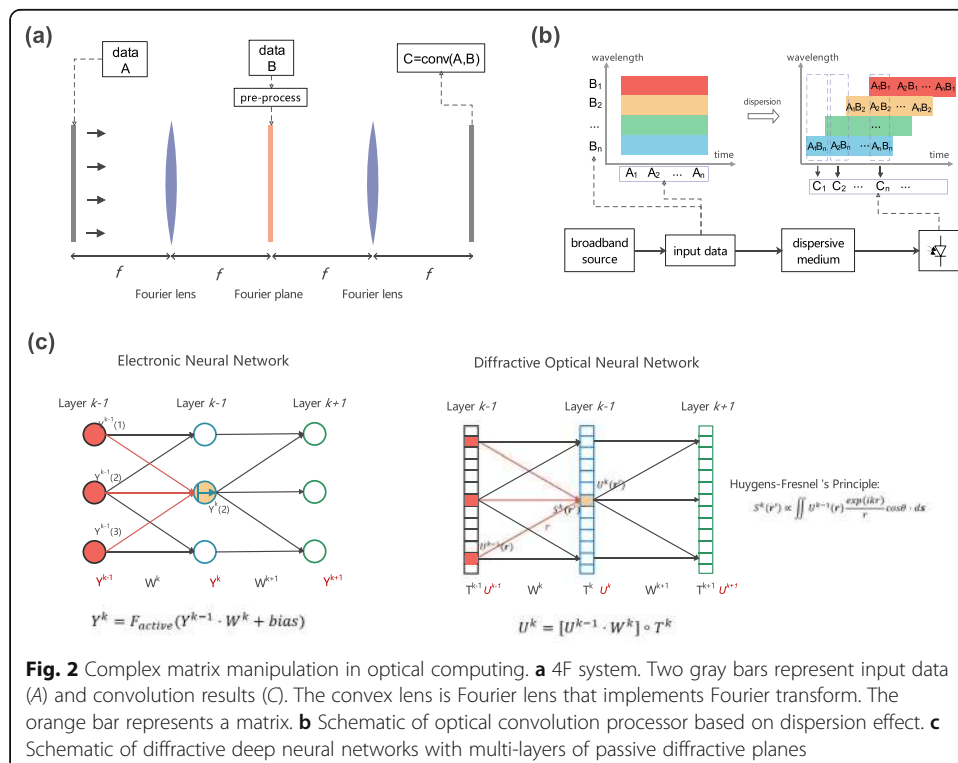
In recent years, many creative works have been proposed and demonstrated in waveguide rather than using traditional free-space VMM scheme. D. B. Miller [37] has proposed a method to efficiently design an optical component for universal linear operation, which can be implemented by Mach-Zehnder interferometer (MZI) arrays. The basic idea is decomposing an arbitrary linear matrix into two unitary matrices and one diagonal matrix by using singular value decomposition (SVD) which can be easily realized by MZI arrays. Shen and Harris et al. [38, 39]. demonstrated a deep learning neural network utilizing a programmable nanophotonic processor chip. The chip consists of 56 MZIs and works as one optical interference unit (OIU) with 4 input ports and 4 output ports, shown as Fig. 1(b). In this work, two OIUs have been used to implement an effective arbitrary linear operator with 4*4 matrix size for inference process of ANNs, and a 76.7% correctness for vowel recognition is achieved compared with 91.7% in a digital computer. Later, Shen and Harris founded startup Lightelligence and Lightmatter respectively to promote this paperwork a step further toward to commercial applications [40, 41]. In 2020, Lightmatter published a board-device demo called 'Mars' on the HotChips 32 forum, which integrated an opt-electrical hybrid chip and other supporting electronic components [42]. The hybrid chip contains a photonics core supporting the multiplication between a 64-length vector and 64*64 matrix. An ASIC chip utilizing14 nm processing technology has been externally integrated for mainly driving active devices in the photonic core. Besides the impressive scale of operating matrix in photonic core, a new technology of nano-optical-electro mechanical system (NOEMS) has been adopted to reduce the power consuming of holding the status of MZIs. Since the matrix's updating rate is lower than vector's inputting rate, the chip's performance can be estimated from 0.4 *TMAC*/s to 4 *TMAC*/s depending on the refreshing frequency of weights.

Besides using MZI arrays with SVD method, there are other on chip architectures which can support the directly matrix loading. These architectures are similar to the systolic array in Google's TPU (tensor processing unit) and 'crossbar' design in the computing-in-memory field [43]. Varies types of modulators can substantially replace MZI to achieve multiplication in these architectures mentioned above. Here, the optical microring device is cited as a canonical example since its smaller footprint compared with MZI device. Several remarkable VMM works have demonstrated by combining the optical microring arrays with the WDM scheme [44–47]. A typical scheme is shown in Fig. 1(c), the vector data is loaded on different wavelengths and the matrix is implemented by an optical microring array. The wavelength-selectivity of optical microring can eliminate the crosstalk of data with different wavelengths. Recently, a massively parallel convolution scheme based on a crossbar structure has been proposed and experimentally demonstrated by Feldmann et al. [48]. In this work, a 16*16 'tensor core' based on crossbar architecture has been built on chip. The optical crossbar has been implemented by using crossing waveguides and PCM modulators embedded in the coupled waveguide bends, as shown Fig. 1(d). Moreover, a chip-scale microcomb has been employed as the multi-wavelength light source. With the fixed matrix data and 13 GHz modulation speed of the input vector, the performance of this chip can

reach more than 2 *TMAC*/s. Meanwhile, utilizing the PCM as a nonvolatile memory in computing is a wise approach for DNNs because the optical-electrical conversion overhead of weights data refreshing can be eliminated. Therefore the energy cost of system can be significantly reduce [46, 47, 49, 50].

Fourier transform, convolution and D²NN

VMM is a universal operator which can be used to do complex computing tasks, such as FT and convolution, with consuming more clock cycles. However, these complex computing tasks can be accomplished in one 'clock cycle' by adopting the inherent parallelism of photons. Theoretically, the process of coherent light wave deformed by an ideal lens and the process of FT can be equivalent. Based on this concept, a 4F system (Fig. 2(a)) can be used to do convolution processing. Since convolution is the heaviest burden in a CNN, Wetzstein et al. [51] made a good attempt on exploring in the optical-electrical hybrid CNN based on the 4F system. The weights of the trained CNN network have been loaded on several passive phase masks by elaborately designing the effective point spread function of the 4F system. The 90%, 78% and 45% accuracy have been achieved in the classification of MNIST, QuickDraw and CIFAR-10 standard datasets, respectively. Recently, Sorger et al. [52] demonstrated that the optical-electrical hybrid CNN still works well if the phase information in the Fourier filter plane is abandoned. In Sorger's demo, the weights of CNN have been directly loaded with the amplitude via a high speed DMD in the filter plane. However, it is disputable in theory that the amplitude-only filter can achieve the 98% and 54% classification accuracy of MNIS T and CIFAR-10.

Fig. 2 Complex matrix manipulation in optical computing. **a** 4F system. Two gray bars represent input data (*A*) and convolution results (*C*). The convex lens is Fourier lens that implements Fourier transform. The orange bar represents a matrix. **b** Schematic of optical convolution processor based on dispersion effect. **c** Schematic of diffractive deep neural networks with multi-layers of passive diffractive planes
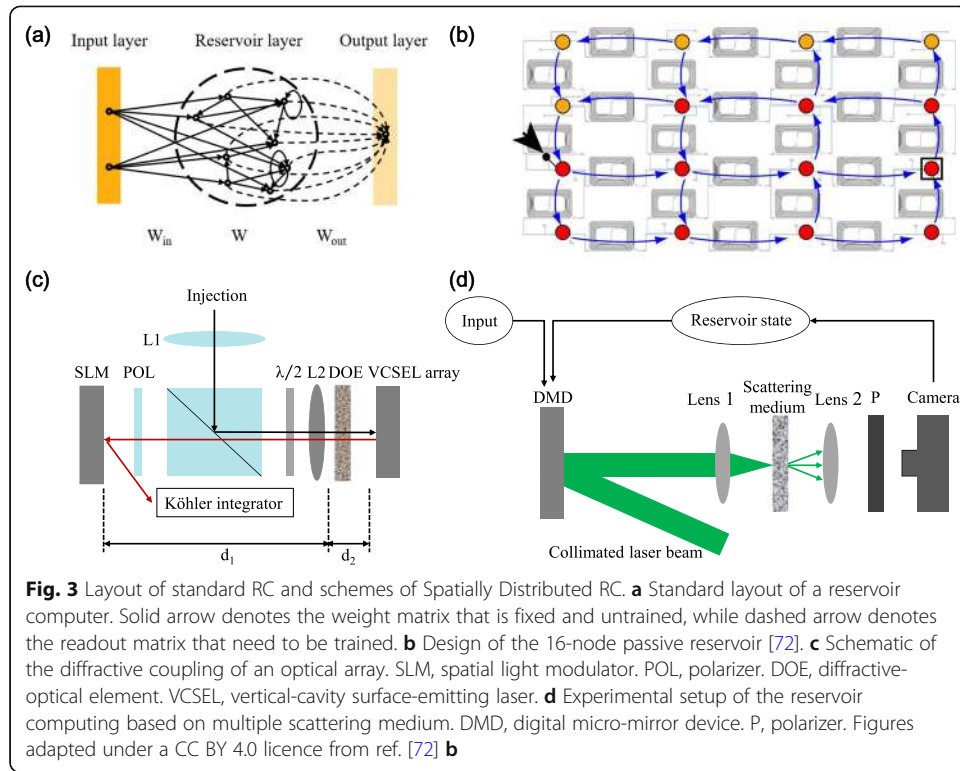
There are other alternative ways to realize FT and convolution in optical apart from the 4F based schemes mentioned above. Since the conventional lens is a bulky device, several types of effective lens, such as gradient index technology, meta-surface and diffraction structure by inverse designed, are considered as alternative devices to implement FT due to their miniaturized feature [53, 54]. However, the accuracy of computing based on these novel approaches has not yet been exploited fully. Besides the ways of effective lens, an integrated optical fast Fourier transform (FFT) approach based on silicon photonics has been also proposed by Sorger et al. [55]. In this paper, a systematic analysis of the speed and the power consuming has been given, and the advantages of integrated optical FFT comparing with P100 GPU (Graphics processing unit) have been figured out.

Apart from the implements of FT based on Fourier lens in space domain, the FT can be implemented in time domain with considering serial data inputting. The dispersion effect, caused by the propagation of multi-wavelength light in a dispersion medium, has been treated as the 'time lens' to achieve FT process in [56–58]. Recently this scheme is further used for the CNN co-processing [59, 60] via loading weights data and feature map data in wavelength domain and time domain, respectively. As shown in Fig. 2(b), the data rectangle is deformed to a shear form since the spectrum disperses in a dispersive medium, and the convolution results are finally detected by using a wide spectrum detector. In Ref. [60], an effective performance of ~ 5.6 *TMAC*/s and 88% accuracy for MNIST recognition have been achieved by utilizing time, wavelength and space dimensions enabled by an integrated microcomb source simultaneously.

In 2018, Ozcan et al. [61] proposed a new network called diffractive deep neural networks (D^2NN) for optical machine learning. This optical network comprises multiple diffractive layers, where each point on a given layer acts as a neuron, with a complex-valued transmission coefficient. According to the Huygens-Fresnel' principle, the behavior of wave propagation can be seen as a full connection network of these neurons (Fig. 2(c)). Although the activation layer has not been implemented, the experimental testing at 0.4 THz has demonstrated a quite good result, 91.75% and 81.1% classification accuracy for MNIST and Fashion-MNIST, respectively. One year later, the numerical work has shown the accuracy has been improved to 98.6% and 91.1% for the MNIST and Fashion-MNIST dataset, respectively. Moreover, that work also has demonstrated 51.4% accuracy for grayscale CIFAR-10 datasets [62, 63]. Besides, the classification for MNIST and CIFAR, the modified D^2NN's ability has also been proved for salient object detection (numerical result, 0.726 F-measurement for video sequence) [64] and human action recognition (> 96% experimental accuracy for the Weizmann and KTH databases) [65].

Optical reservoir computing

Reservoir computing (RC), which find its roots in the concept of liquid-state machine [66] and echo state networks [67], is a novel computational framework derived from recurrent neural networks (RNNs) [68]. It consists of three layers, named as input, reservoir, and output, as shown in Fig. 3(a). Different from general RNNs trained with back-propagation, such as LSTM and gated recurrent units (GRUs), only the readout coefficients denoted by W_{out} from the reservoir layer to the output layer need to be

Fig. 3 Layout of standard RC and schemes of Spatially Distributed RC. **a** Standard layout of a reservoir computer. Solid arrow denotes the weight matrix that is fixed and untrained, while dashed arrow denotes the readout matrix that need to be trained. **b** Design of the 16-node passive reservoir [72]. **c** Schematic of the diffractive coupling of an optical array. SLM, spatial light modulator. POL, polarizer. DOE, diffractive-optical element. VCSEL, vertical-cavity surface-emitting laser. **d** Experimental setup of the reservoir computing based on multiple scattering medium. DMD, digital micro-mirror device. P, polarizer. Figures adapted under a CC BY 4.0 licence from ref. [72] **b**

trained for a particular task for RC. The internal network parameters, namely the adjacency matrix W_{in} from the input layer to the reservoir layer, and the connections inside the reservoir W are untrained, which are fixed and random [67] or in a regular topology [69–71]. In the training phase of conventional reservoir computing architectures, the reservoir state is collected at each discrete time step n following

$$x(n) = f_{NL}(W_{in} \cdot u(n) + W \cdot x(n-1)) \tag{2}$$

where f_{NL} is a vector nonlinear function, $u(n)$ is the input signal, $x(n)$ is the reservoir state. In the case of the supervised learning, the optimal readout matrix W_{out} is obtained by ridge regression in general following

$$W_{out} = M_y \cdot M_x^T \cdot \left(M_x \cdot M_x^T + \lambda \cdot I\right)^{-1} \tag{3}$$

where M_x is the matrix which is concatenated by the reservoir state x with some training input vectors u, M_y is the target matrix that is concatenated by the ground truth corresponding to the training input vectors, I is the identity matrix, and λ is the regularization coefficient which is used to avoid over-fitting. In the testing phase, the predicted output signal $y(n)$ is calculated following

$$y(n) = W_{out} \cdot x(n). \tag{4}$$

Compared with general RNNs, the training time of RC is reduced by several orders of magnitude, which speeds up the time-to-result tremendously. Besides, RC has achieved the state-of-the-art performance for many sequential tasks [73, 74]. Last but not least, RC is very friendly to hardware implementation [73]. Due to the

aforementioned advantages, RC has attracted more and more attentions in research community. It has be utilized in signal equalization [67, 75–81], speech recognition [82, 83], time-series prediction or classification [82, 84–91], and de-noising in temporal sequence [92, 93].

The research on RC focuses on three aspects: the expansion of the application scope of RC, the optimization of the topological structure in the reservoir, and new physical implementation. The first aspect is devoted to using RC to solve specific tasks. The second aspect is aimed to reduce the computing complexity or increase the memory capacity of RC algorithm [69–71, 94–99]. The third aspect is about employing novel mechanism to realize or optimize RC [100, 101]. Limited by the scope of this paper, we concentrate on the third aspect, especially on the optoelectronic/optical implementations of RC.

Due to its inherent parallelism and speed, photonic technology is expertly suited for hardware implementation of RC. Over the past decade, the optoelectronic/optical implementations of RC has aroused great interest of researchers [95]. According to the way to achieve the internal connection in the reservoir, optoelectronic/optical RC can be divided into two categories: spatially distributed RC (SD-RC) and time-delayed RC (TL-RC) [95].

Spatially distributed RC, SD-RC

For SD-RC, it allows for the implementation of various connection topologies of the reservoir layer. In 2008, Vandoorne et al. suggested the implementation of photonic RC in an on-chip network of semiconductor optical amplifiers (SOAs) in numerical simulation, where SOAs are connected in a waterfall topology and the power-saturation behavior of SOA resembles the nonlinear function [100]. Soon after, researchers intended to optically reproduce the performance of the numerical counterparts [102, 103], realizing it is energy-inefficient to driving a SOA into power saturation results. Vandoorne et al. therefore proposed and demonstrated RC on a silicon photonic chip [72], which consists of optical waveguides, optical splitters, and optical combiners as shown in Fig. 3(b). Reservoir nodes are indicated by the colored dots, while blue arrows indicate the topology of the network. The nonlinearity was achieved by the photo detector, for photo detector detects optical power rather than the amplitude. This approach can deal with data in the rate of 0.12 up to 12.5Gbit/s. As for the disadvantages, the number of nodes in the reservoir, namely the reservoir size is restricted by the optical losses. Besides, it is difficult to measure response on all nodes in parallel. In 2015, Brunner and Fischer demonstrated a spatially extended photonic RC which is based on the diffractive imaging of the vertical cavity surface emitting lasers (VCSEL) using a standard diffractive optical element (DOE) [104]. The connection matrix in the reservoir is implemented by coupling between individual lasers of the VCSEL, where the bias current of each laser can be controlled separately. As shown in Fig. 3(c), an image of the VCSEL array is formed on the left side of the imaging lens. By fine-tuning the parameters of the system, after passing through the DOE beam splitter, diffractive orders of one laser will overlap with the non-diffracted image of its neighbors, thus achieving the connection of different neurons. By using the SLM located at imaging plane, the coupling weights can be controlled. The nonlinearity originates from the highly

nonlinear response of the semiconductor lasers. Following the VCSEL array reservoir, a Köhler integrator and detectors are utilized to collect the integrated and weighted reservoir state. The reservoir size of this system is limited by optical aberrations of the imaging setup. Except that, miniaturization is another issue need to be addressed for commercial applications. Brunner et al. further proposed a large scale photonic recurrent neural network with 2025 diffractively coupled photonic nodes using DOE [105] and investigated fundamental and practical limits to the size of photonic networks based on diffraction coupling [106]. They also investigated the noise's influence on the performance of the optoelectronic recurrent neural network [107]. In 2018, Jonathan et al. presented a novel optical implementation of RC using light-scattering media and a DMD [108]. As shown in Fig. 3(d), input and reservoir state are encoded on the surface of the DMD. After illuminating by the collimated laser, the encoded optical pattern pass through the multiple scattering medium, and detected by the camera. The mapping from the input to the reservoir and the internal connection in the reservoir are both realized by the optical transmission in the scattering medium instantly. Researches show the transmission matrix of the multiple scattering media is complex Gaussian matrix [109, 110], thus the internal connection in the reservoir of this setup is random and fixed. The reservoir state are recorded by the camera. One prominent advantage of this approach is that the reservoir size can be scaled easily and be expanded to even millions, which is challenging for the server based on conventional von Neumann computer architectures. Nevertheless, the calculation accuracy is limited by the experimental noise and encoding strategy. They further improved the performance of this system by using phase modulation [111] and demonstrated its feasibility for spatiotemporal chaotic systems prediction [112]. Inspired by this research, Uttam et al. put forward an optical reservoir computer for classification of time-domain waveforms by using multimode waveguide as scattering medium [113].

Time-delayed RC, TL-RC

For TL-RC, a discrete reservoir with a circular connection topology is formed due to the circular symmetry of a single delay line [114]. It uses only a single nonlinear node with delayed feedback. Figure 4 shows the general structure of a delay line based reservoir computer. In essence, TL-RC constitutes an exchange between space and time. In the input layer, a temporal input-mask W^{in} is used to map the input information $u(n)$ to the TL-RC's temporal dimensions, which results in N-dimensional vector $\mathbf{u}^{in} = (u_1^{in}, u_2^{in}, \cdots, u_N^{in})$ at each n, where $n \in \{1, 2, ..., T\}$. Thus, the TL-RC has to run at an N times higher speed compared with an N-node SD-RC, which is demanding for the modulators and bandwidth of the detector. Time multiplexing now assigns each $\mathbf{u}^{in}(n)$ to a temporal position denoted by $l \times \delta\tau$, where $l \in \{1, 2, \cdots, N\}$ denotes the index of the virtual nodes, $\delta\tau$ denotes the temporal separation or distance of virtual nodes. The mask duration τ_m equals $l \times \delta\tau$, while τ_D denotes the duration of the delay in the feedback loop. In this way, the input is mapped to the reservoir layer. Each virtual node can be regarded as a measuring point or tap in the delay line, whose value can be detected by a single detector. In the training phase, the reservoir state is sampled per $\delta\tau$. The samples are then reorganized in a state matrix which is used to calculate the readout matrix. Two mechanisms have been proposed to realize the internal connectivity inside

Fig. 4 Schematic illustration of the time-delay reservoir computer [114]. The input layer is implemented by modulating input $u(n)$ with temporal mask to create input $u^{in}(t)$. τ_m denotes mask duration, τ_D denotes the duration of the delay in the feedback loop, $\delta\tau$ denotes the temporal separation or distance of virtual nodes. The reservoir state is detected during one delay
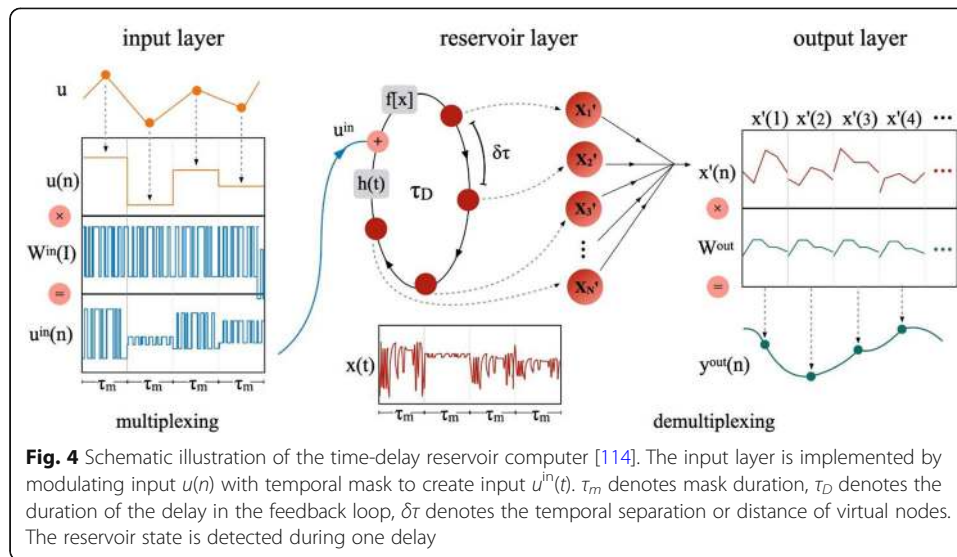
the reservoir. The first uses the system's impulse response function h(t), while the other use the de-synchronization between the input mask duration τ_m and the delay duration τ_D.

The first photonic implementations of RC based on time delay were independently by Larger et al. [115] and Paquot et al. [116]. Both implementations are based on the optoelectronic implementation of an Ikeda-like ring optical cavity. These systems use the concept of dynamical coupling via impulse response function h(t). For this, the temporal duration of a single node $\delta\tau$ is shorter than the system's response time, which results in connections according to the convolution-response h(t) and the neighboring nodes owing to inertia-based coupling. This approach is conductive to maximize the speed of TL-RC.

The other pioneering work was demonstrated by Duport et al. [117]. In this setup, the $\delta\tau$ is significantly larger than the system's response time, while input mask duration τ_m is smaller than the delay duration τ_D. A local coupling is introduced by setting $\delta\tau = \tau_D/(N+k)$, which results in node $x_l(n)$ is delay coupled to the node $x_{l-k}(n-1)$. This approach makes the mathematical model and numerical simulation process simplified. The operational bandwidth is reduced compared with the first approach, which may be profitable for the system's signal to noise ratio.

Following the above mentioned pioneering works, the TL-RC based on optoelectronic oscillators has been tested on various tasks that can be divided into two main categories: classification and prediction. More details can be found in the Yanne's review [118]. Except for the optoelectronic implementation, another branch of TL-RC is all-optical RC. In this branch, the nonlinear node is implemented by optical components such as semiconductor optical amplifier [117], semiconductor saturable absorber mirror [119], external-cavity semiconductor laser [120–122], and vertical cavity surface-emitting lasers [123].

The main advantages of optical/optoelectronic implementation of RC are the low power consumption and high processing speed, which results from the parallelism and speed of light. Integration or miniaturization of the system are the main challenges that optoelectronic/ optical RC need to be solved before commercial applications. More

importantly, the killer application of optoelectronic/optical RC are urgently to be demonstrated.
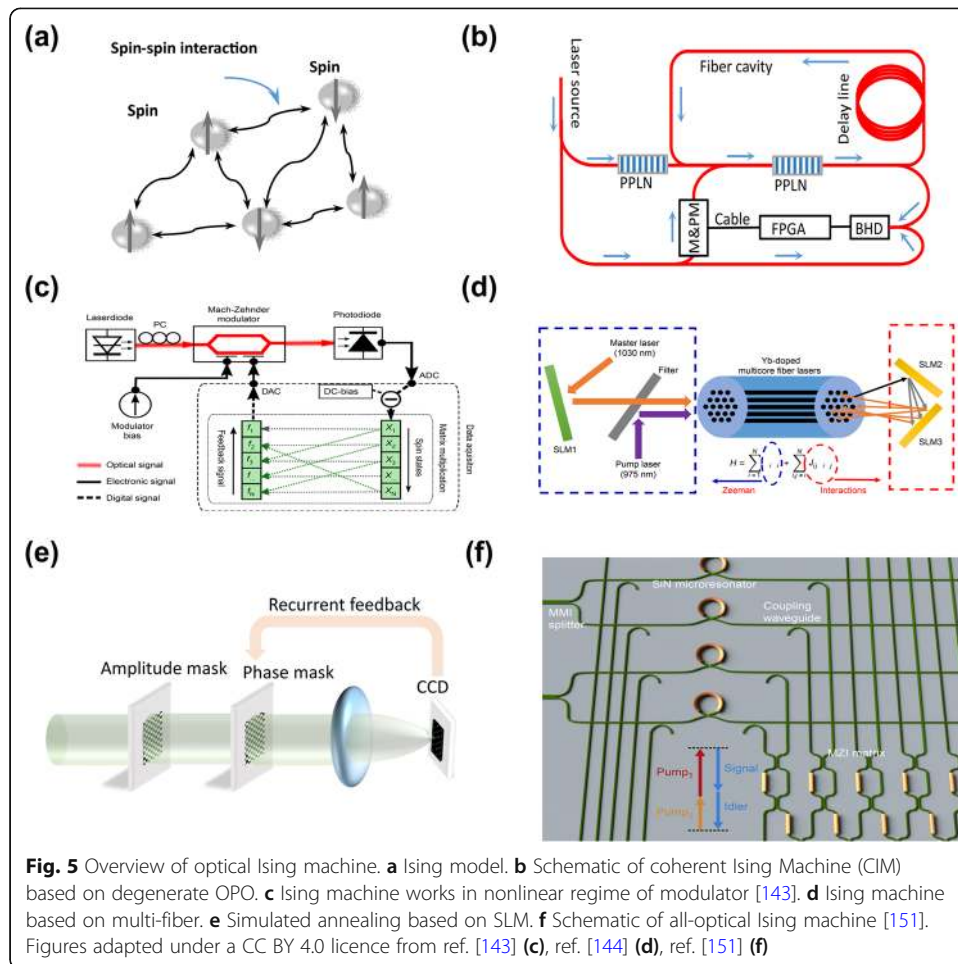
Photonic Ising machine

Numerous important applications, such as circuit design, route planning, sensing, and drug discovery can be mathematically described by combinatorial optimization problems. Many of such problems are known to be non-deterministic polynomial time (NP)-hard or NP-complete problems. However, it is a fundamental challenge in computer science to tackle these NP problems by conventional (von Neumann) computing architecture since the number of computational states grows exponentially with the problem size. This challenge motivates large amount of research work attempting to develop non von Neumann architectures. Fortunately, Ising model provides a feasible way to efficiently solve these computational-hard problems by searching the ground-state of the Ising Hamiltonian [124, 125]. Various schemes of simulating Ising Hamiltonian have been proposed and experimentally demonstrated in different physical systems, such as superconducting circuits [126], trapped ions [127], electromechanical oscillators [128], CMOS devices [129], memristors [130], polaritons [131] and photons [132–152]. Among these systems, photonic system has been considered as one of the most promising candidates due to its unique features, such as inherent parallelism, low latency and nearly free of environment noise, namely thermal and electromagnetic noise. In this section, the brief reviews of recent progress of photonic Ising machine (denoted as PIM hereafter) have been given and the main hurdles that hamper its practical applications have been clarified.

Before reporting research progress during last decade, the concept of Ising model is explained as follow. Figure 5(a) explicitly illustrates an Ising model with $N = 5$ spin nodes [138]. Each node occupies one spin state, either spin-up ($\sigma_i = +1$) or spin-down ($\sigma_i = -1$). $J_{i,j}$ represents interaction between two connected spins σ_i and σ_j. The Hamiltonian of Ising model without external field is given by

$$H = -\sum_{1 \leq i < j \leq N} J_{i,j} \sigma_i \sigma_j.$$

(5)

Driven by the interaction network and the underlying annealing mechanism, the Ising model could gradually converges into a particular spin configuration that minimizes the energy function (H). Three annealing mechanisms are illustrated in Fig. 5(b). One mechanism is simulated annealing (denoted as SA hereafter) relies on a specific annealing algorithm. Other two annealing mechanisms belong to a broad class of physical annealing (denoted as PA hereafter). Specific speaking, one is quantum annealing that harnesses quantum tunneling effect to identify the minimum state. The last one is optical parametric oscillation (OPO) gained network which relies on the mode selection in the dissipative system [132–141]. Here, apart from the OPO network, there are other peculiar mechanism being used to realize physical annealing as well, such as nonlinear dynamics in opto-electronic oscillators (OEO) [143].

Figure 5(a) and (b) indicates four indispensable elements of Ising machine: spin node, interaction network, feedback link and annealing mechanism. Taking advantages of various [143] degrees of freedom and appropriate technologies, numerous schemes have

Fig. 5 Overview of optical Ising machine. **a** Ising model. **b** Schematic of coherent Ising Machine (CIM) based on degenerate OPO. **c** Ising machine works in nonlinear regime of modulator [143]. **d** Ising machine based on multi-fiber. **e** Simulated annealing based on SLM. **f** Schematic of all-optical Ising machine [151]. Figures adapted under a CC BY 4.0 licence from ref. [143] **(c)**, ref. [144] **(d)**, ref. [151] **(f)**

been experimentally demonstrated during last decade [132–156]. Figure 5(c) to (f) show several exceptional works of PIM [132–152]. Meanwhile, the experimental data of relevant works is summarized in Table 1. Additionally, scalability and robustness are included in our discussion with the consideration of potential practical applications. These experimental demonstrations can be classified into three classes: fiber-based systems, free-space systems and chip-based systems. Each system is briefly explained in the next paragraph.

Fiber-based systems are shown in Fig. 5(b) and (c). Each spin node is represented by an optical pulse and their interaction network is implemented by optical delay [133, 134, 137, 138] or field programmable gate array (FPGA) [135, 136, 142, 143]. One advantage of fiber-based system is the excellent scalability that allows large-scale Ising model by increasing cavity length or repetition rate, while it suffers robustness issue result from a relatively short coherent time of photon. A mitigated approach is encoding the spin state in microwave signal since its coherent time is way longer than an optical signal [142]. Moreover, temporal multiplexing scheme constraints the scope of its applications as sequential processing sacrifices large part of annealing time. Figure 5(d) and (e) illustrate free-space systems. Spin node and interaction network are implemented by a fiber-core (or a pixel) and a SLM, respectively. In spatial domain, free-space system allows large-scale Ising model annealing simultaneously. Nevertheless, inevitable fluctuations in practical environment will ruin the interaction network

Table 1 Experimental data of different schemes shown in Fig. 5 [a]

Item/Scheme	OPO	Microwave	Modulator	Multi-fiber	SLM	MZI	Ring
Multiplexing	Temporal	Temporal	Temporal	Spatial	Spatial	Spatial	Spatial
Spin number	4-50 k	10 k	100	16	100-75 k	4	2
Annealing time	μs~ms	s	ms~s	μs	s	ms~s	ns~μs
Connection type	Full/Chain	Chain/Partial	Full	Full	Full	Full	Chain
Reconfigurable	Y	N	Y	Y	N	Y	N
Annealing Mech.	PA	PA	PA	PA	SA	SA	PA
Scalability	Excellent	Excellent	Excellent	Good	Excellent	Excellent	Excellent
Robustness	Poor	Good	Excellent	Poor	Poor	Excellent	Excellent
Reference	[132–141]	[142]	[143]	[144]	[145–148]	[149, 150]	[151, 152]

[a] *OPO* optical parametric oscillation, *PA* physical annealing, *SA* stimulated annealing, *SLM* spatial light modulator, *MZI* Mach-Zehnder interferometer

as it relies on the accurate alignment. Chip-based systems are shown in Fig. 5(f). A fully reconfigurable interaction network is implemented by MZI matrix [156, 157]. And the spin node can be built by a scalable building block, such as micro-ring resonator [151, 152]. Benefiting from the advanced CMOS technologies, chip-based system could potentially shrink a clumsy system into one monolithic/hybrid chip so that it is nearly immune from environmental fluctuation. Compared with the spin node demonstrated in other two classes, chip-based system is the "ugly-duckling" of approach to PIM. It will grow into a swan after we tackle several technical challenges. These challenges will be included into the following discussion.

Based on these extensive research woks, the technical roadmap of PIM becomes crystal clear. It is to develop a highly scalable, reconfigurable and robust PIM that could find an optimal (or a near optimal) solution of a large-scale combinatorial optimization problem in a polynomial time. Table 1 indicates the fiber-based scheme [141–143] and the chip-based scheme [149, 151] are two promising pathways as they satisfy scalability and robustness simultaneously. However, both schemes are severely limited by the scale of the interaction network since practical applications requires large amount of spin node. In fiber-based scheme, a creative solution is rebuilding the feedback signal after balanced-homodyne detection (BHD) and VMM in FPGA [135, 136, 142, 143]. The cost is extra process time required for synchronization between the optical signals within cavity and the external feedback signals. Besides additional time consumption, electro-optical conversion and VMM in FPGA are the potential bottleneck for the large-scale PIM. One plausible solution is utilizing $N-1$ optical delay lines with modulator in each line so that generate feedback signal instantaneously [139].

In chip-based scheme, the interaction network requires an overwhelming number of optical unit ($\propto N^2$, where N represents spin number) [156, 157]. To the best of our knowledge, the largest MZI matrix (64*64) developed by Lightmatter is still smaller than the dimension of practical models [42]. Alternatively, nonlinear effect, such as frequency conversion via $\chi^{(2)}$ / $\chi^{(3)}$ medium [154, 158, 159], could be a viable approach to build interaction network on a large scale. Meanwhile, the giant model of practical problems can be split into many sub-models so that we can solve these sub models sequentially or simultaneously by chip-based systems with a comparable matrix size. Besides the aforementioned technical challenges, experimental verification of the parallel search or the ergodicity of spin configuration in PIM, particularly in coherent Ising

machine (CIM) [139], is another haunting research work. Because this work would explicitly explain the advantage of PIM over von Neumann computing architecture.

The promising results of PIM achieved over last decade indicate a feasible way to solve computational hard-problems. However, this research direction needs continuous research effort to build a scalable, reconfigurable and robust PIM which will make profound impact on our society.

The new challenges and opportunities for optics

As explained in the Chapter 2, analog optical computing is considered as an alternative approach to execute complex computing in the post-Moore era. Compared with electrical computing, one prominent advantage of optical computing is negligible energy consumption when multiplication is performed in optical domain. However, the actual benefit of such a hybrid opt-electrical system should be systematically analyzed, especially the cost of transferring data between different domains and formats has not yet been discussed. In this chapter, the energy consumption and calculation precision in the hybrid opt-electrical computing system are discussed in "Hybrid computing system" section. In "New challenges and prospects" section, we prospect the new challenges and opportunities of analog optical computing in the future.

Hybrid computing system

In the section, energy consumption of hybrid computing system and the speed-up factor, S, have been clearly explained in the first half. Then, the calculation precision of analog optical computing has been analyzed and the potential solutions to suppress errors are proposed at the end of this section.

The aforementioned difficulties, such as coherent storage and logic operation, indicates a hybrid architecture would be a promising solution for analog optical computing. A typical architecture is illustrated in Fig.6(a). The gray and orange parts indicate electrical and optical domain. Presume this hybrid architecture is implementing large-scale VMM. The electrical processor, like CPU, offers external support, such as data reading/ storing, logic operation and pre/post processing. Assisted with DACs (digital to analog convertors) and ADCs (analog to digital convertors), the vector data is regenerated by an array of light sources (referred as Tx in Fig. 6(a)), and the matrix is loaded into modulators (referred as MD in Fig. 6(a)). The calculation results are collected by detectors (referred as Rx in Fig. 6(a)). Such a system could be an exceptional accelerator in specific scenarios since large amount of repeatable tasks are implemented in optical domain. While, a rigorous and systematical analysis is indispensable before practical applications.

In the following paragraphs, the performance and power consumption of the hybrid optical computing system are explicitly discussed. Similar to CPU, a clock frequency of an optical processor unit (denoted as OPU hereafter) is defined as $F_{clc} = \frac{1}{T_{clc}}$, where T_{clc} is the clock time of OPU. Practically, T_{clc} is constrained by the response time of opt-electric devices (such as tunable laser, modulator and photon detector) or electric converters (DAC, ADC and amplifier), rather than the propagation time of optical length. The performance of an OPU is defined as:

Fig. 6 Overview of optical-electrical hybrid computing system. **a** Schematic diagram of architecture for the optical-electronic hybrid system. **b** Ratio of power cost to performance. *S* is the speed-up factor in Eq. (6). $\tilde{E}$ represents energy budget per channel per symbol. The bar chart below represents a typical energy-per-symbol distribution. **c** Schematic illustrations for finite precision analysis in OPU. The upper panel shows the propagation routes of data in a VMM, with the blue solid arrow line and the green dash arrow line indicating correct and crosstalk routes respectively. The bottom left panel shows the deviation between actual physical quantity and ideal data. The bottom right panel depicts the relationship of accumulated error and bit precision in computing. **d** Convolution result of OPU with equivalent 4bit output precision

$$Perf = \frac{operations}{time\ cost} \sim F_{clc} \cdot N \cdot S(N) \quad (6)$$

Here, N is the number of lanes in the processor, and $S(N)$ is an effective speed-up factor that indicates the number of operations per lane and per clock time. Moreover, S factor also represents the fan-in/out in specific computing process, such as VMM. Apparently, improving the performance by increasing the N and F_{clc} is a conventional and reliable way both for CPU and OPU, while the effective speed-up factor $S(N)$ is the key to release unprecedented computing capabilities of the OPU due to the bosonic characteristic of photon. A more comprehensive discussion of S factor is conducted in the paragraph after Table 2.

In this hybrid system, energy consumption in optical domain is negligible. The main power consumption comes from the O/E (& E/O) conversion and A/D (& D/A) conversion. The entire power consumption of OPU can be written as:

$$P = P_{Tx} + P_{MD} + P_{AD} + P_{DA} + P_{TIA}. \tag{7}$$

The terms P_{Tx}, P_{MD}, P_{AD}, P_{DA} and P_{TIA} represent the power of transmitters, modulators, ADCs, DACs and TIAs (Transimpedance Amplifier), respectively. To further

Table 2 Typical value of energy consumption per symbol operating in each device of OPU [a]

Emission domain	Processing domain	Detecting domain
$p_{static}^{Tx} \Big/ F_{clc}$	$E_{symb}^{Tx} \sim E_{symb}^{MD}$	E_{symb}^{TIA}
10^0 [b]	$10^{-4} \sim 10^0$ [c] $10^{-3} \sim 10^{-2}$ [d]	$10^0 \sim 10^1$ [e]
/	E_{symb}^{DA}	E_{symb}^{AD}
/	$10^0 \sim 10^3$ [f]	$10^0 \sim 10^3$ [f]

[a] The unit of variables in this table is pJ/symb or mW/GHz

[b] It's an estimated value assuming that a 10^0 mW semiconductor light emitter is employed in each lane and F_{clc} = 1 GHz

[c] Using the silicon modulators based on carrier effect as reference [160–166]

[d] Using novel modulators based on surface plasmon polariton or hybrid plasmonic mechanism [160]

[e] The voltage feedback TIA's power is usually restricted by the gain-bandwidth product. E_{symb}^{TIA} is the estimated value based on the data from ref. [167, 168], assuming the gain~10^4 and the bandwidth~10^0 GHz

[f] E_{symb}^C (C indicates AD or DA) is usually proportional to 2^n (n means n-bit conversion resolution). The value here is estimated both from the datasheet of commercial products [42, 169, 170] and academic works [171] at the condition of 8-bit conversion resolution

simplify the followed discussion, presume these devices can operate at high speed and they have been optimized to be power efficient. Thereby, P_{MD}, P_{AD}, P_{DA} and P_{TIA} are determined by their dynamic power, which is proportional to $CV^2 \times Frequency$ [160, 161, 167, 168, 171]. The variable C and V represent the capacitance and driving voltage, respectively. P_{Tx}, the power of transmitters, can be divided into two parts: the static and dynamic part. So is P_{MD}, the dynamic part is also proportional to F_{clc}. Assuming there are no additional amplifiers embedded in the hybrid system, and each electro-optical device is driven by an independent DAC or ADC. Therefore, the total power of system can be re-organized as:

$$P_{Tx} + P_{MD} + P_{AD} + P_{DA} + P_{TIA} \gtrsim p_{static}^{Tx} N_{Tx} + E_{symb}^{Tx} N_{Tx} F_{clc} + E_{symb}^{MD} N_{MD} F_{MD}$$
$$+ E_{symb}^{DA} N_{Tx} F_{clc} + E_{symb}^{DA} N_{MD} F_{MD} + E_{symb}^{AD} N_{Rx} F_{clc}. \tag{8}$$

Here, p_{static}^{Tx} is the static power in one Tx. E_{symb}^{X} represents the energy cost per symbol operating in a single device X (X indicates Tx, MD, DA or AD). N_Y is the total amount of device Y (indicates Tx, MD, Rx). F_{MD} is the operating frequency of MD.

In this review, a conventional term, operation power per second (W/Tops), is used as an appropriate benchmark since energy consumption of most devices in the system is proportional to the operation numbers. In a semi-quantitative view, the power of one ADC is comparable with that of one DAC at the same precision, architecture design and manufacture procedure (i.e. $E_{symb}^{DA} \sim E_{symb}^{AD} = E_{symb}^{C}$, the superscript C means converter). In addition, we assume $N_{Tx} = N_{Rx} = N_{lane}$. Then, the operation power per second can be described as:

$$\frac{Power}{Perf} = \frac{P_{Tx} + P_{MD} + P_{AD} + P_{DA} + P_{TIA}}{F_{clc} \cdot N_{lane} \cdot S} \gtrsim \left[\frac{p_{static}^{Tx}}{F_{clc}} \right] \cdot \frac{1}{S}$$
$$+ \left[E_{symb}^{TIA} + E_{symb}^{Tx} + 2E_{symb}^{C} + \left(E_{symb}^{MD} + E_{symb}^{C} \right) \cdot \frac{N_{MD}}{N_{lane}} \cdot \frac{F_{MD}}{F_{clc}} \right] \cdot \frac{1}{S}. \tag{9}$$

If ultra-low power modulators are used, E_{symb}^{Tx} and E_{symb}^{MD} can be neglected compared with E_{symb}^{C}. After defining $k = \frac{N_{MD}}{N_{lane}} \cdot \frac{F_{MD}}{F_{clc}}$ and $\tilde{E} = \frac{p_{static}^{Tx}}{F_{clc}} + E_{symb}^{TIA} + E_{symb}^{C}(2 + k)$, the final equation is:

$$\frac{Power}{Perf} \gtrsim \frac{\tilde{E}}{S} = \left[\frac{p_{static}^{Tx}}{F_{clc}} + E_{symb}^{TIA} + E_{symb}^{C}(2+k) \right] \cdot \frac{1}{S}. \tag{10}$$

A lower *Power/Perf* means a higher energy efficiency of the system. Table 2 lists the typical value of energy per symbol operating in each device used in the OPU system, such as Tx, MD, DA, AD and TIA.

This Eq. (10) together with Table 2 show that the system's operation power per second would be mainly constrained by the energy consumption per operation of electrical devices (TIA, DAC, ADC). Obviously, the energy consumption per operation of these electrical devices is difficult to be improved significantly in the post-Moore's era. Therefore, the speed-up factor S is the essential parameter to improve the system's energy efficiency. According to the 10^{0} mW/Gops operation power per second of nowadays AI chips, the competitive operation power per second of a OPU should be $\sim 10^{-1}$ mW/Gops. Fig. 6(b) demonstrates the relationship of OPU's *Power/Perf*, $\tilde{E}$ and speed-up factor S based on Eq. (10). In this figure, the horizontal axis $\tilde{E}$ can be seen as the energy budget per channel per symbol operation for the OPU. To achieve the bellowing 1 mW/Gops *Power/Perf* of OPU, a factor S with the value of tens is needed. Consequently, the $\tilde{E}$ can be higher than 10 pJ/symb which is given as an example by the green dot in Fig.6(b). If the same *Power/Perf* of OPU is achieved with $S = 1$, the total energy consumption of devices per operation per channel will be limited within 1 pJ/symb. In other words, a higher speed-up factor S could bring a lower operation power per second of the system and relax the energy consumption requirement of electrical devices.

Apart from the energy consumption, the calculation precision is another problem which needs to be concerned and investigated. Compared to digital computing, one of the main drawbacks in analog computing is the systematic errors. In this section, the universal finite precision analysis has been discussed in first. Then, the fundamental causes of various errors have been investigated. In final, the criteria of error control, the effects of bit-depth, and the methods of error compensation have been proposed.

It is clear that VMM is one of the most popular parallel optical computing systems. In addition, the main mechanisms of error in optical computing systems, such as error propagating, error converging and signals interfering, can coexist in same VMM system. Therefore, the VMM system has been proposed as the universal instance for the finite precision analysis in here.

As shown in Fig.6(c), the ideal relationship between the input data and the output data of the system can be illuminated as Eq. (1) in Chapter 2.1. However, the modulation, transmission and detection of analog signal are unideal in fact. Therefore, the realistic rules of the information indicated quantities in Fig.6(c) can be written as below:

$$\tilde{C} = \epsilon + (1 + \Delta c) \left[\boldsymbol{T} \cdot \tilde{\boldsymbol{B}} \circ \boldsymbol{S} \cdot \tilde{A} \right]. \tag{11}$$

In Eq. (11), the vector $\tilde{A}$ is optical physical value (intensity or complex amplitude) of the input data A after applying on the Tx array, the matrix $\tilde{\boldsymbol{B}}$ is optical physical value of the input matrix B after applying on the MD array, $\boldsymbol{S}$ is the transfer tensor of optical signal propagating from the Tx array to the MD array, and $\boldsymbol{T}$ is the transfer tensor of optical signal propagating from the MD array to the Rx array. The vector $\tilde{C}$ is the output data of Rx array by detecting the optical signal. Because the Rx array is unideal in

reality, the proportional error of optical-electrical conversion is unneglectable and described as Δc, and the rest parts of systematic error is referred as ϵ. The symbol '∘' refers the Hadamard product operation in Eq. (11). Based on Eq. (11), the detecting output $\widetilde{C}_l$ of anyone receiver l among the Rx array can be written as:

$$\widetilde{C}_l = \epsilon_l^C + (1 + \Delta c_l) \left(\sum_{CR} i + \sum_{XT} i \right)$$
$$\times \left\{ t_{lkj} \cdot \left[(1 + \Delta b_{kj}) B_{kj} + \epsilon_{kj}^B \right] \cdot s_{kji} \cdot \left[(1 + \Delta a_i) A_i + \epsilon_{kj}^A \right] \right\}. \tag{12}$$

The variables apart from A_i and B_{kj} cited in Eq. (12) have been normalized with dimensionless (A_i and B_{kj} are the element of input vector and matrix, respectively). Δa_i and Δb_{kj} represent the proportional error of the corresponding element in $\tilde{A}$ and $\tilde{B}$, respectively. Other errors in vector $\tilde{A}$ and matrix $\tilde{B}$ are indicated as ϵ_{kj}^A and ϵ_{kj}^B. s_{kji} and t_{lkj} represent the element of transfer tensor S and T, respectively. $\widetilde{C}_l$ is the realistic output with errors both from ideal propagation paths $\Sigma_{CR}(error)$ and unideal propagation paths $\Sigma_{XT}(error)$, which are indicated by the blue solid line with arrow and the green dash line with arrow in Fig. 6(c) respectively.

Based on the Eq. (12), the summarized error $\Delta C_l = \widetilde{C}_l - C_l$ can be rewritten as expanded polynomial with containing higher order terms. In a well-designed system, the deviation value of each variable could be far less than 1. The errors of variable deviation with higher order can be neglected and the polynomial of ΔC_l can be shorted as below:

$$\Delta C_l = \Delta^{(2)} + \Delta^{(1)} + \Delta^{(0)} + \Delta^{XT}, \tag{13}$$

$$\Delta^{(2)} = \sum_{CR} i \; B_{kj} A_i \left(\Delta t_{lkj} + \Delta s_{kji} + \Delta a_i + \Delta b_{kj} + \Delta c_l \right), \tag{14}$$

$$\Delta^{(1)} = \sum_{CR} i \; A_i \epsilon_{kj}^B + B_{kj} \epsilon_i^A, \tag{15}$$

$$\Delta^{(0)} = \epsilon_l^C, \tag{16}$$

$$\Delta^{XT} = \sum_{XT} i \; B_{kj} A_i \left(t_{lkj}^{XT} + s_{kji}^{XT} \right). \tag{17}$$

$\Delta^{(2)}$ describes the two main deviation errors: the response factor deviations (Δa_i, Δb_{kj}, Δc_l) of active devices and the transmission factor deviations (Δs_{kji}, Δt_{lkj}) of passive devices, between theory and reality. $\Delta^{(1)}$ gives the error caused by the limited linearity and extinction ratio of modulators. The extinction ratio in here is defined as $\epsilon^{ER} = 2^{bit\ depth}/ER$ (ER is the value of extinction ration, e.g. $\epsilon^{ER} = 0.16$ under $ER = 20$ dB, *bit depth* = 4). $\Delta^{(0)}$ indicates the background error of detectors and backend circuits. Δ^{XT} shows the crosstalk errors of the system. On the ideal propagation paths, s_{kji}^{XT} and t_{lkj}^{XT} must be zero. However, the crosstalk error can be accumulated on the unideal propagation paths, especially in spatial optical systems. All the errors of optical computing system discussed above can be classified as systematic error and random error. Table 3 shows the details for these two kinds of errors.

Table 3 Classification and sources of error parts in OPU [a]

Error parts	Factors		Error type	Sources	Calibrating or compensating
$\Delta^{(2)}$	$\Delta s,$ Δt	/	systematic	• non-uniformity of passive devices	possible
	$\Delta a,$ $\Delta b,$ Δc	$\delta^s a, \delta^s b,$ $\delta^s c$	systematic	• non-uniformity of active devices	possible
				• unflatness of source spectrum [b]	
				• unflatness of spectral response [b]	
		$\delta^r a, \delta^r b,$ $\delta^r c$	random	• fluctuation of electrical supplier	difficult
				• fluctuation of light source	
				• environmental perturbation	
				• additional noise from amplifier	
				• shot noise of electrons & photons	
				• thermal noise	
$\Delta^{(1)}$	$\epsilon^A,$ ϵ^B	$\epsilon^{ER\_A},$ ϵ^{ER_B}	systematic	• limited extinction ratio	depends
		$\epsilon^{NL\_A},$ ϵ^{NL_B}	systematic	• limited linearity of electro-optical response	possible
Δ^{XT}	$t^{XT},$ s^{XT}	/	systematic	• crosstalk between routes	difficult
				• crosstalk between wavelength channels [b]	
$\Delta^{(0)}$	ϵ^C	/	systematic + random	• dark current in detector	difficult
				• other gound noise	

[a] The subscript of each variable listed in the table is omitted for notional simplicity
[b] Additional considerations are required when WDM technology is applied to implement computing task

Due to the lack of Boole logic and limited SNR, the integer number is an appropriate format for analog optical computing rather than floating point. Presume 8 bit is the required calculation precision, if the length of the input vector is 16, then, each element in vector A and matrix B only need 2 bit precision. The aforementioned error ΔC_l includes systematic error $\delta^s C_l$ and random error $\delta^r C_l$. Without loss the generality, the normal distribution is applied to described $\delta^r C_l$ and its standard deviation is σC_l. The detected result and error margin are shown in Fig.6 (c). In order to obtain correct value, error should be carefully controlled within the region of six sigma ($\pm 3\sigma$), correspond 99% correctness. And its error can be described by

$$\delta^s C_l + 3\sigma C_l < 0.5. \tag{18}$$

After deducing with Eqs. (13-18), a general guidance of suppressing error is obtained. When the major error induced by the poor uniformity, the overall deviation should satisfy $\overline{\Delta s} + \overline{\Delta t} + \overline{\Delta a} + \overline{\Delta b} + \Delta c_l < \frac{0.5}{255}$ (nearly 0.2%). If extinction ratio plays a key role, ϵ^{ER} for input vector A and matrix B satisfy $\frac{1}{ER_A} + \frac{1}{ER_B} < \frac{0.5}{255}$. This criterion indicates the average extinction ratio is 30 dB. When cross-talk noise dominates the error, the entire cross-talk exists in the transfer tensor S and T should be suppresses less than 0.1%. Furthermore, the random error with independent lane is written as

$$\sigma^2 C_l = \left(\sigma^2 a + \sigma^2 b\right) \sum_{\substack{i \\ CR}} \left(B_{kj} A_i\right)^2 + C_l^2 \sigma^2 c_l, \tag{19}$$

where $\sigma^2 C_l$ is the variance of the random error. In most applications, $B_{kj}A_i$ of different lane are independent with one another. In this scenario, the expectation value of $\sqrt{\dfrac{\sum_i (B_{kj}A_i)^2}{CR}}$ would be several times smaller than the expectation value of C_l.

Thereby, the standard variance of detection module (σc_l) is more stringent than other modules ($\sqrt{\sigma^2 a + \sigma^2 b}$). For example, in calculation for 8 bit output (correspond 255 intervals), σc_l and $\sqrt{\sigma^2 a + \sigma^2 b}$ should be controlled within 0.06%, and 0.2%, respectively.

In a practical system, the major part of systematic error ($\Delta^{(2)}$) comes from the poor uniformity of each module, such as input laser sources and modulator array, and its typical value is $0.1 \sim 0.2$. Fortunately, this part can be compensated or suppressed with specific design and algorithm. Beside $\Delta^{(2)}$, part of the $\Delta^{(1)}$ error, such as ϵ^{NL-A} and ϵ^{NL_B} induced by the unideal linearity of respond curve, can be overcame by reconfiguring input electrical signal. However, the precision of electrical signal should be higher than the input data. Moreover, limited SNR induces ϵ^{ER-A} and ϵ^{ER-B} which cannot be eliminated by fine adjustment in hardware. One potential solution is post-processing by a particular algorithm, but the trade-off is scarifying parts of computing capability. The most challenging task is suppressing the crosstalk noise. The potential route of crosstalk is several times higher than the number of correct routes. The accumulated error can be magnified if t^{XT} and s^{XT} are non-trivial. After eliminating systematic error in analog optical computing, the random noise becomes the main obstacle to improve computing precision, such as fluctuation from electrical power supplier or light source, noise from amplifier, thermal and shot noise. The first two types of random noise can be suppressed by employing special hardware design. Cryogenic environment is a potential solution to mitigate the thermal noise. The shot noise can be circumvented by using an appropriate power scheme, such as increasing the bit interval power (see the bottom right panel in Fig.6(c)), in analog computing. For example, in calculation for 8 bit output (correspond 255 intervals), 10 μW per interval at Rx is sufficient to guarantee high correctness, because the corresponding standard deviation (0.005%) is much smaller than the aforementioned value (0.06%).

The methodology explained above is compatible with the proposed hybrid computing system shown in Fig. 6(a). In our proof-of-principle demonstration, the hybrid system is utilized to implement CNN tasks, such as the handwritten digits recognition task. Since the inference process relies on logic results rather than analytic solutions, CNNs have higher error tolerance than conventional analytic computations in same system. Additionally, the systematic error existing in our experiment setup is suppressed by retraining the weight parameters of CNNs. Thanks to the retraining method and high tolerance feature, the proposed hybrid system achieves 4 bit output precision in optical convolution and 96.5% accuracy in the recognition of handwritten digits (MNIST dataset), as shown in Fig.6(d). This experimental demonstration offers a solid experimental foundation to analyze the achievable highest precision of optical computing. Therefore, it is essential to figure out relevant scenarios which can be applied with limited precision.

New challenges and prospects

Following the discussion above, there're some general challenges for the variable approaches of optical computing. Firstly, the manufacture technology for large scale

integration of optical-electrical chip is firmly needed to improve the parallelism of optical computing system in hardware level. Furthermore, the optical-electrical co-package technology is also need to reduce the cost of transferring the data between electrical and optical domain.

Secondly, the modern optical transmitters and modulators are designed for optical communication, rather than computing tasks. For example, optical computing system requires much higher extinction ratio and linearity of optical device than optical communication in most applications, because the input data of most applications is high bit depth. In addition, the higher extinction ratio and linearity of optical devices can support high efficiency optical coding for data input, the systematic throughput will be improved.

Thirdly, new architecture design is essential. The conventional computing architecture is difficult to take the advantages of optical computing as the optical-electrical conversion could heavily limit the energy efficiency of the hybrid computing system. The new architecture design could has large speed-up factor S (Eq. (6), i.e. process much more operations with few active devices) and retain the configurability as much as possible meanwhile.

In last, there is few explorations in algorithms which are suitable for analog optical computing. Currently, algorithms are designed based on the Boole logics which is suitable for digital computing system. However, they are difficult to match the operators provided by optical computing. If the algorithms are developed for optical computing, the operation complexity and the executing time of them will be much shorter than that of current ones.

Through there are many challenges, the opportunities of optical computing has been rising. Firstly, many fabrications have been involved in developing the larger scale integration of optical-electrical chips. For example, the Lightmatter released the world first 4096 MZI integrated chip 'Mars' with proving the feasibility of large scale integration, and brought more confidence for the people researching in optical computing. In addition, the WDM and MDM mentioned in before and the spatial optical system are also compatible for the parallelism improving.

Secondly, the low extinction ratio and linearity of optical devices can be compensated by using the higher speed optical device with low bit depth optical coding directly. For example, a 2GHz optical modulator with OOK and a 1GHz optical modulator with PAM4 are equivalent in data input efficiency. However, this kind of compensation is only feasible in certain computing processes which can be converted to the linear combination of series low bit depth operations in time domain. In contrast, employing low bit depth quantization for the input data of applications is a pervasive solution for making the modern optical devices to be practicable in optical computing.

Thirdly, to reduce the overhead from optical-electrical conversion in hybrid computing system, optical signal looping needs to be fully utilized for keeping the data in optical domain as long as possible. Because of the high propagation speed of light, the time delay caused by optical signal looping can be negligible. The stream processing methodologies can inspire the new architectures.

Lastly, the algorithms developed for optical computing could consider the complex operators provided in optical domain. Some sets of Boole logic operators in current algorithms can be replaced by one complex operator to reduce the complexity and

execution time in total. Therefore, combining the complex operators with the Boole logic operators in an algorithm is the potential way to develop the suitable algorithms for optical computing.

Obviously, the opportunities of optical computing have been rising. The growing demand of artificial neural network and its computing hunger would continuously drive the researches in optical computing patterns. The optical sensing and optical communication may give another chance for optical computing to be employed. In addition, the approaches of high complexity computing in optical domain, such as Fourier transform, convolution and equation solving, could effectively improving the systematic efficiency. In a word, the optical computing has been considering as the "elixir" in the post Moore's era.

Conclusions

In this paper, a systematic review has been presented on the state-of-the-art analog optical computing, mainly focusing on the fundamental principles, optical architectures, and their new challenges. Firstly, a brief introduction of the slowing down of Moore's law has been given, which is mainly hindered by the 'heat wall' and the difficulty of manufacturing. Meanwhile, the challenges from growing demands of information processing have been discussed. And the attempt to improve the computing capability also have been investigated.

Then, the state-of-the-art analog optical computing, as one approach of 'Beyond Moore', is reviewed in three directions: vector/matrix manipulation, reservoir computing and photonic Ising machine. The vector/matrix manipulation by optics includes the VMM and other more complex processing, such as FT, convolution, and even directly applied in neural network by stacking diffractive layers. The optical reservoir computing is introduced and divided into SD-RC and TD-RC. After that, we review the principle of photonic Ising machine and take a brief comparison of varies schemes. After talking about the ability of analog optical computing, some preliminary discussion of computing efficiency is introduced, mainly about the ratio of performance and power dissipation. The power dissipation in electric convertors predominates in the hybrid computing system and the architectures with higher speed-up factor will take more advantages. Moreover, a comprehensive discussion of systematic and random error indicates achieving high precision optical computing require dedicated work in both hardware and algorithm.

To promote analog optical computing into practical application, the problems of large scale integration technologies, appropriate devices, and suitable algorithms are need to be solved essentially. In fine, the opportunities of optical computing in the post-Moore era is rising, and the prospects of optical computing are bright.

Abbreviations
CMOS: Complementary-metal-oxide-semiconductor; CPU: Central processor unit; ITRS: International Technology Roadmap of Semiconductors; WDM: Wavelength division multiplexing; MDM: Mode division multiplexing; EUV: Extreme ultraviolet; EBL: Electron beam lithography; FinFET: Fin field-effect transistor; DNN: Deep neural network; CNN: Convolution neural network; LSTM: Long short-term memory; MEMS: Micro-electromechanical system; CNT: Carbon nanotubes; SRAM: Static random access memory; NVM: Non-volatile memory; PCM: Phase change material; RRAM: Resistive random-access memory; VMM: Vector matrix multiplier; FT: Fourier transformation; SLM: Spatial light modulator; MAC: Multiply-accumulate operation; DSP: Digital signal processor; GaAs: Gallium arsenide; DMD: Digital micro-mirror device; LCoS: Liquid-crystal-on-silicon; MZI: Mach-Zehnder interferometer; SVD: Singular value decomposition; OIU: Optical interference unit; NOEMS: Nano-optical-electro mechanical system; TPU: Tensor processing unit; NPU: Neural network processing unit; FFT: Fast Fourier transform; GPU: Graphics

processing unit; D^{2}NN: Diffractive deep neural networks; RC: Reservoir computing; RNN: Recurrent neural network; GRU: Gated recurrent unit; SD-RC: Spatially distributed reservoir computing; TL-RC: Time-delayed reservoir computing; SOA: Semiconductor optical amplifiers; VCSEL: Vertical cavity surface emitting lasers; DOE: Diffractive optical element; NP: Non-deterministic polynomial time; PIM: Photonic Ising machine; PA: Physical annealing; OPO: Optical parametric oscillation; OEO: Opto-electronic oscillators; CIM: Coherent Ising machine; FPGA: Field programmable gate array; BHD: Balanced-homodyne detection; DAC: Digital to analog convertor; ADC: Analog to digital convertor; TIA: Transimpedance Amplifier; OPU: Optical processor unit; WPE: Wall-plug efficiency; QCSE: Quantum-confined Stark effect; SNR: Signal noise ratio

Authors' contributions
Methodology, XD, CL; writing—original draft preparation, CL (Introduction, chapter 2.1, chapter 3), XZ (chapter 2.3), JL (chapter 2.2), TF (chapter 2.1), XD (chapter 1); writing—review and editing, CL, XZ, JL, XD; supervision, XD. All authors read and approved the final manuscript.

Availability of data and materials
The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Declaration

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

References
1. Dennard RH, Gaensslen FH, Yu H-N, Rideout VL, Bassous E, LeBlanc AR. Design of ion-implanted MOSFET's with very small physical dimensions. IEEE J Solid State Circuits. 1974;9(5):256–68. https://doi.org/10.1109/JSSC.1974.1050511.
2. Tallents G, Wagenaars E, Pert G. Lithography at EUV wavelengths. Nat Photonics. 2010;4(12):809–11.
3. Fan D, Ekinci Y. Photolithography reaches 6 nm half-pitch using EUV light. In: Extreme Ultraviolet (EUV) Lithography VII: International Society for Optics and Photonics. Bellingham, Washington USA: 2016. p. 97761V.
4. Lee H, Yu L-E, Ryu S-W, Han J-W, Jeon K, Jang D-Y, et al. Sub-5nm all-around gate FinFET for ultimate scaling. In: 2006 Symposium on VLSI technology, 2006 digest of technical papers; 2006. p. 58–9.
5. Shrestha A, Mahmood A. Review of deep learning algorithms and architectures. IEEE Access. 2019;7:53040–65. https://doi.org/10.1109/ACCESS.2019.2912200.
6. Shoeybi M, Patwary M, Puri R, LeGresley P, Casper J, Catanzaro B. Megatron-LM: Training multi-billion parameter language models using model parallelism. ArXiv190908053 Cs. 2020;
7. AI and Compute. OpenAI. 2018. https://openai.com/blog/ai-and-compute/: online.
8. Hochreiter S, Schmidhuber J. Long short-term memory. Neural Comput. 1997;9(8):1735–80.
9. Gers FA, Schmidhuber J, Cummins F. Learning to forget: continual prediction with LSTM; 1999. p. 850–5.
10. Lecun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition. Proc IEEE. 1998;86(11):2278–324. Accessed 23 Mar 2021.
11. Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. Commun ACM. 2017;60(6):84–90. https://doi.org/10.1145/3065386.
12. International Technology Roadmap for Semiconductors. 2011. http://www.itrs.net: online.
13. Franklin AD. The road to carbon nanotube transistors. Nature. 2013;498(7455):443–4.
14. Hutchby JA, Bourianoff GI, Zhirnov VV, Brewer JE. Extending the road beyond CMOS. IEEE Circuits Devices Mag. Washington, D.C: 2002;18(2):28–41.
15. Nikonov DE, Young IA. Overview of beyond-CMOS devices and a uniform methodology for their benchmarking. Proc IEEE. 2013;101(12):2498–533.
16. Chen A. Beyond-CMOS technology roadmap: The ConFab; 2015.
17. Ahopelto J, Ardila G, Baldi L, Balestra F, Belot D, Fagas G, et al. NanoElectronics roadmap for Europe: from nanodevices and innovative materials to system integration. Solid State Electron. 2019;155:7–19.
18. Roy K, Chakraborty I, Ali M, Ankit A, Agrawal A. In-memory computing in emerging memory technologies for machine learning: an overview. In: 2020 57th ACM/IEEE Design Automation Conference (DAC); 2020. p. 1–6.
19. Ankit A, Hajj IE, Chalamalasetti SR, Ndu G, Foltin M, Williams RS, et al. PUMA: a programmable ultra-efficient memristor-based accelerator for machine learning inference. In: Proceedings of the twenty-fourth international conference on architectural support for programming languages and operating systems. New York: Association for Computing Machinery; 2019. p. 715–31. (ASPLOS '19).
20. Wong H-SP, Raoux S, Kim S, Liang J, Reifenberg JP, Rajendran B, et al. Phase change memory. Proc IEEE. 2010;98(12):2201–27.
21. Wong H-SP, Lee H-Y, Yu S, Chen Y-S, Wu Y, Chen P-S, et al. Metal–Oxide RRAM. Proc IEEE. 2012;100(6):1951–70.
22. Ambs P. Optical computing: a 60-year adventure. Adv Opt Technol. 2010;2010:1–15.
23. Vander Lugt A. A review of optical data-processing techniques. Opt Acta Int J Opt. 1968;15(1):1–33.

24. Goodman JW, Dias AR, Woody LM. Fully parallel, high-speed incoherent optical method for performing discrete Fourier transforms. Opt Lett. 1978;2(1):1–3.
25. Casasent D. Coherent optical pattern recognition: a review. Opt Eng. 1985;24(1):240126.
26. McCall S, Gibbs H, Venkatesan T. Optical transistor and bistability. J Opt Soc Am 1917–1983. 1975;65:1184.
27. Jain K, Pratt GW Jr. Optical transistor. Appl Phys Lett. 1976;28(12):719–21. https://doi.org/10.1063/1.88627.
28. Athale RA, Lee SH. Development of an optical parallel logic device and a half-adder circuit for digital optical processing. Opt Eng. 1979;18(5):185513.
29. Jenkins BK, Sawchuk AA, Strand TC, Forchheimer R, Soffer BH. Sequential optical logic implementation. Appl Opt. 1984; 23(19):3455–64.
30. Tanida J, Ichioka Y. Optical-logic-array processor using shadowgrams. III. Parallel neighborhood operations and an architecture of an optical digital-computing system. JOSA A. 1985;2(8):1245–53. https://doi.org/10.1364/JOSAA.2.001245.
31. Tanida J, Ichioka Y. OPALS: optical parallel array logic system. Appl Opt. 1986;25(10):1565–70. https://doi.org/10.1364/AO.25.001565.
32. Awwal AAS, Karim MA. Polarization-encoded optical shadow-casting: direct implementation of a carry-free adder. Appl Opt. 1989;28(4):785–90. https://doi.org/10.1364/AO.28.000785.
33. Main T, Feuerstein RJ, Jordan HF, Heuring VP, Feehrer J, Love CE. Implementation of a general-purpose stored-program digital optical computer. Appl Opt. 1994;33(8):1619–28. Accessed 23 Mar 2021.
34. Miller DAB. Are optical transistors the logical next step? Nat Photonics. 2010;4(1):3–5.
35. Tamir DE, Shaked NT, Wilson PJ, Dolev S. High-speed and low-power electro-optical DSP coprocessor. JOSA A. 2009; 26(8):A11–20. https://doi.org/10.1364/JOSAA.26.000A11.
36. Zhu W, Zhang L, Lu Y, Zhou P, Yang L. Design and experimental verification for optical module of optical vector-matrix multiplier. Appl Opt. Washington, D.C: 2013;52(18):4412–8. https://doi.org/10.1364/AO.52.004412.
37. Miller DAB. Self-configuring universal linear optical component [invited]. Photonics Res. 2013;1(1):1–15. https://doi.org/10.1364/PRJ.1.000001.
38. Shen Y, Skirlo S, Harris NC, Englund D, Soljačić M. On-chip optical neuromorphic computing. In: Conference on lasers and electro-optics (2016), paper SM3E2: Optical Society of America; 2016. p. SM3E.2.
39. Shen Y, Harris NC, Skirlo S, Prabhu M, Baehr-Jones T, Hochberg M, et al. Deep learning with coherent nanophotonic circuits. Nat Photonics. 2017;11(7):441–6.
40. Lightmatter. Lightmatter. Washington, D.C. https://lightmatter.co/: online.
41. Lightelligence - Empower AI with light. Lightelligence - Empower AI with light. https://www.lightelligence.ai: online.
42. Ramey C. Silicon photonics for artificial intelligence acceleration: HotChips 32. In: 2020 IEEE hot chips 32 symposium (HCS): IEEE Computer Society; 2020. p. 1–26.
43. Zhou J, Kim K, Lu W. Crossbar RRAM arrays: selector device requirements during read operation. IEEE Trans Electron Devices. 2014;61(5):1369–76.
44. Yang L, Ji R, Zhang L, Ding J, Xu Q. On-chip CMOS-compatible optical signal processor. Opt Express. 2012;20(12):13560–5. https://doi.org/10.1364/OE.20.013560.
45. Tait AN, de Lima TF, Zhou E, Wu AX, Nahmias MA, Shastri BJ, et al. Neuromorphic photonic networks using silicon photonic weight banks. Sci Rep. 2017;7(1):1–10.
46. Chakraborty I, Saha G, Sengupta A, Roy K. Toward fast neural computing using all-photonic phase change spiking neurons. Sci Rep. 2018;8(1):12980. https://doi.org/10.1038/s41598-018-31365-x.
47. Feldmann J, Youngblood N, Wright CD, Bhaskaran H, Pernice WHP. All-optical spiking neurosynaptic networks with self-learning capabilities. Nature. 2019;569(7755):208–14.
48. Feldmann J, Youngblood N, Karpov M, Gehring H, Li X, Stappers M, et al. Parallel convolutional processing using an integrated photonic tensor core. Nature. 2021;589(7840):52–8. https://doi.org/10.1038/s41586-020-03070-1.
49. Ríos C, Youngblood N, Cheng Z, Gallo ML, Pernice WHP, Wright CD, et al. In-memory computing on a photonic platform. Sci Adv. 2019;5(2):eaau5759.
50. Wu C, Yu H, Lee S, Peng R, Takeuchi I, Li M. Programmable phase-change metasurfaces on waveguides for multimode photonic convolutional neural network. Nat Commun. 2021;12(1):96.
51. Chang J, Sitzmann V, Dun X, Heidrich W, Wetzstein G. Hybrid optical-electronic convolutional neural networks with optimized diffractive optics for image classification. Sci Rep. 2018;8(1):1–10.
52. Miscuglio M, Hu Z, Li S, George JK, Capanna R, Dalir H, et al. Massively parallel amplitude-only Fourier neural network. Optica. 2020;7(12):1812–9.
53. Wu Y, Zhuang Z, Deng L, Liu Y, Xue Q, Ghassemlooy Z. Arbitrary multi-way parallel mathematical operations based on planar discrete metamaterials. Plasmonics. 2018;13(2):599–607. https://doi.org/10.1007/s11468-017-0550-0.
54. Liao K, Gan T, Hu X, Gong Q. AI-assisted on-chip nanophotonic convolver based on silicon metasurface. Nanophotonics. 2020;9(10):3315–22. https://doi.org/10.1515/nanoph-2020-0069.
55. George JK, Nejadriahi H, Sorger VJ. Towards on-chip optical FFTs for convolutional neural networks. In: 2017 IEEE International Conference on Rebooting Computing (ICRC); 2017. p. 1–4.
56. Park Y, Azaña J. Optical signal processors based on a time-spectrum convolution. Opt Lett. 2010;35(6):796–8.
57. Zhang X, Huo T, Wang C, Liao W, Chen T, Ai S, et al. Optical computing for optical coherence tomography. Sci Rep. 2016;6:37286.
58. Babashah H, Kavehvash Z, Khavasi A, Koohi S. Temporal analog optical computing using an on-chip fully reconfigurable photonic signal processor. Opt Laser Technol. 2019;111:66–74.
59. Huang Y, Zhang W, Yang F, Du J, He Z. Programmable matrix operation with reconfigurable time-wavelength plane manipulation and dispersed time delay. Opt Express. 2019;27(15):20456–67. https://doi.org/10.1364/OE.27.020456.
60. Xu X, Tan M, Corcoran B, Wu J, Boes A, Nguyen TG, et al. 11 TOPS photonic convolutional accelerator for optical neural networks. Nature. 2021;589(7840):44–51.
61. Lin X, Rivenson Y, Yardimci NT, Veli M, Luo Y, Jarrahi M, et al. All-optical machine learning using diffractive deep neural networks. Science. 2018;361(6406):1004–8.
62. Li J, Mengu D, Luo Y, Rivenson Y, Ozcan A. Class-specific differential detection in diffractive optical neural networks improves inference accuracy. Adv Photonics. 2019;1(4):046001.

63. Mengu D, Luo Y, Rivenson Y, Ozcan A. Analysis of diffractive optical neural networks and their integration with electronic neural networks. IEEE J Sel Top Quantum Electron. 2020;15(1):1–14.

64. Yan T, Wu J, Zhou T, Xie H, Xu F, Fan J, et al. Fourier-space diffractive deep neural network. Phys Rev Lett. 2019;123(2): 023901. https://doi.org/10.1103/PhysRevLett.123.023901.

65. Zhou T, Lin X, Wu J, Chen Y, Xie H, Li Y, et al. Large-scale neuromorphic optoelectronic computing with a reconfigurable diffractive processing unit. Nat Photonics. 2021:1–7.

66. Maass W, Natschläger T, Markram H. Real-time computing without stable states: a new framework for neural computation based on perturbations. Neural Comput. 2002;14(11):2531–60.

67. Jaeger H, Haas H. Harnessing nonlinearity: Predicting chaotic systems and saving energy in wireless communication. Science. 2004;304(5667):78–80. https://doi.org/10.1126/science.1091277.

68. Verstraeten D, Schrauwen B, d'Haene M, Stroobandt D. An experimental unification of reservoir computing methods. Neural Netw. 2007;20(3):391–403. https://doi.org/10.1016/j.neunet.2007.04.003.

69. Rodan A, Tino P. Minimum complexity echo state network. IEEE Trans Neural Netw. Piscataway, NJ USA: 2011;22(1):131–44.

70. Rodan A, Tiňo P. Simple deterministically constructed cycle reservoirs with regular jumps. Neural Comput. 2012;24(7): 1822–52.

71. Bacciu D, Bongiorno A. Concentric ESN: assessing the effect of modularity in cycle reservoirs. In: 2018 International Joint Conference on Neural Networks (IJCNN): IEEE; 2018. p. 1–8.

72. Vandoorne K, Mechet P, Van Vaerenbergh T, Fiers M, Morthier G, Verstraeten D, et al. Experimental demonstration of reservoir computing on a silicon photonics chip. Nat Commun. 2014;5(1):1–6.

73. Tanaka G, Yamane T, Héroux JB, Nakane R, Kanazawa N, Takeda S, et al. Recent advances in physical reservoir computing: a review. Neural Netw. 2019;115:100–23.

74. Vlachas PR, Pathak J, Hunt BR, Sapsis TP, Girvan M, Ott E, et al. Backpropagation algorithms and reservoir computing in recurrent neural networks for the forecasting of complex spatiotemporal dynamics. Neural Netw. 2020;126:191–217.

75. Antonik P, Duport F, Hermans M, Smerieri A, Haelterman M, Massar S. Online training of an opto-electronic reservoir computer applied to real-time channel equalization. IEEE Trans Neural Netw Learn Syst. 2017;28(11):2686–98. https://doi.org/10.1109/TNNLS.2016.2598655.

76. Skibinsky-Gitlin ES, Alomar ML, Frasser CF, Canals V, Isern E, Roca M, et al. Cyclic Reservoir Computing with FPGA Devices for Efficient Channel Equalization. In: Rutkowski L, Scherer R, Korytkowski M, Pedrycz W, Tadeusiewicz R, Zurada JM, editors. Artificial intelligence and soft computing. Cham: Springer International Publishing; 2018. p. 226–34. (Lecture Notes in Computer Science).

77. Katumba A, Yin X, Dambre J, Bienstman P. A neuromorphic silicon photonics nonlinear equalizer for optical communications with intensity modulation and direct detection. J Light Technol. 2019;37(10):2232–9.

78. Argyris A, Bueno J, Fischer I. PAM-4 transmission at 1550 nm using photonic reservoir computing post-processing. IEEE Access. 2019;7:37017–25.

79. Da Ros F, Ranzini SM, Bülow H, Zibar D. Reservoir-computing based equalization with optical pre-processing for short-reach optical transmission. IEEE J Sel Top Quantum Electron. 2020;26(5):1–12. https://doi.org/10.1109/JSTQE.2020.2975607.

80. Li J, Lyu Y, Li X, Wang T, Dong X. Reservoir computing based equalization for radio over fiber system. In: 2021 23rd International Conference on Advanced Communication Technology (ICACT); 2021. p. 85–90.

81. Martinenghi R, Rybalko S, Jacquot M, Chembo YK, Larger L. Photonic nonlinear transient computing with multiple-delay wavelength dynamics. Phys Rev Lett. 2012;108(24):244101.

82. Deihimi A, Orang O, Showkati H. Short-term electric load and temperature forecasting using wavelet echo state networks with neural reconstruction. Energy. 2013;57:382–401. https://doi.org/10.1016/j.energy.2013.06.007.

83. Abreu Araujo F, Riou M, Torrejon J, Tsunegi S, Querlioz D, Yakushiji K, et al. Role of non-linear data processing on speech recognition task in the framework of reservoir computing. Sci Rep. 2020;10(1):328. https://doi.org/10.1038/s41598-019-56991-x.

84. Pathak J, Hunt B, Girvan M, Lu Z, Ott E. Model-free prediction of large spatiotemporally chaotic systems from data: a reservoir computing approach. Phys Rev Lett. 2018;120(2):024102. https://doi.org/10.1103/PhysRevLett.120.024102.

85. Zhou H, Huang J, Lu F, Thiyagalingam J, Kirubarajan T. Echo state kernel recursive least squares algorithm for machine condition prediction. Mech Syst Signal Process. 2018;111:68–86.

86. Griffith A, Pomerance A, Gauthier DJ. Forecasting chaotic systems with very low connectivity reservoir computers. Chaos Interdiscip J Nonlinear Sci. 2019;29(12):123108.

87. Antonik P, Marsal N, Brunner D, Rontani D. Human action recognition with a large-scale brain-inspired photonic computer. Nat Mach Intell. Manhattan, New York: 2019;1(11):530–7.

88. Arcomano T, Szunyogh I, Pathak J, Wikner A, Hunt BR, Ott E. A machine learning-based global atmospheric forecast model. Geophys Res Lett. 2020;47(9):e2020GL087776.

89. Fourati R, Ammar B, Sanchez-Medina J, Alimi AM. Unsupervised learning in reservoir computing for eeg-based emotion recognition. IEEE Trans Affect Comput. 2020.

90. Del Ser J, Lana I, Manibardo EL, Oregi I, Osaba E, Lobo JL, et al. Deep echo state networks for short-term traffic forecasting: Performance comparison and statistical assessment. In: 2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC): IEEE; 2020. p. 1–6.

91. Zhou Z, Liu L, Chandrasekhar V, Zhang J, Yi Y. Deep reservoir computing meets 5G MIMO-OFDM systems in symbol detection. In: Proceedings of the AAAI Conference on Artificial Intelligence; 2020. p. 1266–73.

92. Gallicchio C, Micheli A, Pedrelli L. Deep reservoir computing: a critical experimental analysis. Neurocomputing. 2017;268: 87–99.

93. Sun W, Su Y, Wu X, Wu X, Zhang Y. EEG denoising through a wide and deep echo state network optimized by UPSO algorithm. Appl Soft Comput. 2021;105:107149.

94. Xue Y, Yang L, Haykin S. Decoupled echo state networks with lateral inhibition. Neural Netw. 2007;20(3):365–76. https://doi.org/10.1016/j.neunet.2007.04.014.

95. der Sande GV, Brunner D, Soriano MC. Advances in photonic reservoir computing. Nanophotonics. 2017;6(3):561–76.

96. Gallicchio C, Micheli A, Pedrelli L. Design of deep echo state networks. Neural Netw. 2018;108:33–47.
97. Gallicchio C, Micheli A. Richness of deep echo state network dynamics. In: Rojas I, Joya G, Catala A, editors. Advances in computational intelligence. Cham: Springer International Publishing; 2019. p. 480–91. (Lecture Notes in Computer Science).
98. Gallicchio C, Micheli A. Deep echo state network (DeepESN): a brief survey. ArXiv171204323 Cs Stat. 2020;
99. Dale M, O'Keefe S, Sebald A, Stepney S, Trefzer MA. Reservoir computing quality: connectivity and topology. Nat Comput. 2021;20(2):205–16.
100. Vandoorne K, Dierckx W, Schrauwen B, Verstraeten D, Baets R, Bienstman P, et al. Toward optical signal processing using photonic reservoir computing. Opt Express. 2008;16(15):11182–92. https://doi.org/10.1364/OE.16.011182.
101. Bauduin M, Massar S, Horlin F. Non-linear satellite channel equalization based on a low complexity Echo State Network. In: 2016 Annual Conference on Information Science and Systems (CISS); 2016. p. 99–104.
102. Vandoorne K, Dambre J, Verstraeten D, Schrauwen B, Bienstman P. Parallel reservoir computing using optical amplifiers. IEEE Trans Neural Netw. 2011;22(9):1469–81.
103. Salehi MR, Dehyadegari L. Optical signal processing using photonic reservoir computing. J Mod Opt. 2014;61(17): 1442–51.
104. Brunner D, Fischer I. Reconfigurable semiconductor laser networks based on diffractive coupling. Opt Lett. 2015;40(16): 3854–7. https://doi.org/10.1364/OL.40.003854.
105. Bueno J, Maktoobi S, Froehly L, Fischer I, Jacquot M, Larger L, et al. Reinforcement learning in a large-scale photonic recurrent neural network. Optica. 2018;5(6):756–60.
106. Maktoobi S, Froehly L, Andreoli L, Porte X, Jacquot M, Larger L, et al. Diffractive coupling for photonic networks: how big can we go? IEEE J Sel Top Quantum Electron. Piscataway, NJ: 2019;26(1):1–8.
107. Andreoli L, Porte X, Chrétien S, Jacquot M, Larger L, Brunner D. Boolean learning under noise-perturbations in hardware neural networks. Nanophotonics. 2020;9(13):4139–47.
108. Dong J, Gigan S, Krzakala F, Wainrib G. Scaling up Echo-State Networks with multiple light scattering. In: 2018 IEEE Statistical Signal Processing Workshop (SSP): IEEE; 2018. p. 448–52.
109. Popoff SM, Lerosey G, Carminati R, Fink M, Boccara AC, Gigan S. Measuring the transmission matrix in optics: an approach to the study and control of light propagation in disordered media. Phys Rev Lett. 2010;104(10):100601.
110. Popoff SM, Lerosey G, Fink M, Boccara AC, Gigan S. Controlling light through optical disordered media: transmission matrix approach. New J Phys. 2011;13(12):123021.
111. Dong J, Rafayelyan M, Krzakala F, Gigan S. Optical reservoir computing using multiple light scattering for chaotic systems prediction. IEEE J Sel Top Quantum Electron. 2019;26(1):1–12.
112. Rafayelyan M, Dong J, Tan Y, Krzakala F, Gigan S. Large-scale optical reservoir computing for spatiotemporal chaotic systems prediction. Phys Rev X. 2020;10(4):041037. https://doi.org/10.1103/PhysRevX.10.041037.
113. Paudel U, Luengo-Kovac M, Pilawa J, Shaw TJ, Valley GC. Classification of time-domain waveforms using a speckle-based optical reservoir computer. Opt Express. 2020;28(2):1225–37. https://doi.org/10.1364/OE.379264.
114. Brunner D, Penkovsky B, Marquez BA, Jacquot M, Fischer I, Larger L. Tutorial: photonic neural networks in delay systems. J Appl Phys. Bellingham, Washington: 2018;124(15):152004. https://doi.org/10.1063/1.5042342.
115. Larger L, Soriano MC, Brunner D, Appeltant L, Gutiérrez JM, Pesquera L, et al. Photonic information processing beyond Turing: an optoelectronic implementation of reservoir computing. Opt Express. 2012;20(3):3241–9. https://doi.org/10.1364/OE.20.003241.
116. Paquot Y, Dambre J, Schrauwen B, Haelterman M, Massar S. Reservoir computing: a photonic neural network for information processing. In: Nonlinear optics and applications IV: International Society for Optics and Photonics; 2010. p. 77280B.
117. Duport F, Schneider B, Smerieri A, Haelterman M, Massar S. All-optical reservoir computing. Opt Express. 2012;20(20): 22783–95.
118. Chembo YK. Machine learning based on reservoir computing with time-delayed optoelectronic and photonic systems. Chaos Interdiscip J Nonlinear Sci. 2020;30(1):013111.
119. Dejonckheere A, Duport F, Smerieri A, Fang L, Oudar J-L, Haelterman M, et al. All-optical reservoir computer based on saturation of absorption. Opt Express. 2014;22(9):10868–81.
120. Brunner D, Soriano MC, Mirasso CR, Fischer I. Parallel photonic information processing at gigabyte per second data rates using transient states. Nat Commun. 2013;4(1):1–7.
121. Nakayama J, Kanno K, Uchida A. Laser dynamical reservoir computing with consistency: an approach of a chaos mask signal. Opt Express. 2016;24(8):8679–92. https://doi.org/10.1364/OE.24.008679.
122. Bueno J, Brunner D, Soriano MC, Fischer I. Conditions for reservoir computing performance using semiconductor lasers with delayed optical feedback. Opt Express. 2017;25(3):2401–12. https://doi.org/10.1364/OE.25.002401.
123. Vatin J, Rontani D, Sciamanna M. Enhanced performance of a reservoir computer using polarization dynamics in VCSELs. Opt Lett. 2018;43(18):4497–500.
124. Cuevas GD l, Cubitt TS. Simple universal models capture all classical spin physics. Science. 2016;351(6278):1180–3.
125. Lucas A. Ising formulations of many NP problems. Front Phys. 2014;2.
126. Johnson MW, Amin MHS, Gildert S, Lanting T, Hamze F, Dickson N, et al. Quantum annealing with manufactured spins. Nature. 2011;473(7346):194–8. https://doi.org/10.1038/nature10012.
127. Kim K, Chang M-S, Korenblit S, Islam R, Edwards EE, Freericks JK, et al. Quantum simulation of frustrated Ising spins with trapped ions. Nature. 2010;465(7298):590–3. https://doi.org/10.1038/nature09071.
128. Mahboob I, Okamoto H, Yamaguchi H. An electromechanical Ising Hamiltonian. Sci Adv. 2016;2(6):e1600236.
129. Yamaoka M, Yoshimura C, Hayashi M, Okuyama T, Aoki H, Mizuno H. A 20k-spin ising chip to solve combinatorial optimization problems with CMOS annealing. IEEE J Solid State Circuits. 2016;51(1):303–9.
130. Cai F, Kumar S, Van Vaerenbergh T, Liu R, Li C, Yu S, et al. Harnessing intrinsic noise in memristor hopfield neural networks for combinatorial optimization. ArXiv190311194 Cs. 2019;
131. Kalinin KP, Berloff NG. Simulating Ising and n-state planar Potts models and external fields with nonequilibrium condensates. Phys Rev Lett. 2018;121(23):235302. https://doi.org/10.1103/PhysRevLett.121.235302.

132. Wang Z, Marandi A, Wen K, Byer RL, Yamamoto Y. Coherent Ising machine based on degenerate optical parametric oscillators. Phys Rev A. 2013;88(6):063853.
133. Marandi A, Wang Z, Takata K, Byer RL, Yamamoto Y. Network of time-multiplexed optical parametric oscillators as a coherent Ising machine. Nat Photonics. 2014;8(12):937–42. https://doi.org/10.1038/nphoton.2014.249.
134. Takata K, Marandi A, Hamerly R, Haribara Y, Maruo D, Tamate S, et al. A 16-bit coherent Ising machine for one-dimensional ring and cubic graph problems. Sci Rep. 2016;6(1):34089. https://doi.org/10.1038/srep34089.
135. Inagaki T, Haribara Y, Igarashi K, Sonobe T, Tamate S, Honjo T, et al. A coherent Ising machine for 2000-node optimization problems. Science. 2016;354(6312):603–6.
136. McMahon PL, Marandi A, Haribara Y, Hamerly R, Langrock C, Tamate S, et al. A fully programmable 100-spin coherent Ising machine with all-to-all connections. Science. 2016;354(6312):614–7.
137. Inagaki T, Inaba K, Hamerly R, Inoue K, Yamamoto Y, Takesue H. Large-scale Ising spin network based on degenerate optical parametric oscillators. Nat Photonics. 2016;10(6):415–9. https://doi.org/10.1038/nphoton.2016.68.
138. Takesue H, Inagaki T. 10 GHz clock time-multiplexed degenerate optical parametric oscillators for a photonic Ising spin network. Opt Lett. 2016;41(18):4273–6. https://doi.org/10.1364/OL.41.004273.
139. Yamamoto Y, Aihara K, Leleu T, Kawarabayashi K, Kako S, Fejer M, et al. Coherent Ising machines—optical neural networks operating at the quantum limit. Npj Quantum Inf. 2017;3(1):1–15.
140. Takesue H, Inagaki T, Inaba K, Ikuta T, Honjo T. Large-scale coherent ising machine. J Phys Soc Jpn. 2019;88(6):061014. https://doi.org/10.7566/JPSJ.88.061014.
141. Hamerly R, Inagaki T, McMahon PL, Venturelli D, Marandi A, Onodera T, et al. Experimental investigation of performance differences between coherent Ising machines and a quantum annealer. Sci Adv. 2019;5(5):eaau0823.
142. Cen Q, Hao T, Ding H, Guan S, Qin Z, Xu K, et al. Microwave photonic ising machine. ArXiv201100064 Phys. 2020
143. Böhm F, Verschaffelt G, Van der Sande G. A poor man's coherent Ising machine based on opto-electronic feedback systems for solving optimization problems. Nat Commun. 2019;10(1):3538.
144. Babaeian M, Nguyen DT, Demir V, Akbulut M, Blanche P-A, Kaneda Y, et al. A single shot coherent Ising machine based on a network of injection-locked multicore fiber lasers. Nat Commun. 2019;10(1):3516.
145. Pierangeli D, Marcucci G, Conti C. Large-scale photonic Ising machine by spatial light modulation. Phys Rev Lett. 2019; 122(21):213902.
146. Pierangeli D, Pierangeli D, Marcucci G, Marcucci G, Conti C, Conti C. Adiabatic evolution on a spatial-photonic Ising machine. Optica. 2020;7(11):1535–43.
147. Pierangeli D, Marcucci G, Brunner D, Conti C. Noise-enhanced spatial-photonic Ising machine. Nanophotonics. 2020;3:4109–16.
148. Pierangeli D, Rafayelyan M, Conti C, Gigan S. Scalable spin-glass optical simulator. Phys Rev Appl. 2021;15(3):034087. https://doi.org/10.1103/PhysRevApplied.15.034087.
149. Prabhu M, Roques-Carmes C, Roques-Carmes C, Shen Y, Shen Y, Shen Y, et al. Accelerating recurrent Ising machines in photonic integrated circuits. Optica. 2020;7(5):551–8.
150. Roques-Carmes C, Shen Y, Zanoci C, Prabhu M, Atieh F, Jing L, et al. Heuristic recurrent algorithms for photonic Ising machines. Nat Commun. 2020;11(1):249. https://doi.org/10.1038/s41467-019-14096-z.
151. Okawachi Y, Yu M, Jang JK, Ji X, Zhao Y, Kim BY, et al. Demonstration of chip-based coupled degenerate optical parametric oscillators for realizing a nanophotonic spin-glass. Nat Commun. 2020;11(1):4119. https://doi.org/10.1038/s41467-020-17919-6.
152. Okawachi Y, Yu M, Luke K, Carvalho DO, Ramelow S, Farsi A, et al. Dual-pumped degenerate Kerr oscillator in a silicon nitride microresonator. Opt Lett. 2015;40(22):5267–70.
153. Kako S, Leleu T, Inui Y, Khoyratee F, Reifenstein S, Yamamoto Y. Coherent ising machines with error correction feedback. Adv Quantum Technol. 2020;3(11):2000045.
154. Kumar S, Zhang H, Huang Y-P. Large-scale Ising emulation with four body interaction and all-to-all connections. Commun Phys. 2020;3(1):1–9.
155. Takesue H, Inaba K, Inagaki T, Ikuta T, Yamada Y, Honjo T, et al. Simulating Ising spins in external magnetic fields with a network of degenerate optical parametric oscillators. Phys Rev Appl. 2020;13(5):054059. https://doi.org/10.1103/PhysRevApplied.13.054059.
156. Tezak N, Van Vaerenbergh T, Pelc JS, Mendoza GJ, Kielpinski D, Mabuchi H, et al. Integrated coherent Ising machines based on self-phase modulation in microring resonators. IEEE J Sel Top Quantum Electron. 2020;26(1):1–15.
157. Clements WR, Humphreys PC, Metcalf BJ, Kolthammer WS, Walmsley IA. Optimal design for universal multiport interferometers. Optica. 2016;3(12):1460–5.
158. Bell BA, Wang K, Solntsev AS, Neshev DN, Sukhorukov AA, Eggleton BJ. Spectral photonic lattices with complex long-range coupling. Optica. 2017;4(11):1433–6. https://doi.org/10.1364/OPTICA.4.001433.
159. Wang K, Bell BA, Solntsev AS, Neshev DN, Eggleton BJ, Sukhorukov AA. Multidimensional synthetic chiral-tube lattices via nonlinear frequency conversion. Light Sci Appl. 2020;9(1):132.
160. Liu K, Ye CR, Khan S, Sorger VJ. Review and perspective on ultrafast wavelength-size electro-optic modulators. Laser Photonics Rev. 2015;9(2):172–94.
161. Zhou Z, Yin B, Deng Q, Li X, Cui J. Lowering the energy consumption in silicon photonic devices and systems [invited]. Photonics Res. 2015;3(5):B28–46. https://doi.org/10.1364/PRJ.3.000B28.
162. Chaisakul P, Marris-Morini D, Frigerio J, Chrastina D, Rouifed M-S, Cecchi S, et al. Integrated germanium optical interconnects on silicon substrates. Nat Photonics. 2014;8(6):482–8.
163. Webster M, Gothoskar P, Patel V, Piede D, Anderson S, Tummidi R, et al. An efficient MOS-capacitor based silicon modulator and CMOS drivers for optical transmitters. In: 11th International Conference on Group IV Photonics (GFP); 2014. p. 1–2.
164. Xuan Z, Ma Y, Liu Y, Ding R, Li Y, Ophir N, et al. Silicon microring modulator for 40 Gb/s NRZ-OOK metro networks in O-band. Opt Express. 2014;22(23):28284–91.
165. Dubé-Demers R, LaRochelle S, Shi W. Ultrafast pulse-amplitude modulation with a femtojoule silicon photonic modulator. Optica. 2016;3(6):622–7.
166. Chaisakul P, Vakarin V, Frigerio J, Chrastina D, Isella G, Vivien L, et al. Recent progress on Ge/SiGe quantum well optical modulators, detectors, and emitters for optical interconnects. Photonics. 2019;6(1):24.

167. Romanova A, Barzdenas V. A review of modern CMOS transimpedance amplifiers for OTDR applications. Electronics. 2019;8(10):1073.
168. Kobayashi KW. State-of-the-art 60 GHz, 3.6 k-Ohm transimpedance amplifier for 40 Gb/s and beyond. In: IEEE Radio Frequency Integrated Circuits (RFIC) Symposium, 2003: IEEE; 2003. p. 55–8. Accessed 8 May 2021.
169. Data Converters | Overview |TI.com. https://www.ti.com/data-converters/overview.html: online. Accessed 8 May 2021
170. High Speed A/D Converters >10 MSPS | Analog Devices. https://www.analog.com/en/products/analog-to-digital-converters/high-speed-ad-10msps.html: online.
171. Juanda FNU, Shu W, Chang JS. A 10-GS/s 4-bit single-core digital-to-analog converter for cognitive ultrawidebands. IEEE Trans Circuits Syst II Express Briefs. 2017;64(1):16–20.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.